Nanomedicine, Volume I: Basic Capabilities
Nanomedicine, Volume I:
Basic Capabilities

Robert A. Freitas Jr.
Research Fellow
Institute for Molecular Manufacturing
Palo Alto, California, U.S.A.
To Nancy, Barbara, and Bob

sine qua non

## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xvii</td>
</tr>
<tr>
<td>Foreword by K. Eric Drexler</td>
<td>xviii</td>
</tr>
<tr>
<td>Preface and Acknowledgements</td>
<td>xxi</td>
</tr>
</tbody>
</table>

### Chapter 1. The Prospect of Nanomedicine

#### 1.1 A Noble Enterprise

1. The Evolution of Scientific Medicine
   - Prehistoric Medicine
   - Ancient Mesopotamian Medicine
   - Ancient Egyptian Medicine
   - Ancient Greek Medicine
   - Ancient Alexandrian Medicine
   - Ancient Roman Medicine
   - Medicine in the Middle Ages
   - Renaissance and Pre-Modern Medicine

#### 1.2 Current Medical Practice

1.2.1 The Evolution of Scientific Medicine
   - Prehistoric Medicine
   - Ancient Mesopotamian Medicine
   - Ancient Egyptian Medicine
   - Ancient Greek Medicine
   - Ancient Alexandrian Medicine
   - Ancient Roman Medicine
   - Medicine in the Middle Ages
   - Renaissance and Pre-Modern Medicine

1.2.2 Volitional Normative Model of Disease

#### 1.3 The Nanomedical Perspective

1.3.1 Nanomedicine and Molecular Nanotechnology
1.3.2 Nanomedicine: History of the Idea
   - The Biological Tradition
   - The Mechanical Tradition

### Chapter 2. Pathways to Molecular Manufacturing

#### 2.1 Is Molecular Manufacturing Possible?

#### 2.2 Top-Down Approaches to Nanotechnology

#### 2.3 Bottom-Up Pathways to Molecular Manufacturing

2.3.1 Biotechnology
2.3.2 Molecular and Supramolecular Chemistry
2.3.3 Scanning Probe Technology

#### 2.4 Molecular Components and Molecular Assemblers

2.4.1 Molecular Mechanical Components
2.4.2 Molecular Assemblers

### Chapter 3. Molecular Transport and Sortation

3.1 Human Body Chemical Composition

3.2 Diffusion Transport

3.2.1 Brownian Motion
3.2.2 Passive Diffusive Intake
3.2.3 Active Diffusive Intake
   - Diffusive Stirring
   - Diffusive Swimming
3.2.4 Diffusion Cascade Sortation
3.2.5 Nanocentrifuge Sortation

3.3 Membrane Filtration

3.3.1 Simple Nanosieving
3.3.2 Dynamic Pore Sizing
3.3.3 Gated Channels

3.4 Receptor-Based Transport

3.4.1 Transporter Pumps
3.4.2 Sorting Rotors
3.4.3 Internal Transport Streams

3.5 Molecular Receptor Engineering

3.5.1 Physical Forces in Molecular Recognition
3.5.2 Ligand-Receptor Affinity
3.5.3 Ligand-Receptor Specificity
3.5.4 Ligand-Receptor Dynamics
3.5.5 Diamondoid Receptor Design
3.5.6 Minimum Feature Size and Positioning Accuracy
Chapter 4. Nanosensors and Nanoscale Scanning 93

4.1 Nanosensor Technology 93

4.2 Chemical and Molecular Nanosensors 93

4.2.1 Broadband Receptor Arrays 93

4.2.2 Narrowband Receptor Arrays 94

4.2.3 Counting Rotors 94

4.2.4 Chemical Assay 95

4.2.5 Chemical Nanosensor Theoretical Limits 95

4.2.6 Spatial Concentration Gradients 96

4.2.7 Temporal Concentration Gradients 96

4.2.8 Chemotactic Sensor Pads 96

4.2.9 Receptor Sensors 96

4.3 Displacement and Motion Sensors 97

4.3.1 Displacement Sensors 97

4.3.2 Velocity and Flow Rate Sensors 97

4.3.3 Acceleration Sensors 98

4.3.3.1 Box-Spring Accelerometers 98

4.3.3.2 Displacement Accelerometers 99

4.3.3.3 Fluid Acceleration Sensors 99

4.3.3.4 Pivot Gyrosopic Accelerometers 99

4.3.3.5 Accelerative Onset 99

4.3.4 Angular Displacement 100

4.3.4.1 Gimballed Nanogyrosopes 100

4.3.4.2 Nanopendulum Orientation Sensing 100

4.3.4.3 Nanopendulum Tachometry 102

4.4 Force Nanosensors 102

4.4.1 Minimum Detectable Force 102

4.4.2 Nanogravimeters 102

4.4.3 Single-Proton Massometer 103

4.4.4 Isotope Discrimination 104

4.5 Pressure Sensing 105

4.5.1 Minimum Detectable Pressure 105

4.5.2 Spatial Pressure Gradients 106

4.5.3 Temporal Pressure Gradients 106

4.5.4 Ullage Sensors 106

4.6 Thermal Nanosensors 107

4.6.1 Minimum Detectable Temperature Change 107

4.6.2 Piston-Based Temperature Sensors 107

4.6.3 Thermal-Expansion Temperature Sensors 108

4.6.4 Mechanocalorimetric Temperature Sensors 108

4.6.5 Spatial Thermal Gradients 108

4.6.6 Temporal Thermal Gradients 108

4.7 Electric and Magnetic Sensing 109

4.7.1 Electric Fields 109

4.7.2 Magnetic Fields 110

4.7.3 Optical Sensing 111

4.7.4 Particulate and High-Energy Radiation 111

4.8 Cellular Bioscanning 111

4.8.1 Cellular Topographies 112

4.8.2 Transcellular Acoustic Microscopy 112

4.8.3 Magnetic Resonance Cytotomography 113

4.8.4 Near-Field Optical Nanoimaging 114

4.8.5 Cell Volume Sensing 114

4.8.6 Noninvasive Neuroelectric Monitoring 115

4.8.6.1 Electric Field Neurosensing 115

4.8.6.2 Magnetic Field Neurosensing 115

4.8.6.3 Neurothermal Sensing 115

4.8.6.4 Direct Synaptic Monitoring 115

4.8.6.5 Other Neurosensing Techniques 115

4.8.7 Cellular RF and Microwave Oscillations 116

4.9 Macroscanning 116

4.9.1 Acoustic Macroscanning 116

4.9.1.1 Cytos-Auscultation 116

4.9.1.2 Blood Pressure and Pulse Detection 116

4.9.1.3 Respiratory Audition 117

4.9.1.4 Mechanical Body Noises 117

4.9.1.5 Vocalizations 117

4.9.1.6 Environmental Sources 118

4.9.2 Proprioceptive Macroscanning 118

4.9.2.1 Kinesthetic Macrosensing 118

4.9.2.2 Orientational Macrosensing 119

4.9.2.3 Body Weight Measurement 119

4.9.2.4 Gravitational Geographic Macrosensing 119

4.9.3 Electric/Magnetic Macrosensing 119

4.9.3.1 Vascular-Interstitial Closed Electric Circuits 119

4.9.3.2 Electric/Magnetic Geographic Macrosensing 120

4.9.3.3 Piezoelectric Stress Macrosensing 120

4.9.4 Optical Macroscanning 120

4.9.5 Neural Macroscanning 121

4.9.6 Other Macroscanning 121

Chapter 5. Shapes and Metamorphic Surfaces 123

5.1 Flexible Form and Function 123

5.2 Optimum Nanorobot Shape 123

5.2.1 Free-Floating Solitary Nanodevices 123

5.2.2 Actively Swimming Nanodevices 124

5.2.3 Intracellular Nanodevices 125

5.2.4 Tessellating Nanodevice Aggregates: Nanotissues 125

5.2.4.1 Tiling Nondeforming Surfaces 125

5.2.4.2 Tiling Deforming Surfaces 126

5.2.5 Space-Filling Nanodevice Aggregates: Nano-Organs 127

5.3 Metamorphic Surfaces 128
Basic Capabilities • Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.3.1 General Design Considerations</td>
<td>129</td>
</tr>
<tr>
<td>5.3.1.1 Dimple Size</td>
<td>129</td>
</tr>
<tr>
<td>5.3.1.2 Keyhole Passage</td>
<td>129</td>
</tr>
<tr>
<td>5.3.1.3 Extensibility</td>
<td>129</td>
</tr>
<tr>
<td>5.3.1.4 Reactivity</td>
<td>129</td>
</tr>
<tr>
<td>5.3.2 Metamorphic Surface Configurations</td>
<td>130</td>
</tr>
<tr>
<td>5.3.2.1 Accordion Model</td>
<td>130</td>
</tr>
<tr>
<td>5.3.2.2 Parasol Model</td>
<td>130</td>
</tr>
<tr>
<td>5.3.2.3 Telescoping Model</td>
<td>132</td>
</tr>
<tr>
<td>5.3.2.4 Flexible Fabric Model</td>
<td>132</td>
</tr>
<tr>
<td>5.3.2.5 Block Exchange Model</td>
<td>133</td>
</tr>
<tr>
<td>5.3.3 Metamorphic Power and Control</td>
<td>133</td>
</tr>
<tr>
<td>5.3.4 Engulf Formations</td>
<td>134</td>
</tr>
<tr>
<td>5.3.5 Reconfiguring Surface-Penetrating Elements</td>
<td>134</td>
</tr>
<tr>
<td>5.3.6 Presentation Semaphores</td>
<td>135</td>
</tr>
<tr>
<td>5.3.7 Chromatic Modification</td>
<td>135</td>
</tr>
<tr>
<td>5.4 Metamorphic Bumpers</td>
<td>136</td>
</tr>
<tr>
<td>5.4.1 Nanojunction Mechanisms</td>
<td>136</td>
</tr>
<tr>
<td>5.4.2 Transbumper Communication</td>
<td>137</td>
</tr>
<tr>
<td>5.4.3 Bumper Mechanics</td>
<td>137</td>
</tr>
<tr>
<td>6.1 Nanodevice Energy Resources</td>
<td>139</td>
</tr>
<tr>
<td>6.2 Energy Storage</td>
<td>139</td>
</tr>
<tr>
<td>6.2.1 Gravitational Energy Storage</td>
<td>139</td>
</tr>
<tr>
<td>6.2.2 Mechanical Energy Storage</td>
<td>139</td>
</tr>
<tr>
<td>6.2.2.1 Pendulums and Springs</td>
<td>139</td>
</tr>
<tr>
<td>6.2.2.2 Flywheels</td>
<td>139</td>
</tr>
<tr>
<td>6.2.2.3 Pressurized Fluids</td>
<td>140</td>
</tr>
<tr>
<td>6.2.3 Chemical Energy Storage</td>
<td>140</td>
</tr>
<tr>
<td>6.2.4 Electric and Magnetic Energy Storage</td>
<td>141</td>
</tr>
<tr>
<td>6.2.5 Nuclear Energy Storage</td>
<td>142</td>
</tr>
<tr>
<td>6.3 Power Conversion</td>
<td>142</td>
</tr>
<tr>
<td>6.3.1 Thermal Energy Conversion Processes</td>
<td>142</td>
</tr>
<tr>
<td>6.3.2 Mechanical Energy Conversion Processes</td>
<td>144</td>
</tr>
<tr>
<td>6.3.3 Acoustic Energy Conversion Processes</td>
<td>145</td>
</tr>
<tr>
<td>6.3.4 Chemical Energy Conversion Processes</td>
<td>146</td>
</tr>
<tr>
<td>6.3.4.1 Human Chemical Energy Resources</td>
<td>146</td>
</tr>
<tr>
<td>6.3.4.2 Biological Chemomechanical Power Conversion</td>
<td>147</td>
</tr>
<tr>
<td>6.3.4.3 Artificial Chemomechanical Power Conversion</td>
<td>148</td>
</tr>
<tr>
<td>6.3.4.4 Glucose Engine</td>
<td>149</td>
</tr>
<tr>
<td>6.3.4.5 Chemoelectic Cells</td>
<td>152</td>
</tr>
<tr>
<td>6.3.5 Electrical Energy Conversion Processes</td>
<td>153</td>
</tr>
<tr>
<td>6.3.6 Photonic Energy Conversion Processes</td>
<td>155</td>
</tr>
<tr>
<td>6.3.7 Nuclear Energy Conversion Processes</td>
<td>156</td>
</tr>
<tr>
<td>6.3.7.1 Radionuclides</td>
<td>156</td>
</tr>
<tr>
<td>6.3.7.2 Nuclear Fusion</td>
<td>159</td>
</tr>
<tr>
<td>6.3.7.3 Exothermal Nuclear Catalysis</td>
<td>159</td>
</tr>
<tr>
<td>6.4 Power Transmission</td>
<td>160</td>
</tr>
<tr>
<td>6.4.1 Acoustic Power Transmission</td>
<td>160</td>
</tr>
<tr>
<td>6.4.2 Inductive and Radiofrequency Power Transmission</td>
<td>162</td>
</tr>
<tr>
<td>6.4.3 Tethered Power Transmission</td>
<td>165</td>
</tr>
<tr>
<td>6.4.3.1 Electrical Tethers</td>
<td>165</td>
</tr>
<tr>
<td>6.4.3.2 Electromagnetic Tethers</td>
<td>166</td>
</tr>
<tr>
<td>6.4.3.3 Hydraulic and Acoustic Tethers</td>
<td>167</td>
</tr>
<tr>
<td>6.4.3.4 Gear Trains and Mechanical Tethers</td>
<td>167</td>
</tr>
<tr>
<td>6.4.3.5 Chemical Tethers</td>
<td>169</td>
</tr>
<tr>
<td>6.4.3.6 Power Tether Configurations In Vivo</td>
<td>169</td>
</tr>
<tr>
<td>6.5 Design Energetics Assessment</td>
<td>170</td>
</tr>
<tr>
<td>6.5.1 Power in Biological Cells</td>
<td>170</td>
</tr>
<tr>
<td>6.5.2 Thermogenic Limits in vivo</td>
<td>170</td>
</tr>
<tr>
<td>6.5.3 Nanorobot Power Scaling</td>
<td>172</td>
</tr>
<tr>
<td>6.5.4 Selection of Principal Power Source</td>
<td>173</td>
</tr>
<tr>
<td>6.5.5 Electrical vs. Mechanical Systems</td>
<td>173</td>
</tr>
<tr>
<td>6.5.6 Power Analysis in Design</td>
<td>174</td>
</tr>
<tr>
<td>6.5.7 Global Hypsithermal Limit</td>
<td>175</td>
</tr>
<tr>
<td>7.1 Nanorobot Communications Requirements</td>
<td>177</td>
</tr>
<tr>
<td>7.2 Communication Modalities</td>
<td>177</td>
</tr>
<tr>
<td>7.2.1 Chemical Broadcast Communication</td>
<td>177</td>
</tr>
<tr>
<td>7.2.1.1 Ideal Messenger Molecule</td>
<td>178</td>
</tr>
<tr>
<td>7.2.1.2 Diffusion-Limited Broadcast Rate</td>
<td>179</td>
</tr>
<tr>
<td>7.2.1.3 Instantaneous Stationary Source in Stationary Medium</td>
<td>179</td>
</tr>
<tr>
<td>7.2.1.4 Instantaneous Stationary Source in Flooded Tube</td>
<td>180</td>
</tr>
<tr>
<td>7.2.1.5 Continuous Stationary Source in Stationary Medium</td>
<td>180</td>
</tr>
<tr>
<td>7.2.1.6 Continuous Mobile Source in Stationary Medium</td>
<td>180</td>
</tr>
<tr>
<td>7.2.1.7 Continuous Stationary Source in Nonstationary Medium</td>
<td>180</td>
</tr>
<tr>
<td>7.2.1.8 Assessment of Chemical Broadcast Messaging</td>
<td>180</td>
</tr>
<tr>
<td>7.2.2 Acoustic Broadcast Communication</td>
<td>181</td>
</tr>
<tr>
<td>7.2.2.1 Acoustic Radiators</td>
<td>181</td>
</tr>
<tr>
<td>7.2.2.2 Free-Tissue Acoustic Channel Capacity</td>
<td>181</td>
</tr>
<tr>
<td>7.2.3 Electromagnetic Broadcast Communication</td>
<td>182</td>
</tr>
<tr>
<td>7.2.4 Nanomechanical Communication</td>
<td>183</td>
</tr>
<tr>
<td>7.2.5 Cable Communication</td>
<td>184</td>
</tr>
<tr>
<td>7.2.5.1 Electrical Cables</td>
<td>184</td>
</tr>
<tr>
<td>7.2.5.2 Infrared and Optical Cables</td>
<td>184</td>
</tr>
<tr>
<td>7.2.5.3 Acoustic Cables and Transmission Lines</td>
<td>184</td>
</tr>
<tr>
<td>7.2.5.4 Mechanical Cables</td>
<td>184</td>
</tr>
<tr>
<td>7.2.5.5 Chemomessenger Cables</td>
<td>185</td>
</tr>
<tr>
<td>7.2.6 Communicytes</td>
<td>185</td>
</tr>
<tr>
<td>7.3 Communication Networks</td>
<td>186</td>
</tr>
<tr>
<td>7.3.1 Fiber Networks</td>
<td>186</td>
</tr>
<tr>
<td>7.3.2 Mobile Networks</td>
<td>187</td>
</tr>
</tbody>
</table>
7.3.3 Networks Assessment 188
7.3.4 Dedicated Communication Organs 188
7.4 Communications Tasks 188
7.4.1 Inmessaging from External Sources 189
7.4.2 Inmessaging from Patient or User 189
7.4.2.1 Mechanical Inmessaging 190
7.4.2.2 Kinesthetic Inmessaging 191
7.4.2.3 Acoustic Inmessaging 191
7.4.2.4 Chemical Inmessaging 191
7.4.2.5 Electromagnetic and Thermal Inmessaging 192
7.4.2.6 Neural Inmessaging 192
7.4.2.7 Macroscale Inmessaging Transducers 192
7.4.3 Intradevice Messaging 192
7.4.4 Interdevice Messaging 193
7.4.5 Biocellular Messaging 193
7.4.5.1 Natural Cellular Communication 193
7.4.5.2 Inmessaging from Cells 194
7.4.5.3 Outmessaging to Cells 194
7.4.5.4 Cell Message Modification 194
7.4.5.5 Inmessaging from Neurons 196
7.4.5.6 Outmessaging to Neurons 196
7.4.6 Outmessaging to Patient or User 197
7.4.6.1 Somesthetic Outmessaging 197
7.4.6.2 Kinesthetic Outmessaging 198
7.4.6.3 Auditory Outmessaging 199
7.4.6.4 Gustatory and Olfactory Outmessaging 200
7.4.6.5 Ocular Outmessaging 201
7.4.6.6 Artificial Symptoms 204
7.4.6.7 Macroscale Outmessaging Transducers 204
7.4.7 Outmessaging to External Receivers 206
7.4.8 Transvenue Outmessaging 207

Chapter 8. Navigation 209
8.1 Navigating the Human Body 209
8.2 Human Somatography 209
8.2.1 Navigational Vasculography 210
8.2.1.1 Arteriovenous Macrocirculation 210
8.2.1.2 Arteriovenous Microcirculation 211
8.2.1.3 Lymphatic System 214
8.2.2 Navigational Bronchography 219
8.2.3 Navigational Alimentography 222
8.2.4 Navigational Osteography 227
8.2.5 Organography and Histonavigation 230
8.3 Positional Navigation 234
8.3.1 Dead Reckoning 234
8.3.2 Cartotaxis 234
8.3.3 Microtransponder Networks 235
8.3.4 Vascular Bifurcation Detection 237
8.3.5 Macrotransponder Networks 237
8.3.6 Dedicated Navigational Organs 238
8.4 Functional Navigation 238
8.4.1 Thermographic Navigation 238
8.4.1.1 Thermography of the Human Body 238

Chapter 9. Manipulation and Locomotion 275
9.1 Nanorobot Dexterity and Mobility 275
9.2 Adhesion and Fluid Transport 275
9.2.1 Van der Waals Adhesion Forces 275
9.2.2 Electrostatic Adhesion Forces 277
9.2.3 Immersive Adhesion Forces 277
9.2.4 Capillarity and Nanoscale Fluid Flow 278
9.2.5 Pipe Flow 280
9.2.6 Effervescence and Crystallization 281
9.2.7 Fluid Pumping and Plumbing 282
9.2.7.1 Pressure Release Pumps 282
9.2.7.2 Positive Displacement Pumps 282
9.2.7.3 Turbomolecular Gas Pumps 283
9.2.7.4 Nonmechanical Pumps 283
9.2.7.5 Fluid Mixing 284
9.2.7.6 Nanopumbing and Fluidic Circuits 284
9.2.7.7 Containerized Flow 285
9.3 Nanomanipulators

9.3.1 Nanoscale Manipulators
9.3.1.1 Biological Cilia
9.3.1.2 Nanosclium Manipulators
9.3.1.3 Pneumatic Manipulators
9.3.1.4 Telescoping Manipulators
9.3.1.5 Stewart Platform Manipulators
9.3.1.6 Metamorphic Manipulators

9.3.2 Nanoscale End-Effectors and Tool Tips

9.3.3 Sensors and Manipulator Control

9.3.4 Manipulator Arrays

9.3.5 Materials Disassembly for Disposal
9.3.5.1 Morcellation and Mincing
9.3.5.2 Disposability Engineering
9.3.5.3 Diamondoid Decomposition
9.3.5.3.1 Grinding
9.3.5.3.2 Cleavage
9.3.5.3.3 Thermal Decomposition
9.3.5.3.4 Chemical and Microbial Decomposition

9.4 In Vivo Locomotion

9.4.1 Rheology of Nanorobot-Rich Biofluids
9.4.1.1 Biofluid Viscosity
9.4.1.2 Viscosity of Whole Blood
9.4.1.3 Radial Distribution of Blood Elements
9.4.1.4 Viscosity of Nanorobot-Rich Blood
9.4.1.5 Bloodstream Velocity Profiles
9.4.1.6 Hematocrit Reduction in Narrow Vessels

9.4.2 Sanguinatation
9.4.2.1 Reynolds Number
9.4.2.2 Rotations and Collisions in sanguo
9.4.2.3 Disturbed Flows, Hydrodynamic Interactions, and Entropic Packing

9.4.3 Cytoambulation
9.4.3.1 Ambulatory Contact Event
9.4.3.2 Cell Plasma Membrane Elasticity
9.4.3.2.1 Plasma Membrane Areal Expansion Elasticity
9.4.3.2.2 Plasma Membrane Shear Elasticity
9.4.3.2.3 Plasma Membrane Bending Elasticity
9.4.3.3 Anchoring and Dislodgement Forces
9.4.3.4 Contact Event Cycling
9.4.3.5 Legged Ambulation
9.4.3.6 Tank-Tread Rolling
9.4.3.7 Amoeboid Locomotion

9.4.4 Histonatation
9.4.4.1 Nanorobot Diapedesis
9.4.4.2 ECM Brachiation
9.4.4.3 Intercellular Passage
9.4.4.4 Nanorobot Conjugation and Partition

9.4.5 Cytopenetration
9.4.5.1 Transmembrane Brachiation
9.4.5.2 Metamorphic Screw Drive
9.4.5.3 Solution Wave Drive
9.4.5.4 Viscous Fusion and Endocytotic Entry
9.4.5.5 Cytosolic Leakage During Transit
9.4.5.6 Breach Sealing and Intrusiveness
9.4.5.7 Nuclear Membrane Penetration

9.4.6 In Cyto Locomotion
9.4.7 Cyrocarrige
9.4.7.1 Objectives of Cyrocarrige
9.4.7.2 Cytovehicle Selection
9.4.7.3 Cyrocarrige Initiation
9.4.7.4 Steering and Control
9.4.7.5 Navigation and Sensing
9.4.7.6 Cytovehicular Behavioral Control

9.5 Ex Vivo Locomotion
9.5.1 Dental Walking
9.5.2 Epidermal Locomotion
9.5.3 Nanoflight
9.5.3.1 Nanoflight and Reynolds Number
9.5.3.2 Nanoflight and Gravity
9.5.3.3 Buoyant Nanoballoons
9.5.3.4 Nanoflyer Force and Power Requirements
9.5.3.5 Hovering Flight
9.5.3.6 No-Fly Zones

9.6 Ex Vivo Locomotion
9.6.1 Nano pady
9.6.2 Nonmechanical Computers
9.6.2.1 Molecular Wires
9.6.2.2 Electromechanical Molecular Switching Devices
9.6.2.3 Field-Controlled Molecular Switching Devices

Chapter 10. Other Capabilities

10.1 Nanochronometry
10.1.1 Human Chronobiology
10.1.2 Artificial Nanoscale Oscillators
10.1.2.1 Mechanometric & Photochemical Oscillators
10.1.2.2 Mechanical Oscillators
10.1.2.3 Acoustic Transmission Line Oscillators
10.1.2.4 Quartz Resonators
10.1.2.5 Atomic Frequency Standards
10.1.3 Nanorobot Synchronization
10.1.4 Dedicated Chronometer Organs

10.2 Nanocomputers
10.2.1 Nanomechanical Computers
10.2.2 Nanoelectronic Computers
10.2.2.1 Molecular Wires
10.2.2.2 Electromechanical Molecular Switching Devices
10.2.2.3 Field-Controlled Molecular Switching Devices
### 1. The Prospect of Nanomedicine

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Evolution of the Paradigm of Scientific Medicine</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Deaths Due to Diphtheria in England and Wales</td>
<td>15</td>
</tr>
<tr>
<td>1.3</td>
<td>Branches of Medicine and Associated Disciplines</td>
<td>17</td>
</tr>
<tr>
<td>1.4</td>
<td>Opinion Survey: What Is &quot;Disease&quot;?</td>
<td>19</td>
</tr>
<tr>
<td>1.5</td>
<td>Three Contemporary Branches of &quot;Nanotechnology&quot;</td>
<td>31</td>
</tr>
</tbody>
</table>

### 2. Pathways to Molecular Manufacturing

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The Nippondenso Microcar—Smaller Than A Grain of Rice</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>The Ribosome Acts as a Programmable Nanoscale Assembler of Protein Nanoproducts</td>
<td>43</td>
</tr>
<tr>
<td>2.3</td>
<td>Synthetic Scheme Used to Build 3-Dimensional DNA Cubes in Solution Phase</td>
<td>46</td>
</tr>
<tr>
<td>2.4</td>
<td>Propellanes</td>
<td>48</td>
</tr>
<tr>
<td>2.5</td>
<td>Rotanes</td>
<td>48</td>
</tr>
<tr>
<td>2.6A</td>
<td>Cyclophane</td>
<td>48</td>
</tr>
<tr>
<td>2.6B</td>
<td>Superphane</td>
<td>48</td>
</tr>
<tr>
<td>2.7</td>
<td>The Prismsanes</td>
<td>49</td>
</tr>
<tr>
<td>2.8</td>
<td>Other Unusual Molecular &quot;Parts&quot;</td>
<td>49</td>
</tr>
<tr>
<td>2.9</td>
<td>Calixarenes</td>
<td>49</td>
</tr>
<tr>
<td>2.10</td>
<td>Ferric Wheel</td>
<td>49</td>
</tr>
<tr>
<td>2.11</td>
<td>“Butterfly Molecule” Used as Molecular Tongs</td>
<td>49</td>
</tr>
<tr>
<td>2.12</td>
<td>“Staffane” Rigid Rods of Different Lengths for a Molecular Construction Kit</td>
<td>51</td>
</tr>
<tr>
<td>2.13</td>
<td>C_{60} Fullerene “Buckyball” and a Soccer Ball</td>
<td>51</td>
</tr>
<tr>
<td>2.14</td>
<td>C_{32} and C_{50} Fullerenes</td>
<td>51</td>
</tr>
<tr>
<td>2.15</td>
<td>C_{20} and C_{50} Fullerenes</td>
<td>52</td>
</tr>
<tr>
<td>2.16</td>
<td>Single-Walled Carbon Nanotubes</td>
<td>52</td>
</tr>
<tr>
<td>2.17</td>
<td>Carbon Nanotubes Kink When Bent</td>
<td>52</td>
</tr>
<tr>
<td>2.18</td>
<td>One Possible Saddle-Shaped Fullerene</td>
<td>53</td>
</tr>
<tr>
<td>2.19</td>
<td>Biological “Fullerenes”</td>
<td>53</td>
</tr>
<tr>
<td>2.20A</td>
<td>Fullerene Dimer</td>
<td>53</td>
</tr>
<tr>
<td>2.20B</td>
<td>Fullerene Polyester Polymer</td>
<td>54</td>
</tr>
<tr>
<td>2.20C</td>
<td>Fullerene Dendrimer</td>
<td>54</td>
</tr>
<tr>
<td>2.20D</td>
<td>Fullerene Rotaxane</td>
<td>54</td>
</tr>
<tr>
<td>2.20E</td>
<td>Fullerene-Nucleotide DNA Cleaving Agent</td>
<td>54</td>
</tr>
<tr>
<td>2.20F</td>
<td>Stable Diels-Alder Fullerene Adduct</td>
<td>54</td>
</tr>
<tr>
<td>2.20G</td>
<td>Extended Fullerene Polymers</td>
<td>54</td>
</tr>
<tr>
<td>2.21</td>
<td>Computer Simulation of Fullerene Nanogears of the Same Size</td>
<td>55</td>
</tr>
<tr>
<td>2.22</td>
<td>Computer Simulation of Fullerene Rack and Pinion System</td>
<td>55</td>
</tr>
<tr>
<td>2.23</td>
<td>Computer Simulation of Fullerene Nanogears of Different Sizes</td>
<td>56</td>
</tr>
<tr>
<td>2.24</td>
<td>Schematic of Scanning Tunneling Microscope (STM)</td>
<td>56</td>
</tr>
<tr>
<td>2.25</td>
<td>IBM Logo Spelled Out Using 35 Xenon Atoms Arranged on a Nickel Surface by an STM</td>
<td>57</td>
</tr>
</tbody>
</table>

### 3. Molecular Transport and Sortation

<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Schematic of Diffusion Cascade Sortation Unit</td>
<td>75</td>
</tr>
<tr>
<td>3.2</td>
<td>Schematic of Teragravity Nanocentrifuge</td>
<td>76</td>
</tr>
<tr>
<td>3.3</td>
<td>Variable Size/Shape Apertures Using Two Nanoscale Perforated Sliding Plates</td>
<td>78</td>
</tr>
<tr>
<td>3.4</td>
<td>Circular Dilating “Iris” Diaphragm Mechanism for Dynamic Pore Sizing</td>
<td>79</td>
</tr>
<tr>
<td>3.5</td>
<td>Schematic of Transporter Molecular Pump Operation (Unisport)</td>
<td>80</td>
</tr>
<tr>
<td>3.6</td>
<td>Schematic of Sodium-Potassium Antiporter Ion Pump Operation</td>
<td>81</td>
</tr>
<tr>
<td>3.7</td>
<td>Molecular Sorting Rotor</td>
<td>81</td>
</tr>
<tr>
<td>3.8</td>
<td>Sorting Rotor Cascade</td>
<td>83</td>
</tr>
<tr>
<td>3.9</td>
<td>Molecular Mills for Internal Transport: Simple Transfer Between Two Reservoirs</td>
<td>83</td>
</tr>
<tr>
<td>3.10</td>
<td>Experimental RMS Fluctuations in Ferrocytochrome C</td>
<td>86</td>
</tr>
<tr>
<td>3.11</td>
<td>Imprint Model for Creating Artificial Molecular Receptors</td>
<td>88</td>
</tr>
<tr>
<td>3.12A</td>
<td>2-D Schematic Representation of a 3-D Solid Mosaic Model Artificial Receptor: Multiform Block Design</td>
<td>89</td>
</tr>
<tr>
<td>3.12B</td>
<td>2-D Schematic Representation of a 3-D Solid Mosaic Model Artificial Receptor: Raster Scan Design</td>
<td>89</td>
</tr>
<tr>
<td>3.13</td>
<td>Schematic Representation of Tomographic Model for a Reconfigurable Artificial Molecular Receptor</td>
<td>89</td>
</tr>
</tbody>
</table>
4. Nanosensors and Nanoscale Scanning
Figure 4.1. Schematic of Broadband Chemical Concentration Sensor Array Using 5 Receptor Units with Steric Probes 94
Figure 4.2. Schematic of Narrowband Chemical Concentration NanoSensor Array Using Receptors of a Single Type 95
Figure 4.3. Chemical Concentration Sensor Using Counting Rotors 95
Figure 4.4. Schematic of Box-Spring Omnidirectional Accelerometer 98
Figure 4.5. Spherical Nanopendulum Sensor for Angular Displacement and Rotational Velocity 101
Figure 4.6. Single-Proton Massometer Sensor Using Coiled Suspension-Spring Cantilever 103
Figure 4.7. Two-Dimensional Representation of a Three-Dimensional Micro-Kelvin Piston-Based Coiled-Cylinder Temperature Sensor 107
Figure 4.8. Thermal Expansion Temperature Sensor 108

5. Shapes and Metaphoric Surfaces
Figure 5.1. Regular Tessellations that Fill A Plane Using One Kind of Polygon 125
Figure 5.2. Semiregular and Nonuniform Periodic (Using More Than One Kind of Polygon) Tessellations 126
Figure 5.3. Stationkeeping on Monoaxially Deforming Surfaces Using Metamorphic Bumpers 126
Figure 5.4. Uniform Space Filling Using One Kind of Polyhedron: Triangular, Square, and Hexagonal Prisms 127
Figure 5.5. Uniform Space Filling Using Only the Truncated Octahedron 127
Figure 5.6. Uniform Space Filling Using Only the Rhombic Dodecahedron 127
Figure 5.7. Uniform Space Filling Using Only the Rhombo-Hexagonal Dodecahedron 128
Figure 5.8. Uniform Space Filling Using Only the Non-Regular Octahedron 128
Figure 5.9. Solid and Folding Geometry of the Regular Octahedron 128
Figure 5.10. Solid Geometry of the Space Filling Trapezohedron 128
Figure 5.11. Accordion Model 130
Figure 5.12. Schematic of Parasol Model 131
Figure 5.13. Schematic of Multiplane Parasol Configuration 132
Figure 5.14. Telescoping Model: Various Configurations 132
Figure 5.15. Flexible Fabric Model: Coiled Pleat Configuration 132
Figure 5.16. Block Exchange Model 133
Figure 5.17. Schematic Representation of a Watertight Block Surface Reconfiguration: Removing an Embedded Sensor Element 134
Figure 5.18. Schematic of Presentation Semaphore Mechanism 135

6. Power
Figure 6.1. Pressure-Driven Actuators for Acoustomechanical Power Transduction 145
Figure 6.2. Bacterial Flagellar Proton Gradient 148
Figure 6.3. Sussmann-Katchalsky Chemomechanical Motor 149
Figure 6.4A. Glucose Engine Combustion Chamber (top cutaway) 150
Figure 6.4B. Glucose Engine Combustion Chamber (side cutaway) 151
Figure 6.4C. Glucose Engine 3-D Electrodynamic Suspension (2-D schematic) 151
Figure 6.5. Schematic of Oxyglucose Biofuel Cell with Proton Exchange Nanomembrane 154
Figure 6.6. Submicron Direct-Current Electrostatic Motor 155
Figure 6.7. Schematic of Semiconductor-Junction Nucleoelectric Transducer 158
Figure 6.8. Safety Zone for Human Exposure to Ultrasound 161
Figure 6.9. Received Acoustic Power vs. Acoustic Frequency for Various Path Lengths in the Human Body 162
Figure 6.10. Received Acoustic Power vs. Acoustic Frequency for Various Receiver Volumes and Reflection Losses 163
Figure 6.11. Received Acoustic Power vs. Acoustic Frequency for Various Absorption Coefficients in Human Tissue 163
Figure 6.12. Occupational Exposure Limits to Electromagnetic Radiation as Equivalent Plane-Wave Power Density 164
Figure 6.13. Estimated Total Attenuation Factor $\alpha$ (Scattering + Absorption) for Electromagnetic Radiation Passing Through Human Tissue 165
Figure 6.14. Recommended Maximum Total Nanomachinery Power Consumption in Isolated Cells and in Human Tissue 172
Figure 6.15. Maximum Temperature Change Produced by Nanorobot Power Released in Isolated Cells and in Human Tissue 172

7. Communication
Figure 7.1. Acoustic Pressure ($P_A$) at the Surface of a Cylindrical Vibrating Piston Acoustic Radiator of Radius $R$ and Input Power $P_m$ for $\Delta P_{min} = 10^{-6}$ atm in Water at 310 K 182
Figure 7.2. Bones of the Right Hand, Palmar Surface 189
Figure 7.3. Cross-Section of the Human Ear 199
Figure 7.4. Cross-Section of the Cochlear Duct 200
Figure 7.5. The Human Eyeball 201
Figure 7.6. The Human Retina 203
Figure 7.7A. A Selection of Possible Dermal Display Screens 205
Figure 7.7B. A Dermal Display Screen in Use 206

8. Navigation
Figure 8.1. Human Arterial System 210
Figure 8.2. Human Venous System 211
Figure 8.3. Vein Valve Architecture 211
Figure 8.4. Metarterioles, Venules, and Lymph Capillaries 213
Figure 8.5. Details of a Lymph Capillary 215
Figure 8.6. Main Trunks of the Human Lymphatic System 216
Figure 8.7. Lymphatic Drainage Regions 216
Figure 8.8. Main Organs of the Human Lymphatic System 217
Figure 8.9. Details of a Lymph Node 217
Figure 8.10. Two Patterns for the Junction of the Lumbar Lymphatic Trunks 218
Figure 8.11. Sagittal Section Through Mouth, Nasal Cavity, Pharynx and Larynx 220
Figure 8.12. Lung Lobes and the Bronchial Tree 220
10. Other Basic Capabilities

Figure 10.1. Schematic of Nanomechanical NAND Gate 344

Figure 10.2. Schematic of Nanomechanical Rod Logic Data Storage Registers 344

Figure 10.3. Schematic of Programmable Logic Array (PLA) Finite State Machine Implementing Rod Logic for a Nanomechanical Central Processing Unit 344

Figure 10.4. Schematic Diagram of a Single Displacement Cycle Aynchronous-Input Nanomechanical OR Gate 345

Figure 10.5. Linear Atomic Relay Switch 346

Figure 10.6. Rotational Molecular Relay Switch 347

Figure 10.7. Hinging Molecular Relay Switch 347

Figure 10.8. Molecular Shuttle Switch 347

Figure 10.9. Two-Terminal Molecular Resonant Tunneling Device (RTD) 347

Figure 10.10. Molecular Electrostatic Logic Device 349

Figure 10.11. Phase Diagram for Water and Ice 354
1. The Prospect of Nanomedicine
   Table 1.1. Average Age of Death of Persons Buried in Roman Times  2
   Table 1.2. Leading Causes of Death in the United States, 1890 vs. 1990  9
   Table 1.3. Comparison of Macroscale and Biomolecular Components and Functions  26
   Table 1.4. Medical Challenges of Increasing Difficulty and Possible Approaches  32

2. Pathways to Molecular Manufacturing
   Table 2.1. Information Required to Describe Some Self-Replicating Systems  65

3. Molecular Transport and Sortation
   Table 3.1. Estimated Atomic Composition of the Lean 70-kg Male Human Body  71
   Table 3.2. Estimated Gross Molecular Contents of a Typical 20-micron Human Cell  71
   Table 3.3. Translational Brownian Diffusion Coefficients for Physiologically Important Molecules Suspended in Water at 310 K  73
   Table 3.4. Estimated Time for Diffusion to Reach 90% Completion for Glycine in Aqueous Solution at 310 K  74
   Table 3.5. Sedimentation Coefficients for Particles in Aqueous Suspension at 310 K  76
   Table 3.6. Maximum Binding Energy Available in a Dispersion-Force Receptor with 0.3-nm Contact Boundary for CHON Target Molecules of Various Sizes  87

4. Nanosensors and Nanoscale Scanning
   Table 4.1. Isotopic Variation in Bulk Properties Possibly Useful in Nanomedicine  105
   Table 4.2. Amplitude Absorption Coefficients for Acoustic Waves in Human Body Tissue  117
   Table 4.3. Acoustic Impedance for Specular Reflection by Acoustic Waves Crossing a Material Interface  118

5. Shapes and Metamorphic Surfaces
   Table 5.1. Geometrical Factors for Bloodstream-Traversing Nanorobots of Maximum Volume with Largest Transdevice Diameter L_n = 4 microns  125

6. Power
   Table 6.1. Energy Storage Density for Various Chemical Fuels  141
   Table 6.2. Energy Conversion Matrix: Power Transduction Technologies That Convert One Form of Energy Into Another  143
   Table 6.3. Estimated Acoustic Energy Available From Physiological Sources Within the Human Body  146
   Table 6.4. Estimated Chemical Energy Resources in the Human Body  147
   Table 6.5. Volumetric Radioactive Power Density of Various Radionuclides  157
   Table 6.6. Acoustic Energy Reflected and Transmitted at Various Tissue Interfaces  163
   Table 6.7. Velocity of Sound in Various Organs and Biological Materials at 310 K  163
   Table 6.8. Estimated Power Output and Measured Power Density of Biological Cells and Human Tissues  171

7. Communication
   Table 7.1. Temporal Signature of A Detectable Mechanical Pulse Conducted Mostly Through Bone, which is Applied to One Finger and Received by Nanorobot Sensors Distally Located in Each of the Five Fingers of the Hand  190
   Table 7.2. The Four Major Classes of Human Cell Secretory Signaling Molecules  195
   Table 7.3. Tactile Receptors in the Human Skin  198

8. Navigation
   Table 8.1. Approximate Quantification of the Human Arteriovenous System  212
   Table 8.2. Additional Data on Selected Blood Vessels  212
   Table 8.3. Branching Structure of the Pulmonary Arterial Network Entering the Lungs  213
   Table 8.4. Typical Blood Perfusion Rates for Various Tissues and Organs  214
   Table 8.5. Approximate Quantification of the Human Lymphatic System  215
   Table 8.6. Chemical Composition of Human Thoracic Lymph  218
   Table 8.7. Approximate Quantification of the Human Bronchial System  219
   Table 8.8. Approximate Quantification of the Human Alimentary System  222
   Table 8.9. Mass, Volume, and Scale Size of the Organs of the Human Body  231
   Table 8.10. Vitamin Concentration in Various Human Organs  232
   Table 8.11. Core Temperatures of Organs and Blood Vessels in the Human Body  239
   Table 8.12. Thermophysical Characteristics of Various Body Tissues, Organs, and Other Materials  242
   Table 8.13. Recognized HLA Specificities  249
Table 8.15. Complete Listing of Recognized Human Platelet Alloantigen (HPA) Systems
Table 8.16. Unique and Shared Antigenic Surface Markers of Hepatopoietic and Hemopoietic Human Cells
Table 8.17. Approximate Quantification of the Components of a Typical 20-micron Human Cell
Table 8.18. Mass % Biochemical Composition of Cell and Organelle Membranes
Table 8.19. Some Species of Stable RNAs and their Localizations in Human Cells

9. Manipulation and Locomotion
Table 9.1. Hamaker Constant for Various Materials
Table 9.2. Henry’s Law Constants
Table 9.3. Mechanical Strength of Nanomedical Materials
Table 9.4. Absolute Viscosity of Some Common Materials

Table 9.5. Conservatively Estimated Force, Power, and Power Density for Hovering and Flying Spherical Nanorobots as a Function of Size and Velocity

10. Other Basic Capabilities
Table 10.1. Van der Waals Equation Gas Constants
Table 10.2. Gas Molecules Packed into a Pressure Vessel: Number Density as a Function of Pressure at 310 K, using the van der Waals Equation
Table 10.3. Biological and Biochemical Effects of High Pressure
Table 10.4. Organic Compounds Affecting Eukaryotic Cell Division at Specific Phases of Mitosis
Table 10.5. Inhibitors of Protein or RNA Synthesis
Table 10.6. Summary of Biochemical Composition and Net Digestion Breakdown Products of a Representative 4-micron³ Bacterium
Table 10.7. Molal Freezing and Boiling Point Constants
Table 10.8. Molar Heats of Fusion and Vaporization
The coming ability to carry out targeted medical procedures at the molecular level will bring unprecedented power to the practice of medicine. Within a few short decades we can expect a major revolution in how the human body is healed. Nanomedicine lays the foundations for understanding this revolution and points to where it is taking us.

Throughout science and technology, the race to obtain complete control of the structure of matter is gaining speed and focus. From chemistry to biotechnology, from applied physics to software, increasing resources are being brought to bear on the goal of nanotechnology. The term nanotechnology is used to describe a variety of nanoscale technologies. More precise for present purposes is molecular nanotechnology: the ability to construct objects with atomic-scale control. In lay terms, this is often called “building atom-by-atom”; more accurately, it means being able to bond every atom into a specific, designed location within a larger structure.

It is easy to see that such an ability could lead to, for example, the construction of stronger, more reliable materials, and smaller, faster computer chips. One can also anticipate improved versions of the molecular machines found in nature, similar to today’s work in redesigning enzymes. More advanced work should eventually enable the building of molecular machine systems -- micron-scale and even macro-scale systems of novel molecular machines, performing complex operations as do the natural molecular machine systems in living things. Such systems could in principle manufacture large numbers of atomically-precise products—a process known as molecular manufacturing.

For those with a medical or biological background, a description of such powerful technological abilities at the molecular scale raises questions regarding the potential for applications in living systems, including the human body. The responsible use of technology demands that these interactions be considered in advance of widespread deployment.

For millennia, physicians and their predecessors have worked to aid the human body in its efforts to heal and repair itself. Slowly at first, and with accelerating speed, new methods and instruments have been added to the physician’s toolkit -- microsurgical techniques for physically removing problematic tissue and reconfiguring healthy tissue, antibiotics for jamming the molecular machinery of unwanted bacteria, gene chips for rapid identification of genetic sequences.

In most cases, however, physicians must chiefly rely on the body’s self-repair capabilities. If these fail, external efforts are hopeless. We cannot today put the component parts of human cells exactly where they should be, and configure them as they should be to form a healthy physiological state. There are no tools for working, precisely and with three-dimensional control, at the molecular level.

Nanomedicine points the way to these advances, which will arm physicians with the most important new tools in medicine since the discovery of antibiotics. The comprehensive development of nanomedicine will dominate medical technology research during the first half of the 21st century, and perhaps beyond.

These are not high-risk predictions, but merely the extension of currently-observable progress in biological and medical research. Advances in understanding living systems in general, and the human body in particular, have arrived at astounding speed over the past three decades and show no signs of slowing down. The completion of the Human Genome Project, once thought almost impossibly ambitious, is now widely regarded as both routine and destined to revolutionize medicine. What was visionary a short time ago is now a minimum baseline expectation.

We can expect this now-familiar pattern to be repeated in the field of nanomedicine. So often a goal, achievable in theory but considered at first to be far too difficult, within a few decades becomes first an active goal, and then an achievement.

Given its revolutionary potential, it is not too soon to examine the goals and potential consequences of nanomedicine. Provided that this work is firmly grounded in today’s science—assuming no new scientific principles, but only the gradual accumulation and application of new data—we can be confident that our calculations will give reasonable, even conservative, results.

Of course, not all physicians will choose to join in examining the future of medicine. This is not only understandable, but quite reasonable for those who must treat patients today, with the methods available today. Nor is the medical researcher, working to improve today’s pharmaceuticals, spurred on by the knowledge that his or her success—no matter how dramatic—will eventually be superseded. In both cases, what can be done today, or next year, is the most appropriate focus.

But only a fraction of today’s physicians and researchers need look ahead for the entire field of medicine to benefit. Those practitioners who plan to continue their careers into the timeframe when nanomedical developments are expected to arrive—e.g., younger physicians and researchers, certainly those now in medical and graduate programs—can incrementally speed the development process, while simultaneously positioning their own work for best effect, if they have a sound idea of where the field of medicine is heading. Those farther along in their careers will be better able to direct research resources today, if the goals of nanomedicine are better understood. Nanomedicine helps us to frame the research issues that must be addressed, and to take better-directed steps on the path toward medical nanotechnology.

Finding the right theoretician to describe the foundations of such a field is difficult. The research requires a high degree of multidisciplinary ability—not the typical result of academic programs which reward specialization. A multidisciplinary group might serve, but would likely fail to provide the necessary conceptual integration.
Rather, the ideal author would be a careful researcher with a broad scientific background—in particular, one willing and able to tackle a daunting task requiring a decade of concentrated full-time research, constructing a technical exposition that may ultimately span thousands of pages and citations. Robert Freitas brings all these qualities and more to this challenging project.

Some aspects of the book will initially seem controversial. For example, the advent of nanomedicine will redefine the very concept of "disease" (Section 1.2.2). Today’s medicine is limited to removing tissue, replacing it with transplants and artificial materials, or helping it repair itself. Current repair techniques require that the tissue be metabolizing and functioning, so inactive or structurally intact but non-functioning tissue is declared "dead." Nanotechnology will let us repair non-functioning tissue, leading us to reexamine the concept of clinical death used in medicine today.

This process of redefinition itself is not new in the field of medicine. It has happened many times before and is central to the goal of extending the frontiers of health into new territory. The advance of the physician’s reach down to control at the nanoscale is just one more step on the long evolutionary pathway of medical history.

Also controversial will be the question of how nanomedical techniques should be used. Science discovers natural laws and facts, and technology extends the limits of what was possible with unaltered natural systems. But science and technology do not speak to the issues of what is morally or ethically correct—they can only help to frame the context of those discussions. Seemingly simple questions such as who is sick and who is well, how much we can do for them, and how much it will cost, require a fundamental understanding of the underlying scientific and technical limits of what is possible. Beyond that, the dialogue must include value judgments, and increasingly—if the economic and environmental costs can be handled responsibly—the answers have been provided by the customer, here, the patient.

There is a growing body of humanitarian, religious, and political work in the field of bioethics. A purpose of the present volume is to help frame the issues with a better understanding of what will be possible technically and economically. Thus, it can provide an essential ingredient for responsible ethical discussions while itself staying within the bounds of scientific and engineering analysis.

Leaving aside the substantive ethical debates to come, the technical case presented in Nanomedicine is sure to spark controversy on its own. Those who disagree with the thrust of this book may find segments which, taken out of context, appear unsupported or overstated. This is almost unavoidable in the writing of any book, much less one as ambitious as Nanomedicine. Yet the ideas in this book are supported, as presented in their intended context, and criticisms of them will be most valuable if they keep this context linked to the discussion.

Both the author and publisher of Nanomedicine have graciously agreed to publish the entire book online, accessible at no charge, to better enable debate on, and evolution of, this work. Published critiques, to serve the advance of knowledge, should include a reference to the online location of this material, enabling readers to probe questions in more depth. Online tools now allow third-party annotation of online texts, enabling all concerned parties to contribute to a reasoned discussion of the technical issues raised here. Robert Freitas and I invite readers of this volume to participate in this discussion, with the goal of furthering the evolution of knowledge in the new and vital field of nanomedicine.
Molecular nanotechnology has been defined as the three-dimensional positional control of molecular structure to create materials and devices to molecular precision. The human body is comprised of molecules, hence the availability of molecular nanotechnology will permit dramatic progress in human medical services. More than just an extension of “molecular medicine,” nanomedicine will employ molecular machine systems to address medical problems, and will use molecular knowledge to maintain and improve human health at the molecular scale. Nanomedicine will have extraordinary and far-reaching implications for the medical profession, for the definition of disease, for the diagnosis and treatment of medical conditions including aging, for our very personal relationships with our own bodies, and ultimately for the improvement and extension of natural human biological structure and function.

Nanomedicine, the book, will be published in three Volumes over the course of several years. The present Volume is the first in this series. Readers wishing to keep abreast of the latest developments may visit the nanomedicine website maintained by the Foresight Institute (http://www.foresight.org/Nanomedicine/index.html) and may visit http://www.nanomedicine.com, the first commercial Internet domain exclusively devoted to nanomedicine and the online home of this document. To date, the author has expended ~19,000 man-hours, plus ~1,000 man-hours by reviewers, a total of ~10 man-years of effort.

To hold the book to a manageable length, most technical discussions have been greatly abbreviated, usually omitting lengthy historical surveys and extensive derivations or proofs of formulas, but providing useful pointers to the relevant specialist literature. Equations are presented, whenever possible, with sufficient qualifications to enable the reader to determine the extent of applicability in a given circumstance; however, complete derivations are rarely provided. Care should be taken in applying these mathematical relationships, since in some cases their applicability to submicron systems, while anticipated, has not yet been definitively established by experiment. With more than 270 mathematical equations in this Volume, it was impossible to maintain strict usage consistency of variables across Chapters, although some effort was made to achieve such consistency within individual Chapters. Clarity is enhanced by defining all variables and constants near or immediately after each usage in an equation. In this book, the symbol “~” signifies “approximately,” and scientific exponential notation appears throughout. Cubic volumes are sometimes specified as, for example, (1 nm)$^3$ to signify a 1 nm$^3$ block. Conversions among various units (which were chosen to maximize interdisciplinary understanding) are summarized in Appendix A; common abbreviations, measurement units, and prefixes employed in the text may also be found in Appendix A.

References [####] are used in this book to denote the source of:

1. a direct quotation (enclosed in quotes),
2. a lightly or heavily paraphrased passage (footnoted but not enclosed in quotes), or
3. a specific datum.

Citations are also employed to indicate sources of additional information on a given topic, especially literature review papers. The author apologizes in advance for any inadvertent instances of unattributed usage of previously published material; such events should be few but should be brought to the author’s immediate attention for correction in a future edition of this work. An attempt was made to cite primary sources whenever possible, but some references are made to secondary sources believed by the author to be reliable. Unreferenced in-text attributions to specific named people generally refer to comments made by a technical reviewer of the manuscript, as a personal communication.

Acknowledgments

Naturally I am replete with gratitude to many people and organizations, but my first thanks must go to K. Eric Drexler. I honor Eric for his original vision of the revolution in future medicine that molecular nanotechnology will inevitably bring; for patiently yet doggedly pursuing that vision during the long hard years when this position was not yet scientifically popular; and for describing the inspirational concept of cellular repair machines in his popular writings$^{1,2}$ and then offering detailed technical engineering analysis of molecular machine components, devices, and systems in Nanosystems,$^{10}$ a textbook which has established high standards of scholarship and has paved the way for all future research in this area.

Next, I would like to thank the following 77 people for providing useful references, preprints, publications or information, helpful discussions or newsgroup postings, personal communications, or other assistance related to the project: Rehal Bhojani, Forrest Bishop, Frank Boehm, Jonathan Boswell, Robert J. Bradbury, Claud A. Bramblett, Donald W. Brenner, Fred and Linda Chamberlain, Thomas M.S. Chang, M.D., Scott R. Chubb, Sharon Churchill, Philip G. Collins, Andrew Czarn, Jerry A. Darsey, Thomas Donaldson, K. Eric Drexler, William L. Dye, Mark Dyer, Robert I. Eachus, James von Ehr, James C. Ellenbogen, Gregory M. Fahy, David R. Forrest, Berry Fowler, M.D., John Gilmore, James L. Halperin, Jie Han, Barbara and Danny Haukedalen, Dan Heidel, Tad Hogg, Neil Jacobstein, Ted Kaeiher, Jeffrey D. Kooistra, Tobias A. Knoor, Markus Krummenacker, Ronald G. Landes, M.D., Kevin Leung, Eric W. Lewis, M.D.,
I also thank Conrad Schneiker for his highly useful summary of the early history of nanotechnology, which I leaned on rather heavily in Section 1.3.2; Roy Porter, from whose excellent history of medicine I borrowed extensively in Section 1.2.1; W. J. Bishop, whose fine book on the history of surgery provided most of the material for the first six paragraphs of Chapter 1, and substantial material for Section 1.2.1 as well; and the Foresight Institute for establishing and maintaining the first nanomedicine website, under the technically-savvy webmastering of James B. Lewis.

I extend my heartfelt thanks to the 48 individuals listed below who reviewed or commented on all or part of various Chapters in Volume I (total number in parentheses): Douglas Berger, M.D. (1), Forrest Bishop (2), Robert J. Bradbury (8), David Brin (1), Fred Chamberlain (2), Linda Chamberlain (1), Gino J. Coviello (2), Andrew Czarn (1), Thomas Donaldson (1), K. Eric Drexler (6), William L. Dye (2), Martin Edelstein (1), Gregory M. Fahy (2), Steven S. Flitman, M.D. (1), Tim Freeman (1), Thomas W. Gage (2), Al Globus (2), J. Storr Hall (2), Jan H. Hoh (1), Christopher Jones (3), Tanya Jones (2), Ted Kaelber (1), Tobias A. Knoch (1), Markus Krummenacker (7), Sarma Lakkaraju (1), Ronald G. Landes, M.D. (10), Eugene Leitl (1), James B. Lewis (8), James Logajan (5), David Mathes (4), Thomas McKendree (3), Ralph C. Merkle (10), Hans Moravec (1), Max More (1), Kenneth Philipson, M.D. (1), Christopher J. Phoenix (10), Virginia Postrel (1), Edward M. Reifman, D.D.S. (1), Edward A. Rietman (6), Markus Roberts (1), Patrick Salsbury (3), Salvatore Santoli (3), Bruce Smith (4), Steven S. Smith (1), Jeffrey Soreff (8), Tihamer Toth-Fejel (3), James M. Tour (1), and Steven C. Vetter (1). These reviewers are to be lauded for undertaking a difficult task and should be held blameless for any errors that remain in the manuscript; the author is solely responsible for all errors of fact or judgement within these pages. Reports of errata may be transmitted by mail to the author at the following physical address: Robert A. Freitas Jr., Institute for Molecular Manufacturing, 555 Bryant Street, Suite 253, Palo Alto, CA 94301 USA; or electronically to the author at the following Internet address: rfreitas@imm.org.

My special thanks go to Robert J. Bradbury and Markus Krummenacker for particularly lengthy and detailed Chapter reviews, and for illuminating the often shadowy boundaries between biotechnology and nanotechnology; to Jeffrey Soreff for his stunningly insightful commentary and analysis of the most difficult technical concepts (on occasions almost too numerous to count), for his advice on equations and technical phraseology, and for several capsule descriptions of recent research results that are reported in Chapter 2; to Chris Phoenix for reading the entire manuscript with alacrity; to Ralph C. Merkle for his many years of sage technical advice and his unwavering encouragement when it was needed most; and to K. Eric Drexler and Ralph C. Merkle for generously contributing the Foreword and Afterword, respectively, to this Volume.

My very special thanks go to James L. Halperin, Ralph C. Merkle, many anonymous Senior Associates of the Foresight Institute, and other friends, family, and colleagues who generously provided financial support which enabled the author to complete this Volume; and to the Foresight Institute for establishing the Nanomedicine Book Fund.

I thank my publisher, Ronald G. Landes, M.D., for having the perspicacity to recognize the potential significance of nanomedicine, the trust in the author to impose no major constraints on substance, length or time, the fortitude to publish an unusual book of uncertain marketability in a difficult business environment, and the audacity to dare to present the material contained herein to medical and publishing colleagues for their serious consideration. I also applaud Cynthia Dworaczyk, Michelle Wamsley, Penny King, and the rest of the staff of Landes Bioscience for their excellent and professional work on this project, and Andreas Passens for the cover art.

Finally, and most importantly, I wish to thank my wife, Nancy Ann Freitas, and my parents, Robert A. Freitas Sr. and Barbara Lee Freitas, without whose help, understanding, and encouragement this book could not have been written.

Robert A. Freitas Jr.
Research Fellow
Institute for Molecular Manufacturing
15 April 1999
No new findings will ever be made if we rest content with the findings of the past. Besides, a man who follows someone else not only doesn’t find anything, he is not even looking. “But surely you are going to walk in your predecessors’ footsteps?” Yes indeed, I shall use the old road, but if I find a shorter and easier one I shall open it up. The men who pioneered the old routes are leaders, not our masters. Truth lies open to everyone. There has yet to be a monopoly of truth. And there is plenty of it left for future generations too.

— Seneca, the Younger (4 BC - 65 AD)

We are like children standing on the shoulders of a giant, for we can see all that the giant can see, and a little more.

— Guy de Chauliac, Chirurgia Magna, 1362

The human mind is often so awkward and ill-regulated in the career of invention that it is at first diffident, and then despises itself. For it appears at first incredible that any such discovery should be made, and when it has been made, it appears incredible that it should so long have escaped men’s research. All which affords good reason for the hope that a vast mass of inventions yet remains, which may be deduced not only from the investigation of new modes of operation, but also from transferring, comparing, and applying these already known, by the method of what we have termed literate experience.

— Francis Bacon, Novum Organum: Aphorisms on the Interpretation of Nature and the Empire of Man, 1620
The Prospect of Nanomedicine

1.1 A Noble Enterprise

The history of disease is vastly older than that of humankind itself. Indeed, disease and parasitism have been inseparable companions to life since the dawn of life on Earth. Fossilized bacteria, similar to those responsible for many infections that afflict people today have been found in geological formations that are 500 million years old. Fossil shells dating from an era almost equally remote show clear evidence of disturbance by injury and parasites. Examination of the skeletons of long-extinct dinosaurs and other great reptiles show that these creatures suffered from fractures, bone tumors, arthritis, osteomyelitis, dental caries and other diseases that still plague us in the 20th century. While available fossil evidence is largely limited to changes observable in bones and teeth, it is probably safe to assume that disease processes were equally prevalent in the soft organs and tissues that have not been geologically preserved, and that the general pattern of disease has not changed in its essentials during the hundreds of millions of years that animal life has existed on this planet.

Since its first appearance on the prehistoric stage millions of years ago, the human body has also been constantly subject to assault and injury, invasion by parasites, extremes of heat and cold, and infections. Early man probably suffered from a number of diseases due to nutritional factors and body chemistry disorders. For example, the poor condition of the teeth of an 18-year old Australopithecus who lived 1.75 million years ago, found at Olduvai Gorge by Louis Leakey in 1959, suggests that the hominin had a disease that lasted for many months — most likely gastrointestinal due to malnutrition — with three major attacks of the disease at the ages of two, four, and four and a half. Diseases that may date back more than 25 million years to the ape ancestors of modern apes and man are amoebic dysentery, malaria, pinworm infections, syphilis, yaws and yellow fever. Diseases which may have appeared and evolved with man include leprosy and typhoid; certain modern diseases such as cholera, measles, mumps, smallpox, whooping cough, and the common cold require inherited empirical knowledge. It is believed that the art of dressing wounds long constituted the whole of medicine — the use of internal remedies or herbs, and use of the knife or fire, came much later.

The history of disease is vastly older than that of humankind itself. Indeed, disease and parasitism have been inseparable companions to life since the dawn of life on Earth. Fossilized bacteria, similar to those responsible for many infections that afflict people today have been found in geological formations that are 500 million years old. Fossil shells dating from an era almost equally remote show clear evidence of disturbance by injury and parasites. Examination of the skeletons of long-extinct dinosaurs and other great reptiles show that these creatures suffered from fractures, bone tumors, arthritis, osteomyelitis, dental caries and other diseases that still plague us in the 20th century. While available fossil evidence is largely limited to changes observable in bones and teeth, it is probably safe to assume that disease processes were equally prevalent in the soft organs and tissues that have not been geologically preserved, and that the general pattern of disease has not changed in its essentials during the hundreds of millions of years that animal life has existed on this planet.

As for cancer, Java man, first discovered by Dutch anatomist Eugene Dubois in 1891 and considered to be half a million years old, had a morbid growth on his femur. It is likely that bone cancer and other forms of cancer have existed from the earliest times. The remains of Neanderthal, a competing species to Homo sapiens that roamed through Europe, Africa, and the Near East during the last glacial period ~75,000 years ago, show clear evidence of arthritis, tooth loss, and suppurative bone disease. The high rate of broken bones and early death suggests that Neanderthals engaged in more close-quarter combat with large animals than did modern humans, who had figured out safer strategies. Human bones unearthed from the New Stone Age period reveal that Neolithic man suffered from arthritis, congenital dislocations and fractures, sinusitis, tuberculosis of the spine, and tumors.

In considering the question of how our earliest ancestors dealt with these conditions, we are on somewhat uncertain ground, since little direct evidence has been preserved. Anthropologists point out that in pre-Neanderthal hunter-gatherer tribes, a sick or lame person is a serious handicap to a group on the move. In the event of major illness or mortal wounds, sufferers may either leave the group, be abandoned, or, as with lepers in medieval Europe, may be ritually expelled, becoming “culturally dead” before they are biologically dead. Early hunter-gatherer hominid bands were more likely to abandon their seriously sick than to succor them, although Cro-Magnons and Neanderthals evidently were the first to care for their wounded and disabled, and to bury their dead.

But moderate wounds, bruises, fractures, or foreign bodies such as arrowheads or thorns are tangible things that demand attention. Thus the art of surgery must first have originated as a response to immediate crises. The first and most obvious course of action of a wounded man would be to protect the site of injury from the influence of external forces or agents. For this there was, and remains to this day, only one means — the application of a dressing. Many observations were made and many substances tried, and in time a body of experience was accumulated and passed on orally to others for use in similar emergencies, eventually creating a considerable sum of inherited empirical knowledge. It is believed that the art of dressing wounds long constituted the whole of medicine — the use of internal remedies or herbs, and use of the knife or fire, came much later.

Present-day primitive and folk medicines provide additional clues to early medical practice. For example, many early peoples developed effective methods to control bleeding — the use of cobwebs is an ancient folk remedy, as is the application of tourniquets, packing with absorbent materials, the laying on of snow or, at the other temperature extreme, the application of cautery by hot knife or spear. The Masai and Akamba tribes treated sword wounds by slapping on a poultice of cow dung and dust. The suturing of wounds is practiced by some primitive peoples and may have been known to prehistoric man. Bone needles furnished with an eye have been found in paleolithic deposits in France and England, and some Indian tribes suture with threads of sinew or bone needles (the needles are left in and the thread is twisted around them). One of the strangest suturing techniques, observed among primitive tribal cultures in such widely separated places as India, East Africa and Brazil, is the sealing of wounds, especially abdominal wounds, using termites or ants. The edges of the wound are drawn closely together and the insect is allowed to bite through them both, firmly securing the flesh on two sides. Once
attached, the insect’s body is severed, allowing only the jaws to remain in place, holding the wound shut.\textsuperscript{2158}

While able to deal with wounds, early man did not admit the existence of disease from “natural causes”. Internal diseases were generally ascribed to malevolent influences exercised by a supernatural entity or a human enemy. As centuries passed, many internal diseases were eventually recognized and simple treatments empirically established — for example, Celsus (ca. 30 AD), a Roman medical writer, described ligature (tying off blood vessels), suturing the large intestine, eye operations such as crouching for cataract, tonsillectomy, and bladder surgery to remove stones.\textsuperscript{2200} But the rational basis for internal medicine awaited basic physiological knowledge such as the circulation of the blood, finally proven by William Harvey (1578-1657) in 1628, and the discovery and acceptance of the theory of infection by microorganisms in the 1800s. Throughout most of human history, life has been short indeed — Table 1.1 shows the average age of death for a sampling of persons buried in Roman times. (Some variation may represent regional differences in burial practices.) Even by the 17th century, more than half of all children never lived past the age of ten, often succumbing to diseases such as cholera, diphtheria, scarlet fever, or whooping cough, or being scarred for life by smallpox. It was a common and widely accepted fact of life that people would die at all ages, although wealthy families could leave town every year during the cholera season.\textsuperscript{1724}

One of the constant missions of human civilization has been the avoidance and elimination of animals that prey on humans. Cave-dwelling carnivorous saber-toothed tigers, having occupied the upper echelons of the food chain for 30-35 million years, finally became extinct not more than about ten thousand years ago, most likely at the hands of newly-arrived human hunters crossing the Siberian land bridge into post-Pleistocene North America.\textsuperscript{2186} As late as medieval times, wolves ranged freely over Europe, remaining abundant in France through 1500 AD — in winter, audacious wolf packs would enter Paris and eat children, dogs, and even adults who were alone on the streets.\textsuperscript{724} With technological advances, exant tiger and wolf species are now largely confined to artificial habitats or isolated nature preserves and no longer pose any serious threat to human health. People in most places are not eaten by wolves; indeed today, a few venturesome individuals actually keep and breed wolves with dogs, as pets.

Bacteria are among the last remaining “wild animals” on Earth that threaten man. As our instrumentalities continue to progress from the macroscale to the microscale, and finally to the molecular or nanoscale, all of the remaining natural “wild things” that endanger human life and health — whether viruses, bacteria, protozoa, metazoan parasites,\textsuperscript{253,3254} or even our own pathological native cells — will be confined and tamed, reconstructed or eliminated. Using smaller tools, we hunt smaller prey. As with sabertooths in the post-Neolithic era, and with wolves in post-medieval times, people of the 21st century will no longer fear or need suffer predation by wild microbes or tumor cells run amok.

Humanity is poised at the brink of completion of one of its greatest and most noble enterprises. Early in the 21st century, our growing abilities to swiftly repair most traumatic physical injuries, eliminate pathogens, and alleviate suffering using molecular tools will begin to coalesce in a new medical paradigm called nanomedicine. Nanomedicine may be broadly defined as the comprehensive monitoring, control, construction, repair, defense, and improvement of all human biological systems, working from the molecular level, using engineered nanodevices and nanostructures.

### Table 1.1. Average Age of Death of Persons Buried in Roman Times\textsuperscript{201}

<table>
<thead>
<tr>
<th>Burial Location</th>
<th>Average Age of Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>City of Rome</td>
<td>29.9 years</td>
</tr>
<tr>
<td>Iberia</td>
<td>31.4 years</td>
</tr>
<tr>
<td>Britain</td>
<td>32.5 years</td>
</tr>
<tr>
<td>Germany</td>
<td>35.0 years</td>
</tr>
<tr>
<td>North Africa</td>
<td>46.7 years</td>
</tr>
</tbody>
</table>

In this book we shall explore in preliminary fashion a few of the technical details, specific requirements, and physical limitations of these engineered nanodevices, and the meaning of nanomedicine within the context of the traditional healing arts. Since at this writing these nanodevices cannot yet be built, our subject matter is of necessity somewhat speculative, albeit thoroughly grounded in the best scientific and engineering knowledge currently available. Medical philosopher Edmond A. Murphy notes that “to the physician, the more remote the area of speculation from the welfare of the patient, the more clearly [this speculation] must be shown to give promise of ultimate benefit.”\textsuperscript{2228} The author believes that nanomedicine clearly meets, indeed vastly exceeds, this stringent requirement for relevance and utility.

After the manner of many contemporary medical school programs, the present trilogy progresses from a general emphasis on basic science and engineering capabilities in Volume I, to nanomedical tools and systems in Volume II, leading finally to a sharp focus on clinical nanomedicine in Volume III.

#### 1.2 Current Medical Practice

In order to fully appreciate the changes that nanomedicine will inevitably bring, it is useful first to review the history and development of current medical practice. As Winston Churchill once remarked: “The further backward you look, the further forward you can see.” The author unapologetically favors what medical anthropologists would regard as the “Western” healing tradition. The great medical theorist Otto E. Guttentag\textsuperscript{2234} agreed: “Contemporary Western medicine involves a type of healing that potentially and actually exceeds all other approaches in maintaining that optimal status of selfhood we call being healthy and in eliminating that reduced status of selfhood we call being sick.”

After summarizing the history of scientific medicine and placing nanomedicine in its proper historical context (Section 1.2.1), we define exactly what is meant by “medicine” and present a new model for “disease” (Section 1.2.2). The modern medical treatment methodology is examined in light of the changes that will be wrought by nanomedicine (Section 1.2.3), and this is followed by a discussion of several important issues in the evolution of medical bedside practice (Section 1.2.4) and the changing view of the human body (Section 1.2.5).

Many fascinating and important practical issues such as medical privacy rights, patient compliance, medical ethics and bioethics per se, nursing, the role of medical bureaucracies in the doctor-patient interaction, and the question of medical causality\textsuperscript{2236} are not addressed here or are deferred to Chapter 31.
1.2.1 The Evolution of Scientific Medicine

A study of the history of “scientific” or Western medical practice suggests a continuity in certain aspects of the medical paradigm, even since ancient times — but reveals changes in attitudes, techniques, and instrumentalities, and a gradual evolution in the epistemology and ontology of medical logic. Figure 1.1 presents a simple model of this historical progression, the details of which will soon become clear. Note that 2010 is marked as a possible date for the first applications of nanomedicine. These will most likely be ex vivo applications only — in vivo nanomedical treatments may come much later.

Another important reason to study the history of medicine is to gain a deeper appreciation of the long, hard struggle to improve human health, a struggle that is expected finally to culminate in victory in the 21st century. If the ~10 billion people that have ever lived survived an average of 40 years and spent 5% of their lives in sickness or physical misery from disease, then ~200 trillion man-hours of suffering have been paid to achieve this remarkable result, a not inconsiderable price.

There is no pretense to completeness here. For example, Chinese, Indian and Islamic contributions are omitted, not because they are unimportant, but because they did not significantly alter the evolutionary pathway of the Western medical paradigm. A great deal of physiology, pathology, neurology, systematics, and many other important medical-related disciplines are also neglected in the interest of brevity. Much of the following will most likely be ex vivo applications only — in vivo nanomedical treatments may come much later.

Another important reason to study the history of medicine is to gain a deeper appreciation of the long, hard struggle to improve human health, a struggle that is expected finally to culminate in victory in the 21st century. If the ~10 billion people that have ever lived survived an average of 40 years and spent 5% of their lives in sickness or physical misery from disease, then ~200 trillion man-hours of suffering have been paid to achieve this remarkable result, a not inconsiderable price.

There is no pretense to completeness here. For example, Chinese, Indian and Islamic contributions are omitted, not because they are unimportant, but because they did not significantly alter the evolutionary pathway of the Western medical paradigm. A great deal of physiology, pathology, neurology, systematics, and many other important medical-related disciplines are also neglected in the interest of brevity. Much of the following discussion, including most of the anecdotal quotations, is liberally and appreciatively drawn from comprehensive works by W.J. Bishop, Roy Porter, William H. McNeill, and the *Cecil Textbook of Medicine*.

1.2.1.1 Prehistoric Medicine

The elementary nature of prehistorical medical practice has already been mentioned. There is direct evidence that Stone Age man used a natural fungus as a treatment for intestinal parasites. Besides basic wound-tending, the practice of circumcision is an age-old form of simple surgery with a rational, if sometimes controversial, hygienic basis.

An even more dramatic surgical operation, for which there is considerable prehistoric fossil evidence, is trephining the skull to remove a round piece of bone from it. This can be done using flint instruments, either by gradually scratching through the skin and bones of the skull, getting gradually deeper and deeper, or by drilling a series of small holes in a circle in the skull, then cutting the small bridges between to remove a disk of bone. Skulls have been found with multiple holes. The additional fact that some hole edges show callus formation (evidence of healing) indicates survival of the patient. Trephined skulls have been found in Western Europe, including England, North Africa, Asia, the East Indies, New Zealand, and the Americas from Alaska in the north down through the continent to Peru in the south. The practice was probably regarded as therapeutic, either to remove a depressed fracture, to try to cure mental illness, or to relieve severe headache or epilepsy, presumably by letting out the demon possessing the patient.

1.2.1.2 Ancient Mesopotamian Medicine

The first written records, which came from ancient Babylon and Egypt, contain the earliest references to medical care, obviously codifying earlier practices no record of which has survived. The medicine of the Sumerians of Mesopotamia from ca. 3000 BC was primarily religious. The Mesopotamian peoples saw the hands of the gods in everything. Disease was caused by spirit invasion, sorcery, malice, or the breaking of taboos, and sickness was both judgement and punishment. An Assyrian text circa 650 BC describes epileptic symptoms within a demonological framework: “If at the time of the possession, his mind is awake, the demon can be driven out; if at the time of his possession his mind is not so aware, the demon cannot be driven out.”

But medicine also had an empirical component, with some sicknesses being ascribed to cold, dust and dryness, putrefaction, malnutrition, venereal infection and other natural causes. The Babylonians drew on an extensive pragmatic materia medica — some 120 mineral drugs and twice that number of vegetable items are listed in surviving tablets. Alongside various fats, oils, honey, wax, and milk, were many active ingredients that included mustard, oleander and hellebore (a plant in the buttercup family that is a violent gastrointestinal poison, hence acts as a powerful purgative, though it is lethal in high doses). Colocynth, senna and castor oil were used as laxatives, while wound dressings were compounded with dried wine dregs, salt, oil, beer, juniper, mud or fat, blended with alkali and herbs. With the discovery of distillation, the

*Honey also was known to the ancients as a preservative agent. The corpses of military leaders who had died far from home and needed to be transported long distances were often immersed in honey to prevent decomposition. Famous historical examples from later Greek and Roman times include the Spartan commander Agesilaus (d. 362 BC) and Alexander the Great (d. 323 BC).*
Mesopotamians made essence of cedar and other volatile oils. Turpentine, asafetida, henbane, myrrh, mint, poppy, fig, and mandrake are also mentioned. Dog dung and other fecal ingredients were used to drive off demons.

Surgical conditions such as wounds, fractures and abscesses were also treated by Mesopotamian surgeons. Practitioners were priests, and after 2000 BC, they were ruled by the strict laws included in the Code of Hammurabi. The Code laid down rewards for success and severe punishment for failure, and contained laws relating to medical practice which show that medicine and surgery were highly organized professions. Fees were regulated on a sliding scale of rewards based on the patient's rank, and severe penalties were laid down for failure:

"Concerning the wounds resulting from operations it is written: if a physician shall produce on anyone a severe wound with a bronze operating knife and cure him, or if he shall open an abscess with the operating knife and preserve the eye of the patient, he usually shall receive 10 shekels of silver [more than a craftsman’s annual pay]; if it is a slave, his master shall usually pay 2 shekels of silver to the physician."

"If a physician shall make a severe wound with an operating knife and kill him, or shall open an abscess with an operating knife and destroy the eye, his hands shall be cut off."

"If a physician shall make a severe wound with a bronze operating knife on the slave of a free man and kill him, he shall replace the slave with another slave. If he shall open an abscess with a bronze operating knife and destroy the eye, he shall pay the half of the value of the slave."

The Hammurabic Code also mentions the Gallabu, or barber-surgeons, whose province was minor surgery, including dentistry and the branding of slaves. If Herodotus (ca. 485-425 BC) may be believed, Babylonian medicine must have declined in the 5th century BC. Herodotus states that there were no physicians, but that the people brought their sick into the marketplace in order that passers-by might make suggestions or offer cures.

1.2.1.3 Ancient Egyptian Medicine

The civilization of ancient Egypt dates from around 3000 BC. As in Mesopotamia, ancient Egyptian medicine was religious-empirical. An examination of the preserved bodies of members of the Royal family has provided much information on the diseases of ancient Egypt, including congenital deformities such as clubfoot, dental caries, gallstones, bladder and kidney stones, rheumatoid arthritis, mastoiditis, numerous eye diseases, and bone fractures, some of which show treatment by quite sophisticated splinting. Other evidence suggests that the average ancient Egyptian was extremely diseased. For example, a weaver who died in the 11th century BC, aged 14-18 years, evidently suffered from schistosomiasis, tapeworm likely associated with malnutrition, anthracosis of the lungs presumably due to environmental pollution from cooking and heating, pulmonary silicosis, and possibly malaria and fleas.

The Ebers Papyrus (ca. 1550 BC), deriving from Thebes, is the principal medical document and may be the oldest surviving medical book. Over 20 meters long, it deals with scores of diseases and proposes remedies involving spells and incantations, but also includes many rational treatments. The Ebers Papyrus covers 15 diseases of the abdomen, 29 of the eyes, and 18 of the skin, and, perhaps unsurprisingly to the modern consumer, lists no fewer than 21 cough treatments. About 700 drugs and 800 formulations are mentioned, mainly involving herbs but also including mineral and animal remedies. For example, to cure night blindness the patient should eat fried ox liver — possibly a tried-and-tested procedure, since liver is rich in Vitamin A, lack of which causes the illness. Eye disorders were also common, for example:

"To drive away inflammation of the eyes, grind the stems of the juniper of Byblos, steep them in water, apply to the eyes of the sick person and he will be quickly cured. To cure granulations of the eye, prepare a remedy of callyrium, verdigris, onions, blue vitriol, powdered wood, and mix and apply to the eyes."

For stomach ailments, a decoction of cumin, goose-fat and milk was recommended, but other remedies sound more exotic, including a drink prepared from black as testicles. A mixture of vulva and penis extracts and black lizard was supposed to cure baldness. Also good for hair growth was a compound of hippopotamus, lion, crocodile, goose, snake and ibex fat — merely assembling these ingredients might promise a hair-raising experience! Egyptian medicine credited many vegetables and fruits with healing properties, and also tree products such as sycamore bark and resins such as myrrh, frankincense and manna. As in Mesopotamia, plant extracts — notably senna, colocynth and castor oil — were employed as purgatives, and oil of camomile to improve digestion. Recipes included ox spleen, pig’s brain, stag’s horn, honey-sweetened tortoise gall, and the blood and fats of various animals. Antimony, copper salts, alum, and other minerals were recommended as astringents or disinfectants. Containing ingredients from leeks to lapis lazuli — including garlic, onion, tamarisk, cereals, spices, condiments, resins, gums, dates, hellebore, opium and cannabis — compound drugs were administered in the form of pills, ointments, poultices, fumigations, inhalations, gargles and suppositories; they might even be blown into the urethra through a tube. Among the most interesting of the healers whose names have been recorded for posterity are Peseshet, head female physician or overseer, proof of the existence, as in Mesopotamia, of female healers; and Iri, Keeper of the Royal Rectum, presumably the pharaoh’s enema expert.

The Edwin Smith Papyrus (ca. 1600 BC), discovered by its American namesake at Luxor in 1862, may be the world’s earliest surviving surgical text, and contains material probably derived from even more ancient times. The “book of wounds” comprises 48 case reports, which commence with the top of the head and proceed systematically downward — nose, face, ears, neck and chest, mysteriously stopping in mid-sentence at the spine, presumably when the scribe was interrupted at his work. The only surgical conditions treated were wounds, fractures, abscesses, and circumcisions. In the Papyrus, the method of presentation is first to set out the title or the chief symptom, followed by the further symptoms, and then the examination, diagnosis, prognosis, and treatment. The advice given is entirely rational, as will be seen by the following directions for the treatment of a fractured humerus:

"Instructions concerning a Break in his Upper Arm. If thou examinest a man having a break in his upper arm, and thou findest his upper arm hanging down, separated from its fellow, thou shouldest say concerning him: One having a break in his upper arm. An ailment which I will treat."

"Thou shouldest place him prostrate on his back, with something folded between his two shoulder blades; thou shouldest spread out with his two shoulders in order to stretch apart his upper arm until that break falls into its place. Thou shouldest make for him two splints of linen, and thou shouldest apply for him one of them both on the inside of his arm, and the other of them both on the underside of his arm. Thou shouldest bind it with ymrw [an unidentified mineral substance], and treat it afterward with honey every day until he recovers."
An interesting point in the case histories is that after the diagnosis the
writer gives a decision about his further course of action. The
verdict may take one of three forms: (1) an ailment which I will
treat; (2) an ailment with which I will contend; and (3) an ailment
not to be treated. Like modern military triage, the “hopeless” patient
is left to his inevitable fate. This guarded attitude on the part of
the medical man was widespread in antiquity. While present-day doctors
generally do everything possible to alleviate symptoms to the very
end, even when the patient has no chance of recovery, the view in
ancient times was that hopeless cases were not to be touched. This
attitude was entirely practical. A doctor in attendance at the courts
of ancient Egypt might expect rich rewards if his patient recovered,
but if a patient died under his care, the unfortunate physician ran a
grave risk of impalement.

1.2.1.4 Ancient Greek Medicine

By 1000 BC the communities later collectively known as the
Greeks were emerging around the Aegean Sea. How much medical
knowledge they took from Egypt remains controversial, but the
contrasts between the two are striking. Little is known of Greek
medicine before the appearance of written texts in the 5th century
BC. Archaic Greece had its folk healers, including priest healers
employing divination and herbs. From early times (the first Olympic
games were recorded in 776 BC), the love of athletics produced
instructors in exercise, bathing, massage, gymnastics and diet. The
Homeric epics (ca. 600 BC) offer glimpses of early Greek medicine.
Scholars count 147 cases of battle wounds in the
_Iliad_, including
106 spear thrusts, 17 sword slashes, 12 arrow shots, and 12 sling
shots. Among the arrow wound survivors was King Menelaus of
Thyestes, who had 106 spear thrusts and 17 sword slashes. The
verse may take one of three forms: (1) an ailment which I will
treat; (2) an ailment with which I will contend; and (3) an ailment
not to be treated. Like modern military triage, the “hopeless” patient
is left to his inevitable fate. This guarded attitude on the part of
the medical man was widespread in antiquity. While present-day doctors
generally do everything possible to alleviate symptoms to the very
end, even when the patient has no chance of recovery, the view in
ancient times was that hopeless cases were not to be touched. This
attitude was entirely practical. A doctor in attendance at the courts
of ancient Egypt might expect rich rewards if his patient recovered,
but if a patient died under his care, the unfortunate physician ran a
guarded attitude on the part of the
medical man was widespread in antiquity. While present-day doctors
generally do everything possible to alleviate symptoms to the very
end, even when the patient has no chance of recovery, the view in
ancient times was that hopeless cases were not to be touched. This
attitude was entirely practical. A doctor in attendance at the courts
of ancient Egypt might expect rich rewards if his patient recovered,
but if a patient died under his care, the unfortunate physician ran a
grave risk of impalement.

1.2.1.4 Ancient Greek Medicine

By 1000 BC the communities later collectively known as the
Greeks were emerging around the Aegean Sea. How much medical
knowledge they took from Egypt remains controversial, but the
contrasts between the two are striking. Little is known of Greek
medicine before the appearance of written texts in the 5th century
BC. Archaic Greece had its folk healers, including priest healers
employing divination and herbs. From early times (the first Olympic
games were recorded in 776 BC), the love of athletics produced
instructors in exercise, bathing, massage, gymnastics and diet. The
Homeric epics (ca. 600 BC) offer glimpses of early Greek medicine.
Scholars count 147 cases of battle wounds in the
*Iliad*, including
106 spear thrusts, 17 sword slashes, 12 arrow shots, and 12 sling
shots. Among the arrow wound survivors was King Menelaus of
Sparta, whose physician extracted the arrow, sucked out the blood
and applied a salve. As with other medical interventions in Homer,
this shows no Egyptian influence, supporting the idea that, even if
Greek practice owed much to Egypt, it rapidly went its own way.

Various Greek gods and heroes were identified with health and
disease, the chief being Asclepius, who even had the power to raise
the dead. A heroic warrior and blameless physician, Asclepius was
the son of Apollo, sired upon a mortal mother, who was taught
herbal remedies by Chiron and then generously used them to heal
humans. Incensed at being cheated of death, Hades, the ruler of the
underworld, appealed to the supreme god, Zeus, who obligingly
dispatched Asclepius with a thunderbolt, though he was later
elevated to godhood. A different version appears in Homer, who
portrays Asclepius as a tribal chief and a skilled wound healer whose
sons became physicians and were called Asclepiads, and from whom
all Asclepian practitioners descended. As the tutelary god of medicine,
Asclepius is usually portrayed with a beard, staff and snake — the
origin of the caduceus symbol of the modern physician, with its
two snakes intertwined, double-helix like, on a winged staff. The
god was often shown accompanied by his daughters, Hygeia (health
or hygiene) and Panacea (cure-all).

For all that, Hippocratic medicine, the foundation of Greek
written medicine, explicitly grounds the art upon a quite different
basis — a healing system independent of the supernatural and built
upon natural philosophy. The beginnings of true medical science in
the West were established when the reliance on superstition that
underpinned tribal medicine was replaced by civilized and rational
curiosity about the cause of illness. The growth of civilized thought
allow for argument on medical cause and cure, with great doctrinal
multiplicity. The separation of medicine from religion reveals
another distinctive feature of Greek healing: its openness, a quality
characteristic of Greek intellectual activity in general, owing to
political diversity. There was no imperial Hammurabic Code and,
unlike Egypt, no state medical bureaucracy, nor were there
documents or professional qualifications. Those calling themselves
doctors (iatroi) had to compete with bone-setters, exorcists, root-
cutters, incantatory priests, gymnasts, and showmen, exposed to
the quips of playwrights and the criticism of philosophers. Medicine
was open to all.

Empedocles (fl. 450 BC) may have been the first to advance
some of the key physiological doctrines in Greek medicine, including
innate heat as the source of living processes such as digestion, the
cooling function of breathing, and the notion that the liver makes
the blood that nourishes the tissues. His contemporary, Alcmaeon
of Croton (fl. 470 BC), believed that the brain, not the heart, was
the chief organ of sensation. Alcmaeon’s examination of the eyeball
led him to discern the optic nerve leading into the skull, a genuine
observational basis. He gave similar explanations for the sensations
of hearing and smelling, because the ear and nostrils suggested passages
leading to the brain. Most of such knowledge depended heavily on
wound observation and animal dissection, for in the classical period
the dignity of the human body forbade dissection.

All we know for sure of Hippocrates (ca. 460-377 BC), who
“taught all that were prepared to pay,” is that he was born on the
island of Cos and lived a long and virtuous life. The sixty or so
works comprising the Hippocratic Corpus (ca. 440-340 BC) derive
from a variety of hands, and, as with the books of the Bible, they
became jumbled up, fragmented, then pasted together again in
antiquity. What is now called the Corpus was gathered around 250
BC in the Library at Alexandria, with further texts added later still.
Some volumes are philosophical, others are teaching texts or case
notes. What unites them all is the conviction that health and
disease are capable of explanation by reasoning about nature,
independently of supernatural interference. Man is governed by the
same physical laws as the cosmos, hence medicine must be an
understanding, empirical and rational, of the workings of the body
in its natural environment. Anticipating modern medicine, appeal
to reason, rather than to rules or to supernatural forces, gives
Hippocratic medicine its distinctiveness. It was also patient- rather
than disease-centered; the Hippocratics specialized in medicine by
the bedside, prizing trust-based clinical relations:

“Make frequent visits; be especially careful in your examinations,
counteracting the thing wherein you have been deceived at the change.
Thou will know the case more easily, and at the same time you will
also be more at your ease. For instability is characteristic of the humours,
and so they may also be easily altered by nature and by chance.”

“...Keep a watch also on the faults of the patients, which often make
them lie about the taking of things prescribed. For through not taking
disagreeable drinks, purgative or other, they sometimes die. What they
have done never results in a confession, but the blame is thrown upon
the physician.”

The Hippocratics promulgated the idea of “vis medicatrix
naturae,” or the power of nature to cure itself, and thus the belief
that there was a natural tendency for things to get better on their
own. This tendency could be aided by providing a beneficial
environment for the patient and by improving physical function
with a regime of suitable diet, lifestyle, and exercise — the diatetics.
In extreme cases, further aids to recovery could be sought.
Stubbornly offending “humors” could be removed with the help of
venesection (phlebotomy or bloodletting) and purgatives, sudorifics
applied to induce sweating, and diuretics to increase urination. But
the Hippocratic physician was extremely reluctant to administer
drugs of any kind, because the physician’s goal was to aid nature in
healing the body.
Hippocrates scorned heroic interventions and left risky procedures to others. Their Oath (Chapter 31) explicitly forbade cutting, even for stones, and other texts reserved surgery for those used to handling war wounds. Surgery was regarded as an inferior trade, the work of the hand rather than the head, a fact reflected in its name: “surgery” derives from the Latin “chirurgia,” which comes from the Greek “chéiros” (hand) and “ergon” (work); surgery was handiwork. Hippocratic surgical texts were thus conservative in outlook, encouraging a tradition in which doctors sought to treat complaints first through management, occasionally through drugs, and finally, if need be, by surgical intervention.

The art of diagnosis involved creating a profile of the patient’s lifestyle, work and dietary habits, partly by asking questions and partly by the use of trained senses:

“When you examine the patient, inquire into all particulars; first how the head is...then examine if the hypochondrium (abdomen beneath lower ribs) and sides be free of pain, for...if there be pain in the side, and along with the pain either cough, terminus [painful intestinal colic] or hellebute, the bowels should be opened with clysters [enemas]. The Physician should ascertain whether the patient be apt to faint when he is raised up, and whether his breathing is free...”

Hippocrates prided themselves on their clinical acuity, being quick to pick up telltale symptoms, as with the facies hippocratica, the facial look of those dying from long-continued illness or cholera: “a protrusive nose, hollow eyes, sunken temples, cold ears that are drawn in with the lobes turned outward, the forehead’s skin rough and tense like parchment, and the whole face greenish or black or blue-grey or leaden.” Experience was condensed into aphorisms, as for instance: “When sleep puts an end to delirium, it is a good sign.”

The technique most prized among Hippocrates was the art of prognosis — a secular version of the priestly and oracular prognostications of earlier medicine, and bearing some analogy to the 20th century weatherman, who can give a bright or gloomy forecast but is powerless to change it. Noted one Hippocratic text:

“It appears to me a most excellent thing for the physician to cultivate Prognosis; for by foreseeing and foretelling, in the presence of the sick, the present, the past, and the future, and explaining the omisions which patients have been guilty of, he will be the more readily believed to be acquainted with the circumstances of the sick; so that men will have confidence to entrust themselves to such a physician... Thus a man will be the more esteemed to be a good physician...from having long anticipated everything; and by seeing and announcing beforehand those who will live and those who will die.”

The ultimate significance of Hippocratic medicine was twofold. First, it carved out a lofty role for the selfless physician which would serve as a lasting model for professional identity and conduct. Second, it taught that an understanding of sickness required an understanding of nature.

1.2.1.5 Ancient Alexandrian Medicine

Soon after the death of Aristotle (d. 322 BC) and his most famous pupil, Alexander the Great (d. 323 BC), a great medical school was founded in Egypt at the court of King Ptolemy, at his capital, Alexandria, at the mouth of the Nile. The King’s main cultural creations were the Alexandrian Library* and the Museum (Sanctuary of the Muses), which installed Greek learning in a new Egyptian environment — Archimedes, Euclid, and the astronomer Ptolemy were soon to teach there. The Library became a wonder of the scholarly world, eventually containing, it was said, 700,000 manuscripts, and other facilities including an observatory, zoological gardens, lecture halls and rooms for research.

The two earliest teachers at the Alexandrian medical school were also its greatest — Herophilus of Chalcedon (ca. 330-260 BC) and his contemporary, Erasistratus of Chios (ca. 330-255 BC). Their writings having been lost, we know about them only through later physicians. Herophilus was the first to dissect cadavers in public. He was a student of Praxagoras of Cos (fl. 340 BC), who had improved Aristotelian anatomy by distinguishing arteries from veins, but who saw the arteries as air tubes, similar to the trachea and bronchi, a common error because arteries are devoid of blood in corpses. Herophilus observed that the coats of arteries were much thicker than those of the veins, thus he speculated that the arteries were filled not with air but with blood. Herophilus wrote at least eleven treatises, discovering and naming the prostate and the duodenum (from the Greek for twelve fingers, the length of gut he found). He also wrote on the pulse as a diagnostic guide and on therapeutics, ophthalmology, dietetics, and midwifery. He recognized the brain as the central organ of the nervous system and the seat of intelligence, extending the knowledge of the parts of the brain, certain of which still bear titles translated from those given by him. He was the first to grasp the nature of the nerves, which he distinguished as motor and sensory, though he did not separate them clearly from tendons. Erasistratus surmised that every organ is formed of a three-fold system of “vessels” — veins, arteries, and nerves, dividing indefinitely. These, plaited together, were postulated to make up the tissues. In the brain, Erasistrates observed convolutions, noting that they were more elaborate in man than in animals and associating this with higher intelligence. He distinguished between cerebrum and cerebellum, and is often regarded as an early mechanist because of his model of bodily processes — for instance, digestion involved the stomach grinding food. He was opposed to intrusive remedies such as venesection, and his favorite therapeutic measures were regulated exercise, diet, and the vapor bath — very much in the Hippocratic tradition.

1.2.1.6 Ancient Roman Medicine

Roman tradition held that one was better off without doctors. According to Cato (234-149 BC), citizens had no need of professional physicians because Romans were hale and hearty, unlike the effete Greeks. Apparently Romans enjoyed bad-mouthing Greek physicians. Thus the Romans despised medicine as a profession but this did not prevent them from making use of Greek physicians or even of the services of their own slaves. Galen of Pergamum (130-200 AD) tells us that in his time, large cities such as Rome and Alexandria swarmed with specialists who also travelled about from place to place. Martial (40-104 AD) mentions some of them in an epigram: “Cascellius extracts and repairs bad teeth; you, Hyginus, cauterize ingrowing eyelashes; Fannius cures a relaxed uvula without cutting; Eros removes brand marks from slaves; Hermes is a very Podalirius for ruptures.” Under the Empire, military medicine was highly organized — every cohort had its surgeon, and surgeons of a higher grade were attached to the legions as consultants. Army surgeons ranked as noncombatants and enjoyed many privileges.

In the Roman empire, the earliest scientific teacher was a Greek, Asclepiades of Bithynia (124-40 BC). Asclepiades ridiculed the Hippocratic expectant attitude as a mere “meditation on death,” and urged active measures that the cure might be “seemly, swift and...

*Part of the Library was wrecked by fire in 48 BC during riots sparked by the arrival of Julius Caesar. Much later, Christian leaders encouraged the destruction of the Temple of Muses. As legend has it, in 415 AD the last resident scholar, the female mathematician Hypatia, was hauled out of the Museum by Christian fanatics and beaten to death (or her flesh was ripped off using clamshells, by another account). The 7th century Muslim conquest of the city resulted in the final destruction of the Library.
sure.” Though outside the mainstream, his medical practice is interesting as a modification of the atomic or corpuscular theory, according to which disease results from an irregular or inharmonious motion of the corpuscles of the body. His pupils were numerous, constituting the Methodical school, but his available therapeutic tools were few — he trusted mainly to changes of diet, friction, bathing, exercise, and occasionally emetics, bleeding, and wine. He was also the first to use music in the treatment of the insane.

Back in the mainstream, Galen of Pergamum (130-200 AD) provided the final medical synthesis of antiquity and the effective medical standard for the next 13 centuries. His first medical appointment was as surgeon to the Roman gladiators. He later traveled to Rome and wrote extensively on anatomy, physiology and practical medicine. His fame is due in part to his prolific pen — some 350 authentic titles ranging in topic from the soul to blood-letting polemies survive, about as much as all Greek medical writings together. Galen was a flamboyant character. One of his party tricks, revealing his genius for self-advertisement as well as experiment, was to sever the nerves in the neck of a pig. As these were severed, one by one, the pig continued to squeal; but when Galen cut one off the laryngeal nerves the squealing stopped, impressing the crowd.

Galen justified veneration in terms of his elaborate pulse lore. Written in the early 170s, his sixteen books on the pulse were divided into four treatises, each four books long. In one of these treatises, he explains how to take the pulse and to interpret it, raising key questions, for example: How was it possible to tell whether a pulse was full, rapid, or rhythmical? Such questions he resolved partly from experience and partly by reference to earlier authorities. Galen developed a characteristic physiological scheme that remained in vogue until the 17th century. It supposes three types of so-called spirits associated with three types of the activity of living things. These were the natural spirits formed in the liver and distributed by the veins, the vital spirits formed in the heart and distributed by the arteries, and the animal spirits formed in the brain and distributed by the nerves. Galen's system was an admirable if factually flawed working hypothesis, based on much experimental evidence, and he presented his work as “perfecting” the legacy of Hippocrates.

As in Greece, medicine remained personal in Rome. No medical degrees were conferred or qualifications required. In the absence of colleges and universities, the private face-to-face nature of medical instruction encouraged fluidity and diversity. Students attached themselves to an individual teacher, sitting at his feet and accompanying him on his rounds. Many different sorts of medical care were available, and self-help was universal. Celsus’ On Medicine was written for a non-professional readership as willing to wield the scalpel as the plough. Disease explanations changed little. Public authorities still ascribed famines and pestilences to the gods, and Galen was silent on contagion. The essential trilogy of classical medicine remained dietetics, exercise, and drugs, accompanied by light surgery.

1.2.1.7 Medicine in the Middle Ages

The passage from the glorious days of Rome to the Middle Ages was often violent, especially in the West, with wave after wave of barbarian onsloughts from the East. These culminated in the sack of the Eternal City by Alaric's Goths in 410 AD, which effectively ended the western empire and frayed the thread of learned medicine. Thus Galen had no effective successor. Indeed, medieval medicine may be summed up as a corrupted version of Galenism. The true scientific tradition did not reappear in the West until the 16th century, after a lengthy incubation in the Islamic world. For example, in England the Venerable Bede (ca. 672-735) and his monks possessed many medical writings, and knowledge of plant remedies was extensive, but the English healer used charms and charms, predicated on the belief that certain diseases and bad luck were caused by darts shot by elves, while other ailments involved a “great worm,” a term applied to snakes, insects, and dragons.

By contrast with the naturalist focus of Hippocratic and Galenic medicine, healing became more authoritarian and intertwined with religion, for the rising Church taught that there was a supernatural plan and purpose to everything, including sickness and death. Christian and Jewish healing traditions became more prominent. Disease could be cured by prayers or by invoking the names of saints, by exorcism, by amulets or number magic, or by transferring the sickness to animals, plants, or to the soil. Certain maladies such as leprosy were associated with the Almighty’s punishments for sin, according to the Book of Leviticus:

> “When a man shall have in the skin of his flesh a rising [a swelling], a scab, or bright spot, and it be in the skin of his flesh like the plague [the spots] of leprosy... and the plague in sight be deeper than the skin of his flesh it is a plague of leprosy: and the priest shall look on him, and pronounce him unclean.”

Such polluting diseases were curable by the Lord alone, encouraging some Jewish people to reject human medicine in favor of the divine, citing the fate of King Asa (ca. 914-874 BC), who “sought not the Lord, but his physicians,” and whose foot sores consequently worsened, and he died. On the other hand, while Jewish dietary rituals (e.g. kosher food) are principally expressions of religious precepts about pollution and purification, a useful practical effect is to limit exposure to foodborne diseases such as trichinosis.2215

Many of the distinguished surgeons of the Middle Ages were clerics, but the practice of medicine and surgery by members of the Church was not favored by the hierarchy. Many laws were passed against the practice of medicine for worldly profit. In 1215 the Fourth Lateran Council forbade all sub-deacons, deacons, or priests to practice that part of surgery that had to do with burning and cutting. Finally, Pope Honorius III (d. 1216) prohibited all persons in holy orders from practicing medicine in any form, which meant that the educated classes were prohibited from performing any type of surgery.

In Europe in the Middle Ages, practitioners unsuccessfully sought to cure wounds by treating only the weapons that caused them, using an ointment on the offending knife or sword known as a weapon salve. Physicians practiced urine-gazing (uroscopy), but proper chemical analysis of urine did not develop until the 18th century. Medieval medicine is the source of many humorous “remedies” in contemporary diatribes against modern medicine — such as being strapped into a halter and walked around a pigsty three times for mumps, drinking water from the skull of a bishop to treat rheumatism, or nuzzling a mouse to cure the common cold.2216

But we must recognize that such treatments were not put forth as reasonable science. Rather, they serve as a reminder of the results of injecting mystical or religious elements into medicine as a substitute for good science.

The Middle Ages are also remarkable for their plagues, which became more numerous as people crowded closer and closer together in urban centers of increasing population density. Plague records from China during this period often show 50%-70% mortality rates, and China probably served as the original source of the greatest of the bubonic plagues that swept through Europe in the 14th century, known as the Black Death. The disease broke out in 1346 among the armies of a Mongol prince who laid siege to the
trading city of Caffa in the Crimea. This compelled his withdrawal, but not before the disease had entered Caffa, whence it rapidly spread by ship throughout the Mediterranean.

The initial shock in 1346-1350 was severe. Die-offs varied widely, with some small communities experiencing total extinction and others, such as Milan, being spared entirely. The lethal effect of the plague may have been enhanced by the fact that it was propagated not solely by the bites of fleas carried on infected rodents, but also was transmitted person to person as a result of inhaling bacillus-filled droplets that had been coughed or sneezed into the air by an infected individual. Lung infections of this kind were observed to be 100% lethal in Manchuria in 1921, the only time modern medicine has directly observed airborne plague communication, so it is tempting to assume a similar mortality for pneumonic plague in 14th century Europe. Mortality rates for sufferers from bubonic infection transmitted by flea bite varies from pneumonic plague in 14th century Europe. Mortality rates for communicable, so it is tempting to assume a similar mortality for time modern medicine has directly observed airborne plague communication, so it is tempting to assume a similar mortality for

The plague returned in the 1360s, the 1370s, and thereafter, and European population declined irregularly as a result, reaching a low point in England sometime between 1440-1480. People learned to minimize infection risk via quarantine, a practice which stemmed from Biblical passages prescribing the ostracism of lepers. Plague sufferers were treated as though they were temporary lepers, with a standard 40-day quarantine for people and incoming ships at all major ports. However, since the role of fleas and rats in disease propagation remained unknown until the end of the 19th century, quarantine measures were often ineffectual. Through the 17th century, occasional plague outbreaks that carried off up to a third or a half of a city's population in a single year were considered normal. For example, Venetian statistics show that in 1575-77 and again in 1630-31, a third or more of the city's population died of plague. Pre-modern medicine was powerless against bubonic plague. Before antibiotics reduced the disease to triviality in 1943, the average mortality rate was 60%-70% of those affected, despite all that the best hospital care could accomplish. The last recorded outbreak of plague that ran its course without benefit of penicillin and related antibiotics (which destroy the infection rapidly) occurred in Burma in 1947, with 78% lethality.

1.2.1.8 Renaissance and Pre-Modern Medicine

While internal medicine languished through the Middle Ages, the surgeons of the Renaissance gained wide experience during the many religious wars of the period. They had many new problems to face, including the treatment of wounds caused by firearms. The French had employed gunpowder at the siege of Puigualma in 1338, and cannon were used by the English at the Battle of Crécy in 1346. Gunshot wounds at that time were caused by large missiles of low velocity that caused ragged wounds and carried pieces of clothing into the tissues. These wounds were severe and very liable to become septic. The universal belief among contemporary surgeons was that gunpowder itself was venomous. To neutralize the effect of this venom, the general practice was to cauterize the wound by injecting boiling oil.

The first man to break away from this old doctrine was Ambroise Pare (1510-90). Pare came to Paris in 1532 as an apprentice to a barber-surgeon and then moved to the great Hotel Dieu as resident surgeon. In that immense medieval hospital, the only one in Paris at the time, he gained great experience, and in 1536 he began his career as a military surgeon. As described in his book The Apologie

Pare also went on to show that bleeding after amputations should be arrested, not by the terrible method of the indiscriminate use of the red-hot cautery, but by simply tying of the blood vessels. His most famous phrase, so reminiscent of the old Hippocratic school of thought, was: “I dressed the wound, and God healed him.” Even by the 18th century, the traditional surgeon's day-to-day business eschewed high-risk operations like amputations; rather, it was a round of minor procedures such as venesection, lancing boils, dressing skin abrasions, pulling teeth, managing whitlows, trussing ruptures, and treating skin ulcers. The fatality rates of these procedures were low, for surgeons understood their limits, and the repertoire of thought, was: “I dressed the wound, and God healed him.” In 1633 appeared the earliest book on first-aid for the injured, by one Stephen Bradwell, although the proffered advice sounds impractical and a bit odd to modern ears. For example, the treatment for the “Biting of a Madde Dogge” is to throw the patient into water. “In doing this, if he cannot swim, after he hath swallowed a good quantity of water, take him out again. But if he be skilful in swimming, hold him under the water a little while till he have taken in some pretty quantity.” This procedure may not be wholly irrational — standard 20th century first aid for dog bites includes a thorough cleansing of the wound with water.

Venesecmation remained a popular 17th century universal remedy. As described by the surgeon Richard Wiseman (1622-1676), “a gentleman of about thirty years of age coming out of Hertfordshire through Tottenham and riding upon the causeway near an inn, one emptying a chamber pot out of the window as he was passing by, his horse started and rushed violently between a signpost and a tree which supported part of the sign. The poor gentleman was beaten off his horse and lay stunned upon the ground.” A barber-surgeon was hastily summoned but nothing much was done for the injured man until Wiseman arrived, whereupon:

“I found the gentleman lying upon the ground, the people and chirurgeon gazing upon him. I felt his pulse much oppressed, the right brow bruised and inquired whether they had bled him blood. The chirurgeon replied that he had opened a vein in his arm but it would not bleed. I replied, we must make him bleed through it by splitting his veins. Turning his head on one side, I saw the jugular vein on the bruised side turgid and opened it. He bled freely. After I had taken about twelve ounces, the blood ran down from his arm which had been opened before and would not bleed. We bled him till he came to life, and then he raved and struggled with us.”

The patient’s injuries were dressed and he was subjected to further bleedings, but evidently made a good recovery.

Even by the 18th century, the traditional surgeon’s day-to-day business eschewed high-risk operations like amputations; rather, it was a round of minor procedures such as venesection, lancing boils, dressing skin abrasions, pulling teeth, managing whitlows, trussing ruptures, and treating skin ulcers. The fatality rates of these procedures were low, for surgeons understood their limits, and the repertoire of operations they attempted was small, because of the well-known risks of trauma, blood loss, and sepsis. Internal disorders were treated not by the knife but by medicines and management, since major internal surgery was unthinkable before anesthetics and antiseptic
procedures. Improvements did occur in certain operations such as lithotomy. William Cheselden (1688-1752), a great British surgeon of the 18th century, perfected a technique which enabled him to remove a stone in the bladder in one minute (his record time was 54 seconds), thus reducing mortality from about 50% to under 10%. Cheselden’s results were not bettered until almost the end of the 19th century.

In the 17th century, internal medical treatment was frequently overdone on those affluent enough to afford it. Critics often denounced physicians as meddlesome, capriciously practicing an often dangerous polypharmacy — a blunderbuss approach. The deathbed of Charles II (1630-1685) of England was a conspicuous case of such medical overkill; after the king had suffered a stroke, his doctors moved in, and Sir Raymond Crawford (1865-1938) recreated the scene:

“Sixteen ounces of blood were removed from a vein in his right arm with immediate good effect. As was the approved practice at this time, the King was allowed to remain in the chair in which the convulsions seized him. His teeth were held forcibly open to prevent him biting his tongue. The regimen was, as Roger North pithily describes it, first to get him to wake, and then to keep him from sleeping. Urgent messages had been dispatched to the King’s numerous personal physicians, who quickly came flocking to his assistance; they were summoned regardless of distinctions of creed and politics, and they came. They ordered cupping-glasses to be applied to his shoulders forthwith, and deep scarification to be carried out, by which they succeeded in removing another eight ounces of blood. A strong antimonial emetic was administered, but as the King could be got to swallow only a small portion of it, they determined to render assistance doubly sure by a full dose of Sulphate of Zinc. Strong purgatives were given, and supplemented by a succession of oysters. The hair was shorn close, and pungent blistering agents were applied all over his head. And as though this were not enough, the red-hot cauterization was requisitioned as well.”

One of the dozen attending physicians noted with pride that “nothing was left untried”; the King graciously apologized for being “an unconscionable time a-dying.”

Meanwhile, the common citizen experienced poor health exacerbated by the many new dangers attending the Industrial Revolution. In 1775, Percivall Pott (1714-1788) pointed out that boys chimney-sweeps developed scrotal cancer, due to soot irritation. In his Condition of the Working Classes in England (1844), Friedrich Engels (1820-1895), a Manchester factory owner as well as Karl Marx’s collaborator, described workers who were “pale, lank, narrow-chested, hollow-eyed ghosts,” cooped up in houses that were mere “kennels to sleep and die in.” In 1832, the Leeds physician Charles Turner Thackrah (1795-1833) published The Effects of Arts, Trades, and Professions on Health and Longevity, documenting the diseases and disabilities of various occupations. Apart from factory workers, among those most exposed to harmful substances were cornmillers, maltsters, coffee-roasters, snuff-makers, rag-pickers, papermakers and feather-dressers. Tailors were so subject to anal fistulas that they set up their own “fistula clubs.” Thackrah’s overall verdict was bleak: “Not 10% of the inhabitants of large towns enjoy full health.”

The single worst malady cultivated in populous cities was tuberculosis (TB), a disease characterized by fever, night sweats, and hemoptysis (coughing up blood), called “consumption” because victims were almost literally consumed. By 1800 TB was proclaimed the most common disease, and in 1815 Thomas Young (1773-1829) surmised that tuberculosis brought a premature death to one in four in the general population. Autopsies conducted in the chief Paris hospitals recorded TB as the cause of death in some 40% of cases. In the continental U.S. as late as 1890, the corresponding percentage was about 13%, and tuberculosis was still the leading cause of death (Table 1.2), though the data are partially suspect because cases of lung cancer were sometimes reported as “consumption.”

One important 18th century improvement in internal medicine which decisively saved many lives was the introduction of inoculation and vaccination against smallpox. Smallpox, “the speckled monster,” had become virulent throughout Europe and in bad years accounted for about 10% of all deaths; Queen Mary of England (1662-1694), Louis XV of France (1710-1774), and Queen Anne’s son and sole surviving direct heir (d. 1700) died of it. Doctors had long been aware of the immunizing properties of an attack, and smallpox inoculation seems to have been known and practiced for centuries at a folk level throughout Arabia, North Africa, Persia, and India. Reports of a more elaborate Chinese method, involving the insertion of a suitable infected swab of cotton inside the patient’s nostril, reached London in 1700. But it was a report from Mary Wortley Montagu (1689-1762), wife of the British consul in Constantinople, that Turkish women held smallpox parties at which they routinely performed inoculations with the aim to induce a mild dose so as to confer lifelong protection without poxmarking, that hastened acceptance in the rural medical community. The usual method was to transfer the infection by introducing matter from a smallpox pustule into a slight wound made in the patient’s skin. Occasionally the patient developed a severe case of smallpox from such treatment, and some died. But usually the symptoms were slight — a few score of pox only — and immunity proved equivalent to that resulting from contracting the disease naturally.

Edward Jenner (1749-1823), an English country doctor who performed such inoculations, noticed that cowpox, a cattle disease occasionally contracted by humans, particularly dairy maids, also conferred immunity against smallpox. Suspecting that it might be possible to produce this immunity by arm-to-arm inoculation from

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>1890 Death Rate per 100,000</th>
<th>1990 Death Rate per 100,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumption (TB)</td>
<td>245.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>186.9</td>
<td>31.1</td>
</tr>
<tr>
<td>Heart disease</td>
<td>121.8</td>
<td>310.4</td>
</tr>
<tr>
<td>Diarrheal diseases</td>
<td>104.1</td>
<td>~ 0</td>
</tr>
<tr>
<td>Debiliry and atrophy</td>
<td>88.6</td>
<td>N/A</td>
</tr>
<tr>
<td>Infantile cholera</td>
<td>79.7</td>
<td>~ 0</td>
</tr>
<tr>
<td>Bronchitis</td>
<td>74.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Diphtheria</td>
<td>70.1</td>
<td>~ 0</td>
</tr>
<tr>
<td>Kidney diseases</td>
<td>59.7</td>
<td>8.8</td>
</tr>
<tr>
<td>Meningitis</td>
<td>49.1</td>
<td>0.4</td>
</tr>
<tr>
<td>Apoplexy (stroke)</td>
<td>49.0</td>
<td>57.9</td>
</tr>
<tr>
<td>Cancer</td>
<td>47.9</td>
<td>203.2</td>
</tr>
<tr>
<td>Typhoid Fever</td>
<td>46.3</td>
<td>~ 0</td>
</tr>
<tr>
<td>Old age</td>
<td>44.9</td>
<td>N/A</td>
</tr>
<tr>
<td>Brain diseases</td>
<td>30.9</td>
<td>N/A</td>
</tr>
<tr>
<td>Croup</td>
<td>27.6</td>
<td>~ 0</td>
</tr>
<tr>
<td>Malaria</td>
<td>19.2</td>
<td>0.02</td>
</tr>
<tr>
<td>Influenza</td>
<td>6.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Other</td>
<td>~ 488.2</td>
<td>248.8</td>
</tr>
<tr>
<td>All Causes</td>
<td>~ 1840.0</td>
<td>863.8</td>
</tr>
</tbody>
</table>
the cowpox pustule, and surmising it would be safer than inoculation from smallpox pustules directly, since in humans cowpox was benign, Jenner tried the experiment, and it worked. In 1798 he published his discovery in *An Inquiry into the Causes and Effects of the Variolae Vaccinæ*. By 1799 over 5000 individuals had been vaccinated in England and abroad the practice was taken up remarkably swiftly, being made compulsory in Sweden and supported by Napoleon, who had his army vaccinated. For the first time in history, organized medicine began to contribute to human population growth in a statistically significant fashion.

The Napoleonic Wars spurred new attempts to treat battle-wounded soldiers in a more timely manner. Traditionally, the wounded were left on the field unattended until the end of the battle, but Napoleon’s chief surgeon Dominique Jean Larrey (1766–1842) introduced the use of rudimentary carts called ambulances volantes (little more than horsedrawn rickshaws) as the first “ambulances” to evacuate and transport wounded soldiers from the field to nearby aid stations, even while the battle raged on. In 1792, Larrey organized the first air evacuation, by hot air balloon.

The use of ambulances didn’t catch on until the late 1800s; until then, anyone injured in the streets of Paris, London, New York or Boston depended on the kindness of strangers or a nearby business shop for a place to rest until a doctor could be summoned. The first “modern” ambulance appeared in the city of Cincinnati in 1865, but the first true city ambulance system was developed in association with Bellevue Hospital in New York City in 1866, receiving 1500 requests for transport in its first three years of service. These horse-drawn ambulances, usually provided by local mortuaries, carried a driver and a surgeon, who was on board mainly to pronounce a patient’s death at the scene or upon arrival at the hospital, since little could be done for the seriously injured. The surgeon kept meticulous notes on the ride, recording the time of the call, transport and arrival times, and any other details that “a coroner’s jury might possibly require.”

The period also saw the development of many simple diagnostic tools that are taken for granted today. For example, a French physician, Rene Theophile Hyacinthe Laennec (1781–1826), in his *Treatise On Mediastal Auscultation* (1819), described pathological lesions found in the chest at autopsy and showed how they correlated with disease detected in living patients, establishing for the first time the modern concept of clinicopathological correlation, the cornerstone of modern diagnosis. Laennec also developed an instrument that he named a “stethoscope” to assist him in his examination of patients, especially female patients, against whose chests the direct placing of a male ear was socially taboo. The original device was a straight wooden tube; by mid-century rubber tubing was introduced to create a flexible monaural stethoscope, and in 1852 an American physician, George P. Cammann (1804–1863), devised our familiar two-ear instrument.

The clinical thermometer is of like vintage. Galileo (1564–1642) invented the first thermometer in the late 16th century, but it was not applied to medicine. Early medical thermometers in the 18th century were a foot long and difficult to use at the bedside, and were reportedly carried under the arm “as one might carry a gun.” The short clinical thermometer was devised by Sir Clifford Allbutt (1836–1925) in the 1860s, and was widely used during the American Civil War (1861–1865). The classical work in temperature diagnostics was Carl Wunderlich’s (1815–1877) *The Temperature in Diseases*, published in 1868, which presented data on nearly 25,000 patients and analyzed temperature variations in 32 diseases, showing that temperature readings could differentiate fevers. Other devices emerged later to measure pulse and blood pressure. In 1854, Karl Vierordt (1818–1884) created the sphygmograph, a pulse recorder usable for routine monitoring on humans. Blood pressure was measured using the familiar inflatable band wrapped around the upper arm, called the sphygmomanometer, whose basic design was established in 1896 by Scipione Riva-Rocci (1863–1937). The hypodermic syringe was invented in 1853.

Biochemistry also began to play an increasing diagnostic function. In the 18th century, Matthew Dobson (d. 1784) developed tests for diabetes. In 1827, Richard Bright (1789–1858) showed the kidney complaint subsequently called Bright’s disease could be diagnosed by a simple, single chemical test. Chemical analysis was crucial to Alfred Becquerel’s (1814–1862) urinalysis studies in 1841, establishing the average amounts of water, urea, uric acid, lactic acid, albumin, and inorganic salts secreted over 24 hours, and correlating these with various disease conditions. In 1859, Alfred Garrod (1819–1907) devised a simple chemical test pathognomonic for gout.

1.2.1.9 Fully Invasive Surgery

The fully invasive surgery of the 19th and 20th centuries rests upon the triple foundation of anatomy, anesthesia, and asepsis. Each has an interesting story, described below, presaging the emergence of the rational-scientific approach to medical treatment.

1.2.1.9.1 Anatomy

In surgical practice, lack of accurate anatomical knowledge had long been a great obstacle. Dissection of the human body was first practiced systematically at the great medical school of Alexandria, which flourished from about 300 BC until the death of the last ruler of Ptolemaic Egypt, Cleopatra, in 30 BC. After the decline of Alexandria, dissection was carried on at a few other centers in the Middle East, but in the first two centuries of the Christian era human bodies were replaced on the dissection table by those of apes and other animals. The anatomical knowledge gained by Galen and others from the dissection of animals was an adequate guide for the simple operative procedures carried out at this time because the abdomen, the chest, and the head were rarely opened by the surgeon’s knife.

Anatomical demonstrations of a kind were introduced into some Italian medical schools early in the 14th century but their main purpose was to serve as an aid in memorizing what Galen had written a thousand years before. During the late Middle Ages, and for a long time after, the procedure was for the professor to read from some second- or third-hand manuscript version of Galen while a demonstrator pointed to the part under discussion with a wand. As the text of Galen was often based on the dissection of an ape or pig, there were naturally many occasions when the anatomical structure under examination did not correspond with Galen’s description.

Some of the great artists took up the scientific study of anatomy, but the true founder of modern anatomy was Andreas Vesalius (1514–1564), a native of Brussels who studied medicine in Paris and later taught surgery and anatomy at Padua and Bologna. Vesalius was filled with a passionate desire for anatomical study, and many stories are told of the great risks which he took in obtaining material, including, it is reported, graverobbing. On one occasion he stole the skeleton of a criminal which was hanging on a gallows outside the city wall of Louvain and the trophy proved of great value in his studies. His public dissections during his seven years in Padua drew enormous crowds of students. In 1543 he published his 355-page great work, *De Humani Corporis Fabrica Libri Septem (On the Fabric of the Human Body)*, the outstanding and precise illustrations in which were copied and plagiarized by scholars for more than 100 years, and could be used for teaching even today. For the first time in the
history of medicine, doctors had at their disposal a detailed and accurate anatomical text with illustrations from the hand of a great artist. This book is the foundation stone of anatomy — indeed of all modern medicine — because without a sound knowledge of the structure of the body there can be no real understanding of the body's functions in health and disease.

A tremendous shortage of dead human bodies for dissection, to teach anatomy, remained a problem for centuries. In Edinburgh in 1827, an old man died in William Hare’s (1792-1870) boarding-house; assisted by his lodger William Burke (1792-1829), the two men bypassed the grave and sold the body directly to anatomists. Spurred by success, they turned to murder, luring victims and suffocating them to avoid signs of violence. Sixteen were done to death and their bodies sold, fetching 7 pounds apiece, before Burke and Hare were brought to justice in 1829. Hare turned King’s evidence and Burke was hanged. The last cadaver was found in the dissecting room of a respected anatomist, Robert Knox (1791-1862). Despite his cries of innocence, an incensed crowd burned down his house and he fled to London, his career in ruins, and eventually Knox died in obscurity.

To help pass the English Anatomy Act of 1832 (which awarded to doctors the “unclaimed bodies” of paupers) and to dispel public concern about dissection, the body of the great English philosopher and jurist, Jeremy Bentham (1748-1832), in accordance with his directions, was dissected in the presence of his friends. The skeleton was then reconstructed, supplied with a wax head to replace the original (which had been mumified), dressed in Bentham’s own clothes and set upright in a glass-fronted case. Both this effigy and the head are preserved at University College, London, to this day.

1.2.1.9.2 Anesthesia

Pain did not prevent surgery but made it almost unbearable, and the accompanying trauma often proved dangerous. Before the anesthetic era, which commenced in the late 1840s, surgical operations were agonizing. Of course, if the patient had a broken leg or a major wound, there was no choice but submit to a surgeon's knife. But non-emergency elective operations would only be undergone if the condition itself was so painful or life-threatening that the victim could even consider allowing surgery. In this event, patients would choose a surgeon with the best reputation for quickness — a limb might be removed or a bladder stone evacuated in a couple of minutes. Progress in technique was often rapid. For example, in 1824, Astley Cooper (1768-1841) took 20 minutes to amputate a leg through the hip joint; ten years later, James Syme (1799-1870) was doing it in just 90 seconds.

In pre-anesthetic days, operations were rushed through at lightning speed and under conditions of appalling difficulty. The most hardened surgeons had to steel themselves to perform operations which they knew would cause agony to their patients and nerve-wracking distress to themselves. It is hard for any 20th century inhabitant of an industrialized nation to imagine what a major surgical operation must have meant to the patient in the days before anesthesia. The following is a personal account by a male patient who suffered the removal of a stone from the bladder by Henry Cline (1750-1827), surgeon to St. Thomas’s Hospital and one of the leading operators of the day, on 30 December 1811, just three decades before the widespread adoption of anesthesia:

“My habit and constitution being good it required little preparation of body, and my mind was made up. When all parties had arrived I retired to my room for a minute, bent my knee in silent adoration and submission, and returning to the surgeons conducted them to the apartment in which the preparations had been made. The bandages &c. having been adjusted I was prepared to receive a shock of pain of extreme violence and so much had I overrated it, that the first incision did not even make me wince although I had declared that it was not my intention to restrain such impulse, convinced that such effort of restraint could only lead to additional exhaustion. At subsequent moments, therefore I did cry out under the pain, but was allowed to have gone through the operation with great firmness.”

“The forcing up of the staff prior to the introduction of the gorget gave me the first real pain, but this instantly subsided after the incision of the bladder was made, the rush of urine appeared to relieve it and soothe the wound.”

“When the forceps was introduced the pain was again very considerable and every movement of the instrument in endeavoring to find the stone increased. Still, however, my mind was firm and confident, and, although anxious, I was yet alive to what was going on. After several ineffectual attempts to grasp the stone I heard the operator say in the lowest whisper, ”It is a little awkward, it lies under my hand. Give me the curved forceps, upon which he withdrew the others. Here, I think, I asked if there was anything wrong — or something to that purport — and was reassured by the reply conveyed in the kindest manner, ”Be patient, Sir, it will soon be over.” When the other forceps was introduced I had again to undergo the searching for the stone and heard Mr. Cline say, ”I have got it.” I had probably by this time conceived that the worst was over; but when the necessary force was applied to withdraw the stone the sensation was such as I cannot find words to describe. In addition to the positive pain there was something peculiar in the feel. The bladder embraced the stone as firmly as the stone was itself grasped by the forceps; it seemed as if the whole organ was about to be torn out. The duration, however, of this really trying part of the operation was short and when the words ”Now, Sir, it is all over” struck my ear, the ejaculation of “Thank God! Thank God!” was uttered with a fervency and fulness of heart which can only be conceived...” I never heard what was the precise duration of the operation but conceive it to have been between twelve and fifteen minutes.”

And now, a woman’s point of view. In 1810, Napoleon’s famed military doctor Dominique Jean Larrey performed a radical mastectomy without anesthetic on the popular female novelist Fanny Burney (1752-1840). Burney later wrote a long account of the operation which, despite the excruciating agony, she believed had nevertheless saved her life:

“M. Dubois placed me upon the Mattress, & spread a cambric handkerchief upon my face. It was transparent, however; & I saw through it that the Bed stead was instantly surrounded by the 7 men and my nurse. I refused to be held; but when, bright through the cambric, I saw the glitter of polished steel — I closed my eyes...

“Yet — when the dreadful steel was plunged into the breast — cutting through veins — arteries — flesh — nerves — I needed no injunctions not to restrain my cries. I began a scream that lasted unintermittingly during the whole time of the incision — & I almost marvel that it rings not in my Ears still so excruciating was the agony.

“When the wound was made, & the instrument was withdrawn, the pain seemed undiminished, for the air that suddenly rushed into those delicate parts felt like a mass of minute but sharp & forked poniards [small pointed daggers], that were tearing at the edges of the wound. But when again I felt the instrument, describing a curve, cutting against the grain, if I may so say, while the flesh resisted in a manner so forcible as to oppose & tire the hand of the operator, who was forced to change from the right to the left — then, indeed, I thought I must have expired, I attempted no more to open my eyes...The instrument the second time withdrawn, I concluded the operation over — Oh no! presently the terrible cutting was renewed — & worse than ever, to separate the bottom, the foundation of the dreadful gland from the parts to which it adhered...yet again all was not over...
This chilling account continued for several more pages. When did the practice of anesthesia begin? The Herbal of Dioscorides (ca. 40-90 AD) contains specific directions for giving a decoction (boiled extraction) of mandragora “to such as shall be cut or cauterized,” one of the earliest references to surgical anesthesia. Bernard de Gordon (ca. 1260-1308) tells us that the Salernitans rubbed up poppy seed and henbane and used them as a plaster to deaden the sensibility of a part to be cauterized. Arnold of Villanova (1235-1311) gives the following recipe:

“To produce sleep so profound that the patient may be cut and will feel nothing, as though he were dead, take of opium, mandragora bark, and henbane root equal parts, pound them together and mix with water. When you want to sew or cut a man, dip a rag in this and put it to his forehead and nostrils. He will soon sleep so deeply that you may do what you will. To wake him up, dip the rag in strong vinegar.”

Some of the surgical textbooks of the Middle Ages contain references to anesthetic sponges which were prepared by soaking them in various herbs reputed to have soporific properties. The favorite herb for this purpose was the mandrake. Another simple method of producing analgesia used intermittently from early times was compression. Writing in 1564 about the various uses of the tourniquet, the French surgeon Ambrose Pare noted that “it much dulls the sense of the part by stupefying it.” Amusingly, reliable witnesses claim that as late as the 19th century, a method of anesthesia practiced at the Imperial Court of China was to “knock the patient out by a sudden blow on the jaw.”

The almost complete absence of any mention of pain-relieving drugs in medical literature of the post-medieval period is not easy to explain. It is, however, probable that the action of crude concoctions employed in early times was very uncertain, and that drugged sleep often ended in death. The active ingredients of the many herbs used in medicine had not been isolated and it would have been very difficult to regulate dosages reliably.

The story of inhalation anesthesia begins in 1799 when Sir Humphry Davy (1778-1829) recorded the effects produced by the inhalation of nitrous oxide. He breathed various concentrations of the gas and noted that a headache and the pain associated with the cutting of a wisdom tooth were relieved. Demonstrations of the effects of nitrous oxide were frequently given, bladders filled with “laughing gas” being passed around at lectures. Gas inhalation became a popular party game. A little book of 1839 contains a description of the “irresistibly ridiculous” sight of a large room filled with persons each of whom was sucking from a bladder. As the gas began to take effect, “some jumped over the tables and chairs; some were bent on making speeches; some were very much inclined to fight; and one young gentleman persisted in attempting to kiss the ladies.” At about the same time, “ether frolics” became equally popular.

In January 1842, William E. Clarke (b. 1818), a young American physician of Rochester, New York, who had acquired some knowledge of ether by attendance at ether frolics, administered the chemical on a towel to a Miss Hobbie who then had one of her teeth extracted painlessly. So far as is known this was the first use of ether for a dental or surgical operation. In March 1842, Crawford W. Long (1815-1878) of Danielville, Georgia, who had also witnessed ether frolics “enjoying sweet kisses from the girls,” successfully removed a small tumor from the neck of a patient under the influence of ether. Horace Wells (1815-1848), a dentist of Hartford, Connecticut, attended a public demonstration of the effects of nitrous oxide in December 1844, and the day after he administered the gas to himself and had one of his own teeth pulled out by a colleague. Afterwards, Wells wrote: “I didn’t feel it so much as the prick of a pin.” His former partner, William Thomas Green Morton (1819-1868), also introduced ether into his dental practice in 1846. By February 1847, the Lancet and other medical journals were reporting anesthetic operations from all parts of Great Britain, and ether had been used in most European countries. In June 1847 the news reached South Africa and a leg was amputated painlessly by W.G. Atherstone of Grahamstown.

Sir James Young Simpson (1811-1870), professor of surgery at Edinburgh, introduced chloroform in 1847. Tradition has it that Simpson had been testing chemicals with his assistants when somebody upset a bottle of chloroform; upon bringing in dinner, Simpson’s wife found them all asleep. Unlike ether, chloroform did not irritate the lungs or cause vomiting, and was powerful and easy to administer. In April 1853, Queen Victoria (1819-1901) took chloroform for the birth of Prince Leopold; John Snow (1813-1858) administered the anesthetic. Protests followed — some objections were religious (e.g., the Bible taught that women were supposed to bring forth in “travail and pain”*) but most were medical, putatively on grounds of safety but ringing with naturophilia (Section 1.3.4): “In no case could it be justifiable to administer chloroform in perfectly ordinary labor,” complained the Lancet. Incredibly, some early 19th century surgeons believed that using anesthesia for an operation would weaken a patient’s character.

Of course, general anesthesia could indeed prove dangerous, and deep unconsciousness was unnecessary for less invasive procedures, so the search was on for substances that would numb a particular area for local surgery. Cocaine was isolated in 1859 and was first used in ophthalmologic procedures by Carl Koller (1857-1944). Cocaine became the first widely-used local anesthetic, synthesized in 1885 by the Merck drug company.

1.2.1.9.3 Germ Theory and Antisepsis

Before the middle of the 19th century, pain had been banished from surgical operations, but one grave danger still faced every patient submitting himself to the surgeon’s knife. This was the ever-present risk of sepsis (infection). Hospital diseases such as erysipelas, pyemia, septicaemia and gangrene, were rife. In the 1850s the death rate after amputations varied from 25%-60% in different countries and in military practice it reached the appalling figure of 75%-90%. The first ovariotomies, which were the first abdominal operations performed on a fairly large scale, had a mortality rate of more than 30% even in the most expert hands. That all these diseases were due to some form of “contagion” had long been suspected, but the general view was that whatever agent was responsible was generated spontaneously in wounds. Alternatively, it was theorized that air itself was responsible for suppuration and many attempts were made to exclude the air from wounds by means of elaborate dressings.

Some medical men had postulated the existence of minute particles in the air which carried contagion, the so-called “germ theory.” In 1546, Girolamo Fracastoro of Verona (1478-1553) proposed “seminaria, the seeds of disease which multiply rapidly and propagate their life,” minute bodies passing unseen from the infector to the infected by contact, by clothing or utensils, and by infection at a distance through the air.

What made the germ theory of contagion so difficult to accept was that no one could see the supposed microbes. Magnifying lenses were used in ancient times, and by the beginning of the 17th century

*Others pointed to Genesis 2:21, where Adam is put to sleep as the rib is taken for Eve.
they had been combined in a tube to make the compound microscope. The first man to employ the microscope in investigating the causes of diseases was probably Anthonius Kircher (1601-1680), a learned Jesuit priest. In 1658 Kircher described experiments upon the nature of putrefaction, showing how maggots and other living creatures developed in decaying matter. He also claimed to have found in the blood of plague-stricken patients “countless masses of small worms, invisible to the naked eye.” It is impossible that he could have seen the plague bacillus with the very low power microscopes at his disposal, but he may have seen some of the larger microorganisms and his statements about the doctrine of contagion are even more explicit than those of Fracastoro.

The great pioneer of modern microscopy was Antoni van Leeuwenhoek (1632-1723), a Dutch linen draper, who ground his own lenses and made hundreds of microscopes. Few of his instruments provided a magnification of more than 160X but he is generally credited as making the first observations of germs, reported in his communications to the Royal Society in London. Leeuwenhoek was the first to describe spermatozoa and he gave the first complete account of the red blood corpuscles in 1674; he also found that the film from his own teeth contained “little animals, more numerous than all the people in the Netherlands.”

In 1847, Ignaz Philipp Semmelweis (1818-1865), an assistant at the Vienna General Hospital in the maternity clinic (the world’s largest at the time), was investigating the 29% postpartum mortality rate among women in Ward One, where births were handled by medical students, vs. a 3% rate in Ward Two where births were handled by midwifery pupils. Semmelweis noticed that the appearances of these deaths from puerperal fever looked the same as those observed in the body of an older colleague, forensic medicine Professor Jakob Kolletschka (1803-1847), who had died from a dissection wound suffered in the clinic. He correctly surmised that the postpartum deaths were caused by infection from “putrid particles” carried on the hands of medical students who often shuttled back and forth between the labor wards, the obstetrical clinic, and the autopsy room where dissections were performed. Semmelweis instituted a simple routine of hand-washing with water solutions of chloride of lime, which promptly reduced mortality to 1.27%. Unfortunately, his ideas were met with fierce opposition by the conservative medical community in Vienna, who subjected him to laughter and ridicule. In disgust, Semmelweis left Austria for Budapest. There he became head of the obstetrical division of St. Rochus Hospital, where puerperal fever mortality rates were reduced to below 1% after Semmelweis introduced chlorinated water disinfection.

The man who elucidated the true nature of infection, founded the science of bacteriology, and paved the way for Lister and the antiseptic system in surgery was Louis Pasteur (1822-1895). Pasteur was led to his great discoveries regarding bacteria and other microorganisms by his investigations into the process of fermentation. He showed conclusively that fermentation was brought about by some external agent entering the wine. He proved by painstaking experiments under rigorously controlled conditions that heat and fluids like blood did not putrefy if they were kept in such a way that all air was excluded from them. By taking samples of air at different levels Pasteur showed that contamination became less with increasing altitude. Then he proved that the contaminating agents were living organisms (bacteria) which were everywhere — in every room, in the air, on every article of clothing, on furniture, on the ground, and on the skin. He showed that putrefaction was caused by the presence of bacteria and that this applied to putrefaction in foods (milk, wine, and meat), urine, and in wounds.

The application of Pasteur’s discoveries to surgical practice was the work of Joseph Lister (1827-1912), a young English surgeon who had concluded that it must also be bacteria that caused the suppuration, pus and gangrene which plagued the surgical wards of those days. He determined to prevent the access of organisms by killing them in or on the surface of the wound. Pasteur had shown that heat could kill microbes, but it was impossible to apply heat to a wound without burning the patient, so some chemical substance had to be found. After trying various chemical agents he finally selected carbolic acid, and he insisted that everything which touched the wound, the dressings, the instruments and the fingers, should be treated with this antiseptic. He even produced an antiseptic atmosphere by means of a carbolic spray. The clinical results of Lister’s first antiseptic system in 1865 included 11 compound fracture cases with only one death, a 9% mortality rate, marking a watershed between the primitive and modern eras of surgery.

Throughout the 19th century, the arguments continued as to whether microorganisms seen in a sick patient were merely coincidental with the illness, or resulted from the changes brought about by the illness itself. In 1882 the German microbiologist Robert Koch (1843-1910) formulated three famous postulates to guide scientists searching for disease-causing microbes. Koch argued that to prove an organism causes a disease, microbiologists must show that the organism occurs in every case of the disease; that it is never found as a harmless parasite associated with another disease; and that once the organism is isolated from the body and grown in laboratory culture, it can be introduced into a new host and produce the disease again. (An oft-stated fourth postulate, that the microbe must be isolated again from the second host, was not part of Koch’s original formulation.) Koch and his pupils discovered specific bacillary causes for various diseases, including anthrax, cholera, tuberculosis, gonorrhea, diphtheria, leprosy, typhoid, trypanosomiasis, and malaria.

The theory that specific germs could cause specific disease remained contentious until the beginning of the 20th century. Many scientists of great repute rejected Koch’s conclusions, with one scientist confidently asserting that “no microbe found in the living blood of any animal was pathogenic.” In one celebrated case, Max von Pettenkofer of Bavaria (1818-1901), a distinguished 19th century experimental hygienist, induced Koch to send him a sample of his cholera vibrios culture and then wrote a letter back to Koch in 1892, as follows:

*Herr Doctor Pettenkofer presents his compliments to Herr Doctor Professor Koch and thanks him for the flask containing the so-called cholera vibrios, which he was kind enough to send. Herr Doctor Pettenkofer has now drunk the entire contents and is happy to be able to inform Herr Doctor Professor Koch that he remains in his usual good health.*

Apparently Pettenkofer, aged 74 at the time he wrote the letter, survived this cholera exposure quite well, perhaps possessing the high stomach acidity which sometimes neutralizes the bacillus, though he shot himself to death in Munich 9 years later.

Skeptics notwithstanding, microbes were key. Enormous new vistas now lay open, as surgeons could confidently make an incision through intact skin without incurring an extreme risk of wound infection. The next step was to progress beyond killing wound

*Koch’s postulates are still valuable, especially as ideals, but many diseases do not conform to them so they are no longer considered the essential basis of diagnosis that they once were.*
bacteria with chemical antiseptics, to the prevention of bacterial contamination by eliminating bacteria in the operating theater — aseptic surgery. The use of steam sterilization of instruments, dressings and gowns, the wearing of masks, caps and gloves, air filtration and the other rituals of the operating theater of today were introduced over the decades following Lister's efforts.

Anesthesia and antisepsis enabled surgeons to carry out procedures that had formerly been quite beyond them, including long operations inside the head, the abdomen and the pelvis. Theodor Billroth of Vienna (1829-1894) resected the esophagus in 1872, parts of the intestines in 1878, and the pyloric end of the stomach in 1881. Billroth also made the first complete excision of the larynx. The first successful repair of a gunshot wound on a major artery was performed in Chicago by John B. Murphy (1857-1916) in 1897; Ludwig Rehn (1849-1930) of Frankfurt am Main performed the first successful repair of a cardiac injury in Germany in the same year.

Writing in 1874, Sir John Eric Erichsen (1818-1896), Professor of Surgery at University College, London, had predicted that “the abdomen, the chest, and the brain would be forever shut from the intrusions of the wise and humane surgeon.” By the time of Erichsen’s death 22 years later, surgeons had successfully removed from patients the stomach and large parts of the intestines, a whole lung had been excised, and a brain tumor had been extirpated. These operations were not mere feats of surgical showmanship; they saved the lives and restored the health of thousands of human beings.

1.2.1.10 Cells and Tissues

The doctrine of the essential cellular nature of living things was established by 1840. Modern cell theory began in botany. The Jena botanist Matthias Schleiden (1804-1881) observed that plants were aggregates of cells, existing as self-reproducing living units. Exploring analogies between animals and plants in structure and growth, Theodor Schwann (1810-1882) took up the idea, maintaining that all these phenomena could also be demonstrated in animal structures. Thus living cells were basic to living things, and cells incorporated a nucleus and an outer membrane.

Jacob Henle (1809-85) applied cell biology to man. His three-volume Handbuch der systematischen Anatomie des Menschen (Handbook of Systematic Human Anatomy) (1866-1871) addressed the body from an architectural standpoint, describing its macro- and microscopic structure. Henle discovered kidney tubules and was the first to describe the muscular coat of the arteries, the minute anatomy of the eye and various skin structures, earning him a reputation as the Vesalius of histology (the study of tissues). Physiologists stressed organ and tissue function — Claude Bernard (1813-1878) emphasized the experimental method in establishing biological knowledge and urged that medical practice should be grounded in such knowledge.

Histology was raised to the status of an independent science by the Swiss microanatomist Albert von Kolliker (1817-1905), who wrote the first textbook on the subject, Handbuch der Gewebelehre des Menschen (Handbook of the Tissues of Man) (1852). The medical implications of cell theory were taken up by Rudolph Virchow (1821-1902), who dominated German biomedical research for half a century. Virchow extended the cell concept to diseased tissues; his Die Cellularpathologie (Cellular Pathology) (1858) analyzed such tissues from the point of view of cell formation and cell structure. Virchow initiated the idea that the body may be regarded as a “cell state in which every cell is a citizen.” Disease is often but civil war, and white cells are likened to scavengers or police. Virchow maintained that cells always arose from pre-existing cells through cellular division. Since Kolliker and Virchow, the study of the intimate structure and workings of the cells themselves, as distinct from the tissues, has become a separate science, cytology, further extended to the study of cells in disease, or cytopathology (Chapter 21).

1.2.1.11 Blood Transfusions

Vague references to blood transfusion are found in medieval writings, though physicians historically were far more concerned with taking blood out of the body than with putting it back in. For example, one story tells of an attempt to prolong the life of Pope Innocent VIII (d. 1492) by means of a blood transfusion. By one account, a Jewish physician transfused the aged Pontiff with blood from three small boys, who each received one ducat as their reward; but another account says the blood was drunk, not infused. One of the earliest proposals to transfuse blood was by Andreas Libavius (1540-1616), a physician of Halle in Saxony, in 1615.

The first serious attempts at blood transfusion were made in England and France. In 1657 Sir Christopher Wren (1632-1723) carried out numerous experiments on the injection of liquids into the veins of animals. Richard Lower (1631-1691) of Oxford carried out the first successful transfusion from artery to vein between two dogs in 1665, a feat repeated by the French physician Jean-Baptiste Denys (1625-1704), who went on to attempt transfusion of blood from one kind of animal into another. Finally, before the Royal Society in 1667, Lower transfused a human youth, 15 years of age, from a sheep. The patient was greatly improved and the only ill effect was a feeling of great heat along his arm; subsequent patients were not so lucky. Blood transfusions from lambs and calves to humans continued to be tried in the mid 17th century, but were not very successful, and in 1670 were finally forbidden by law in England.

The first known transfusion of human blood into an already moribund person was attempted by James Blundell in 1818, and on several other occasions during the 1820s, with poor results. Transfusion was carried out on a small scale during the American Civil War, but technical difficulties connected with premature clotting and the occurrence of accidents arising from the use of incompatible blood prevented the rapid acceptance of the process. The cause of many of the untoward effects of blood transfusion was finally explained in 1901 when the presence of agglutinins and iso-agglutinins in the blood was demonstrated by Karl Landsteiner (1868-1943) in Vienna, winning him the Nobel Prize in 1930; in 1907, the four main blood groups were determined by Jan Jansky (1873-1921) of Prague. Rhesus factor was later identified by two American scientists, Philip Levine (1900-1987) and Rufus Stetson (1886-1967). These advances were of fundamental importance and it became possible in the 20th century to eliminate most of the fatalities due to incompatibility, allowing blood transfusions to be practiced reliably and with uniformly good results.

1.2.1.12 20th Century Medicine

The 20th century saw more discoveries and advances in medical science than all previous centuries combined. In this period, medicine became more powerful than ever before as scientists gained knowledge of matters and processes of illness that, at the beginning of the century, were still unknown or mysterious. Unlike the mere palliatives of earlier eras, 20th century physicians could actually cure some diseases, reverse some physical traumas, and save many lives that could not be saved before.

In the first half of the 20th century, the rational scientific paradigm that arose in the 19th century was pursued and extended. Acceptance of the germ theory of infection and the discovery of leukocytes led to the rapid emergence of immunology. This allowed
medical scientists to produce protective vaccines and antisera which largely eliminated many diseases that were previously prevalent and dangerous, including whooping cough, measles, and diphtheria — the first truly effective medical treatments (Figure 1.2). Acceptance of the germ theory also led to the discovery of filterable viruses (organisms that could pass through the pores of all known filters) in the 1890s, and subsequently to the identification of viruses as specific causes of yellow fever, smallpox, typhus (a Rickettsia organism), measles, poliomyelitis, rabies, and viral meningitis. Biochemists synthesized vitamins which were recognized as essential constituents of a healthy diet, thus allowing the elimination of vitamin deficiency diseases such as scurvy, rickets, osteomalacia, beriberi, pellagra, xerophthalmia, nyctalopia, and pernicious anemia, via dietary supplements. Many metabolic diseases became treatable due to biochemical investigations; for example, the discovery of insulin in 1921 by the Canadian physiologists Sir Frederick Banting (1891-1941) and Charles Best (1899-1978) rapidly transformed diabetic patients. Medical investigations; for example, the discovery of insulin in 1921 by the Canadian physiologists Sir Frederick Banting (1891-1941) and Charles Best (1899-1978) rapidly transformed diabetic patients. The first heart transplant in man was achieved by Christiaan Barnard (b. 1922) in 1967. Results were poor at first, but the use of new anti-rejection medications for major mental illnesses including schizophrenia and manic depression.

The second pivotal event was the genetics revolution, starting with the discovery in 1953 of the information-carrying double-helix structure of DNA by Francis Crick (b. 1916) and John B. Watson (b. 1928) followed in the 1980s by the ability to chemically read the genetic code, isolate specific genes and clone them for further study. In the mid-1980s, the Human Genome Project (Chapter 20) was launched, with the objective of fully sequencing every gene in the human genome. The first phase of this project near completion as the 20th century drew to a close.

Molecular biology became the premier scientific discipline of the latter 20th century. By 1998, the compositions of organs, tissues, cells, organelles, and membranes had been identified, and the biosynthesis and catabolism of hundreds of compounds had been elucidated. The regulation of body processes was described at progressively finer levels in biochemical language. Many pharmacologic agents were finally understood in terms of specific molecular loci and mechanisms of action. Advances were particularly rapid in immunology, virology, cellular biology, peptide research, and structural biology. A beginning was made in explaining human behavior in mechanistic terms, as more and more chemical mediators and pharmacologic modifiers were discovered. In biology, these disciplines developed porous boundaries with related disciplines such as physiology, pharmacology, neurosciences, biochemistry and biophysics. All entered a phase of confluence, employing the common language of chemistry.

Thus the late 20th century is best regarded as the molecular age of basic biological science. The molecular influence pervades all the traditional disciplines underlying clinical medicine. As of February 1999, one source listed exactly 1446 genetic disorders, most of which could be linked to a specific human chromosome;
another source stated that ~4000 genetic disorders were known in 1998. There were more than 575 known abnormal human hemoglobins, and for each of these the precise structural defect in the DNA of the mutant gene could be defined. Knowledge of membrane, cytoplasmic, and nuclear receptors for hormones and drugs was exploding, with old as well as new diseases being defined in terms of receptor abnormalities — for example, type II hypercholesterolemia and nephrogenic diabetes insipidus. Recognition of opiate receptors led to the discovery of endogenous peptides (endorphins) with analgesic activity. Their localization promised further understanding of the limbic system, affective states, and addictions. Defects in a subcellular organelle, the peroxisome, were known to be responsible for a growing list of important genetic afflictions such as Refsum disease, Zellweger syndrome, and X-linked adrenoleukodystrophy. The genetic defect responsible for Huntington's disease was discovered, and in the 1990s the genetic defects responsible for many other important neurological disorders were becoming known, including forms of Alzheimer's disease, Charcot-Marie-Tooth disease, and common blinding retinal degenerative disorders such as retinitis pigmentosa, Leber's hereditary optic neuropathy, Norrie's disease, and choroideremia.

DNA sequencing techniques and restriction endonucleases permitted precise identification of the exact structural alteration of the gene in an increasing number of hereditary diseases. For example, Burkitt's lymphoma is characterized by a translocation of the distal end of the long arm of chromosome 8 to loci on chromosomes 14, 22, or 2. Gene therapy — both pharmacologic modification of specific gene action and physical replacement of damaged genetic segments — became possible in experimental systems. Complete maps of the entire genomes of 18 microbial species had been compiled and published by the end of 1998, with more than 60 others in progress. In 1992, wrote the conservative *Cecil Textbook of Medicine* "The expansion of the knowledge bank of the past quarter century justifies great optimism for the eventual control and cure of major diseases and the possible elimination of premature death from illness."

Global changes in progressive aging dysfunction were shown to be strongly related to declining secretion of growth hormone by Rudman in 1990. Shortly thereafter, anti-aging medicine was recognized as a distinct discipline and was promoted by the American Academy of Anti-Aging Medicine (A4M), a group that claimed >6000 physician and scientist members worldwide by 1998 and had held numerous conferences. In a book on the subject, Ronald Klatz, A4M's president, first comprehensively documented the implications of Rudman's work.

While biological science was vaulting forward into the molecular realm at a blistering pace, biomedical engineering lagged considerably behind, though many remarkable successes had been achieved since mid-century. Radiology expanded with sophisticated radiotherapy, ultrasound, scanning and imaging techniques (e.g., CAT, PET, NMR), with submillimeter resolution in living tissues. Surgical instruments became less damaging and less invasive; minimally invasive techniques used OK (orifice and keyhole) surgery and surgery performed under the "eye" of a scan. Fetuses could be screened for abnormality and surgery then performed upon them inside the womb, or they could be removed temporarily from the womb and then returned to continue gestating, after surgery. Prospective parents with fertility difficulties could make use of a wide range of therapies including in vitro fertilization. Implantable pacemakers, defibrillators, and ventricular assist devices were commonplace, and in 1998 artificial wearable/implantable and full/partial replacements for lungs, heart, kidney, liver, and pancreas were either available, in clinical trials, or under development (Volume III). Electron microscopes with 200,000X magnifications allowed many details of internal cellular and viral structures to be resolved as early as 1946; by 1998, atomic-force (AFM) and scanning-tunneling microscopy (STM) permitted direct tactile examination of individual biomolecules in fixed cells.

As the 20th century drew to a close, a few preliminary efforts had been made to apply the molecular approach to medical diagnostic and clinical tools. Biotechnology was one avenue being pursued, with the rational design of artificial enzymes and specified-ligand binding sites having already been achieved in certain limited cases, and the beginnings of gene therapy as noted earlier. Carbon fullerenes (Section 2.3.2) had been used to create a water-soluble inhibitor of HIV protease, and had other biological applications. AFM-based force-amplified biological sensors could detect defined biological species such as cells, proteins, toxins, and DNA at concentrations as low as 10^-18 M (-1/mm 3), and automated laboratory systems for sorting and handling individual cells and viruses were commonplace. Biotech companies such as Physiome Sciences of Princeton NJ had developed three-dimensional computer models of the heart and other organs. Physiome's heart model was based on detailed molecular, biochemical, cellular and anatomical information, including submodels of all the different cell types found in the heart embodying knowledge of the function of each cell type in healthy and diseased hearts, and information on gene function and the causes and effects of congestive heart failure, arrhythmias and heart attacks. Other cell biochemistry simulators included E-CELL (http://www.e-cell.org) and the Virtual Cell (http://www.ncr.am.uchc.edu). There was also much progress and interest in nanostructure analysis and nanomaterials fabrication for medical and biological purposes, with research groups emerging at major universities such as the Cornell Nanofabrication Facility, the University of Michigan Center for Biologic Nanotechnology, the Rice University Center for Nanoscale Science and Technology, the CalTech Materials and Process Simulation Center, the Washington University Nanotechnology Center (St. Louis, MO), the USC Laboratory for Molecular Robotics, the UCLA Exotic Materials Center, the Institute for Molecular Medicine of the University of Oxford, and at many biotechnology-oriented corporations such as Nanogen and Affymetrix.

However, the greatest medical revolution of all awaits the ability to engineer and fabricate whole devices and systems at the molecular scale. Along this course lies nanotechnology and molecular manufacturing, a deep well from which nanomedicine will inevitably spring.

1.2.1.13 21st Century Medicine

It is always somewhat presumptuous to attempt to predict the future, but in this case we are on solid ground because most of the prerequisite historical processes are already in motion and all of them appear to be clearly pointing in the same direction.

Medical historian Roy Porter notes that the 19th century saw the establishment of what we think of as scientific medicine. From about the middle of that century the textbooks and the attitudes they reveal are recognizable as not being very different from modern ones. Before that, medical books were clearly written to address a different mind-set.

But human health is fundamentally biological, and biology is fundamentally molecular. As a result, throughout the 20th century scientific medicine began its transformation from a merely rational basis to a fully molecular basis. First, antibiotics that interfered with pathogens at the molecular level were introduced. Next, the ongoing
revolutions in genomics, proteomics and bioinformatics\textsuperscript{2321} provide detailed and precise knowledge of the workings of the human body at the molecular level. Our understanding of life advanced from organs, to tissues, to cells, and finally to molecules, in the 20th century. By the early 21st century, the entire human genome will be mapped. This map will inferentially incorporate a complete catalog of all human proteins, lipids, carbohydrates, nucleoproteins and other molecules, including full sequence, structure, and much functional information. Only some systemic functional knowledge, particularly neurological, may still be lacking by that time.

This deep molecular familiarity with the human body, along with simultaneous nanotechnological engineering advances (Chapter 2), will set the stage for a shift from today's molecular scientific medicine in which fundamental new discoveries are constantly being made, to a molecular technologic medicine in which the molecular basis of life, by then well-known, is manipulated to produce specific desired results. The comprehensive knowledge of human molecular structure so painstakingly acquired during the 20th and early 21st centuries will be used in the 21st century to design medically-active microscopic machines. These machines, rather than being tasked primarily with voyages of pure discovery, will instead most often be sent on missions of cellular inspection, repair, and reconstruction.

In the coming century, the principal focus will shift from medical science to medical engineering. Nanomedicine will involve designing and building a vast proliferation of incredibly efficacious molecular devices, and then deploying these devices in patients to establish and maintain a continuous state of human healthiness.

The very earliest nanotechnology-based biomedical systems may be used to help resolve many difficult scientific questions that remain. They may also be employed to assist in the brute-force analysis of the most difficult three-dimensional structures among the 100,000-odd proteins of which the human body is comprised, or to help ascertain the precise function of each such protein. But much of this effort should be complete within the next 20-30 years because the reference human body has a finite parts list, and these parts are already being sequenced, geometric and archived at an ever-increasing pace. Once these parts are known, then the reference human being as a biological system is at least physically specified to completeness at the molecular level. Thereafter, nanotechnology-based discovery will consist principally of examining a particular sick or injured patient to determine how he or she deviates from molecular reference structures, with the physician then interpreting these deviations in light of their possible contribution to, or detraction from, the general health and the explicit preferences of the patient.
In brief, nanomedicine will employ molecular machine systems to address medical problems, and will use molecular knowledge to maintain human health at the molecular scale.

1.2.2 Volitional Normative Model of Disease

What, exactly, is "medicine"? Dictionaries give several definitions, ranging from the very restrictive to the most general, as follows: "a drug or remedy"; "any substance used for treating disease"; "any drug or other substance used in treating disease, healing, or relieving pain"; "in a restricted sense, that branch of the healing art dealing with internal diseases"; "treatment of disease by medical, as distinguished from surgical, treatment"; "the branch of this science and art that makes use of drugs, diet, etc., as distinguished especially from surgery and obstetrics"; the study and treatment of general diseases or those affecting the internal parts of the body; "the science of treating disease, the healing art"; "the art and science of preventing or curing disease"; "the department of knowledge and practice dealing with disease and its treatment". Or, most generally, "the science and art of diagnosing, treating, curing, and preventing disease, relieving pain, and improving and preserving health". In this book, we shall adopt the latter, maximally-inclusive, definition of "medicine" (Figure 1.3).

Reviewing Figure 1.3, the contemporary physician might at first be inclined to relegate molecular approaches to some minor subfield, perhaps "nanoanalytics," "nanogenomics," or "nanotherapeutics." This would be a serious mistake, because the application of molecular approaches to health care will significantly impact virtually every category of laboratory and clinical practice across the board. Thus we are led to the broadest possible conception of nanomedicine as "the science and technology of diagnosing, treating, and preventing disease and traumatic injury, of relieving pain, and of preserving and improving human health, using molecular tools and molecular knowledge of the human body."

This brings us to the question of "disease," a complex term whose meaning is still hotly debated among medical academics. Figure 1.4 shows the results of a survey of four different groups of people who were read a list of common diagnostic terms and then asked if they would rate the condition as a disease. Illnesses due to microorganisms, or conditions in which the doctor's contribution to the diagnosis was important, were most likely to be called a disease, but if the cause was a known physical or chemical agent the condition was less likely to be regarded as disease; general practitioners also had the broadest definition of disease.

No less than eight different types of disease concepts are held by at least some people currently engaging in clinical reasoning and practice, including:

1. Disease Nominalism — A disease is whatever physicians say is a disease. This approach avoids understanding and forestalls inquiry, rather than furthering it.

2. Disease Relativism — A disease is identified or labeled in accordance with explicit or implicit social norms and values at a particular time. In 19th century Japan, for example, armpit odor was considered a disease and its treatment constituted a medical specialty. Similarly, 19th-century Western culture regarded masturbation as a disease, and in the 18th century, some conveniently identified a disease called drapetomania, the "abnormally strong and irrational desire of a slave to be free." Various non-Western cultures having widespread parasitic infection may consider the lack of infection to be abnormal, thus not regarding those who are infected as suffering from disease.

3. Sociocultural Disease — Societies may possess a concept of disease that differs from the concepts of other societies, but the concept may also differ from that held by medical practitioners within the society itself. For instance, hypercholesterolemia is regarded as a disease condition by doctors but not by the lay public; medical treatment may be justified, but persons with hypercholesterolemia may not seek treatment, even when told of the condition. Conversely, there may be sociocultural pressure to recognize a particular condition as a disease requiring treatment, such as alcoholism and gambling.

4. Statistical Disease — A condition is a disease when it is abnormal, where abnormal is defined as a specific deviation from a statistically-defined norm. This approach has many flaws. For example, a statistical concept makes it impossible to regard an entire population as having a disease. Thus tooth decay, which is virtually universal in humans, is not abnormal; those lacking it are abnormal, thus are "diseased" by this definition. More reasonably, a future highly-aseptic society might regard bacterium-infested 20th century humans (who contain in their bodies more foreign microbes than native cells; Section 8.5.1) as massively infected. Another flaw is that many statistical measurables such as body temperature and blood pressure are continuous variables with bell-shaped distributions, so cutoff thresholds between "normal" and "abnormal" seem highly arbitrary.

5. Infectious Agency — Disease is caused by a microbial infectious agent. Besides excluding systemic failures of bodily systems, this view is unsatisfactory because the same agent can produce very different illnesses. For instance, infection with hemolytic Streptococcus can produce diseases as different as erysipelas and puerperal fever, and Epstein-Barr virus is implicated in diseases as varied as Burkitt's lymphoma, glandular fever, and naso-pharyngeal carcinoma.

6. Disease Realism — Diseases have a real, substantial existence regardless of social norms and values, and exist independent of whether they are discovered, named, recognized, classified, or diagnosed. Diseases are not inventions and may be identified with the operations of biological systems, providing a reductionistic account of diseases in terms of system components and subprocesses, even down to the molecular level. One major problem with this view is that theories may change over time — almost every 19th century scientific theory may change over time — almost every 19th century scientific theory was either rejected or highly modified in the 20th century. If the identification of disease is connected with theories, then a change in theories may alter what is viewed as a disease. For example, the 19th century obsession with constipation was reflected in the disease labelled "autointoxication," in which the contents of the large bowel were believed to poison the body. Consequently much unnecessary attention was paid to laxatives and purgatives and, when surgery of the abdomen became possible toward the end of the century, operations to remove the colon became fashionable in both England and America.

7. Disease Idealism — Disease is the lack of health, where health is characterized as the optimum functioning of biological systems. Every real system inevitably falls short of the optimum in its actual functioning. But by comparing large numbers of systems, we can formulate standards that a particular system ought to satisfy, in order to be the best of its kind. Thus "health" becomes a kind of
Platonic ideal that real organisms approximate, and everyone is a less than perfect physical specimen. Since we are all flawed to some extent, disease is a matter of degree, a more or less extreme variation from the normative ideal of perfect functioning. This could be combined with the statistical approach, thus characterizing disease as a statistical variation from the ideal. But this view, like the statistical, suffers from arbitrary thresholds that must be drawn to qualify a measurable function as representing a diseased condition.

8. Functional Failure — Organisms and the cells that constitute them are complex organized systems that display phenomena (e.g. homeostasis) resulting from acting upon a program of information. Programs acquired and developed during evolution, encoded in DNA, control the processes of the system. Through biomedical research, we write out the program of a process as an explicit set (or network) of instructions. There are completely self-contained “closed” genetic programs, and there are “open” genetic programs that require an interaction between the programmed system and the environment, e.g. learning or conditioning. Normal functioning is thus the operation of biologically programmed processes, e.g. natural functioning, and disease may be characterized as the failure of normal functioning. One difficulty with this view is that it enshrines the natural (Section 1.3.4) as the benchmark of health, but it is difficult to regard as diseased a natural brunette who has dyed her hair blonde in contravention of the natural program, and it is quite reasonable to regard the mere possession of an appendix as a disease condition, even though the natural program operates so as to perpetuate this troublesome organ.* A second weakness of this view is that disease is still defined against population norms of functionality, ignoring

* The vermiform appendix may have some minor immune function, but it is clearly nonessential and can kill when infected. Yet natural selection has not eliminated it. Indeed, there is evidence for positive selection due to the following accident of physiological evolution. Appendicitis results when inflammation causes swelling, compressing the artery supplying blood to the appendix. High bloodflow protects against bacterial growth, so any reduction aids infection, creating more swelling; if flow is completely cut off, bacteria multiply rapidly until the organ bursts. A slender appendix is especially susceptible, so untreated appendicitis applies positive selective pressure to maintain a larger appendix.2185

---

Fig. 1.4. Opinion survey: what is disease? (modified from Campbell, Scadding, and Roberts²²³²)
individual differences. As a perhaps overly simplistic example, 65% of all patients employ a cisterna chyli in their lower thoracic lymph duct, while 35% have no cisterna chyli (Figure 8.10) — which group has a healthy natural program, and which group is “diseased”?

The author proposes a ninth view of disease, a new alternative which seems most suitable for the nanomedical paradigm, called the “volitional normative” model of disease. As in the “disease idealism” view, the volitional normative model accepts the premise that health is the optimal functioning of biological systems. Like the “functional failure” view, the volitional normative model assumes that optimal functioning involves the operation of biologically programmed processes.

However, two important distinctions from these previous views must be made. First, in the volitional normative model, normal functioning is defined as the optimal operation of biologically programmed processes as reflected in the patient’s own individual genetic instructions, rather than of those processes which might be reflected in a generalized population average or “Platonic ideal” of such instructions; the relative function of other members of the human population is no longer determinative. Second, physical condition is regarded as a volitional state, in which the patient’s desires are a crucial element in the definition of health. This is a continuation of the current trend in which patients frequently see themselves as active partners in their own care.

In the volitional normative model, disease is characterized not just as the failure of “optimal” functioning, but rather as the failure of either (a) “optimal” functioning or (b) “desired” functioning. Thus disease may result from:

1. a failure to correctly specify desired bodily function (specification error by the patient),
2. a flawed biological program design that doesn’t meet the specifications (programming design error),
3. flawed execution of the biological program (execution error),
4. external interference by disease agents with the design or execution of the biological program (exogenous error), or
5. traumatic injury or accident (structural failure).

In the early years of nanomedicine, volitional physical states will customarily reflect “default” values which may differ only insignificantly from the patient’s original or natural biological programming. With a more mature nanomedicine, the patient may gain the ability to substitute alternative natural programs for many of his original natural programs. For example, the genes responsible for appendix morphology or for sickle cell expression might be replaced with genes that encode other phenotypes, such as the phenotype of an appendix-free cecum or a phenotype for statistically replaced with genes that encode other phenotypes, such as the phenotype of a maintenance engineer to a machine or a veterinarian to an animal."

1.2.3 Treatment Methodology

The availability of advanced nanomedical instrumentalities should not significantly alter the classical medical treatment methodology, although the patient experiences and outcomes will be greatly improved. Treatment in the nanomedical era will become faster and more accurate, efficient, and effective. In clinical practice, patient treatment customarily includes up to six distinguishable phases: examination, diagnosis, prognosis, treatment, validation, and prophylaxis. Let us consider each of these, in turn.

1.2.3.1 Examination

The first step in any treatment process is the examination of the patient, including the individual’s medical history, personal functional and structural baseline, and current complaints. In classical medicine, interview and observation have long been the cornerstone of examination. In ancient times this was limited to obvious

*It is often pointed out that sickle cell is advantageous in malaria-infested countries because the trait confers resistance to malaria. This flaw-tolerant view makes a virtue of necessity—a direct cure for malaria will undoubtedly be more efficient. Sickle cell is disadvantageous in hypoxic conditions, which is why no one with this trait can hold a civil airline pilot’s license.2257
manifestations and simple constellations of observables, such as the Hippocratic facies, the four signs of inflammation noted by Celsus, or pulse rate and fever. Clinicians recognize that the traditional taking and interpreting of oral medical histories from new patients is a subtle and complex art, although some aspects of this process might be automated using voice recognition and text preinterpretation software, somewhat easing the physician's burden.

Advancing technology has also brought a plethora of tests that contribute to accurate diagnosis, including auscultation, microscopy and clinical bacteriology in the 19th century, and radiological scanning, clinical biochemistry, genetic testing, and minimally invasive exploratory surgery in the 20th century.

In the 21st century, new tools for nanomedical testing and observation (Chapter 18) will include clinical in vivo cytography; real-time whole-body microbiotic surveys; immediate access to laboratory-quality data on the patient (e.g. blood tests such as blood counts, dissolved gases and solutes, vitamin and ion assays); physiological function and challenge tests; tissue composition including direct organelle counts in specified tissue populations; quantitative flowcharts of in cyto secondary messenger molecules, extracellular hormones and neuuropeptides; per-compartment cytogluucose inventories; and so forth. Before a proper diagnosis can be made, the physician must also establish the patient's personal functional and structural baseline against which any deviations can be noted and corrected, in keeping with the volitional normative model of disease (Section 1.2.2).

The capabilities of nanomedical testing are explored at length in Chapter 18. By way of introduction, it is instructive to think about a trivial class of test procedures that might be used to diagnose a simple infectious disease at several different levels of technological competence. Let us consider a patient who presents with signs and symptoms that are nonspecific in nature but which suggest an infectious process — e.g. nasal congestion, mild fever, discomfort and cough. The initial signs are due in part to the body's inflammatory response and in part to the infectious agent itself. The diagnostic goal is to identify the infectious agent.

In the late 20th century, the usual procedure would be to culture a sample taken from the patient, in the microbiology laboratory, using various broths, petri plates, and biochemical tests. Some infectious agents are easy to demonstrate. Beta hemolytic streptococci from a throat swab will grow overnight on a blood agar plate, and colony counts for E. coli in a urine sample are available in 24 hours. A throat culture that the lab reports as a mixed culture causes no concern, although some aspects of this process might be automated using voice recognition and text preinterpretation software, somewhat easing the physician's burden.

In 1996 when first suggested, but was regarded as a reasonable and likely future application of biotechnology in the early 21st century given rapid progress in single-molecule DNA assay techniques, DNA in the sample immediately available for hybridization to the probes. The swab is swirled in the liquid mix of the prepackaged test kit for 1 minute. The liquid is then poured through a column that separates hybridized DNA molecules (bacterial target DNA sequences bound to probe DNA) from all other debris (taking several minutes). A chemiluminescence detection system for the probes shows two of several possible colors indicating mixed infection. The diagnostic result, available in 10 minutes, indicates a Rhinovirus of a strain known to be epidemic in the geographic area. A significant superinfection with a penicillin-resistant streptococcus is also identified. With a definitive diagnosis, the patient is started on the appropriate antibiotic.

How might nanomedicine handle this test? In the nanomedical era, taking and analyzing microbial samples will be much simpler for the practitioner. Such analysis will be as quick and convenient as the electronic measurement of body temperature using a tympanic thermometer in a late 20th-century clinical office or hospital. As described in Chapter 18, the physician faces the patient and pulls from his pocket a lightweight handheld device resembling a pocket calculator. He unsnaps a self-sterilizing cordless pencil-sized probe from the side of the device and inserts the business end of the probe into the patient's opened mouth in the manner of a tongue depressor. The ramifying probe tip contains billions of nanoscale molecular assay receptors mounted on hundreds of self-guiding retractile stalks. Each receptor is sensitive to one of thousands of specific bacterial membrane or viral capsid ligands (Section 4.2). An acoustic echolocation transceiver provides gross spatial mapping. The patient says "Ahh," and a few seconds later a three-dimensional color-coded map of the throat area appears on the display panel that is held in the doctor's hand. A bright spot marks the exact location where the first samples are being taken. Underneath the color map scrolls a continuously updated microflora count, listing in the leftmost column the names of the ten most numerous microbial and viral species that have been detected, key biochemical marker codes in the middle column, and measured population counts in the right column. The number counts flip up and down a bit as the physician directs probe stalks to various locations in the pharynx to obtain a representative sampling, with special attention to sorets or any signs of exudate. After a few more seconds, the data for two of the bacterial species suddenly highlight in red, indicating the distinctive molecular signatures of specific toxins or pathological variants. One of these two species is a known, and unwelcome, hostile pathogen. The diagnosis is completed, the infectious agent is promptly exterminated (Chapter 19), and a resurvey with the probe several minutes afterwards reveals no evidence of the pathogen.

1.2.3.2 Diagnosis

Diagnosis is the determination of the cause and nature of a disease in order to provide a logical basis for treatment and prognosis. Traditionally the diagnostic process begins with a thorough history taken from the patient and a relevant physical examination. Often this sufficed to make a confident diagnosis, but the cause of some illnesses remained uncertain without recourse to additional information such as blood tests or radiological examinations. Nanotechnology-based diagnosis (Chapter 18) will consist principally of examining the patient to determine how he or she deviates from autogenous reference structures and functions, and then interpreting those deviations as healthy or unhealthy for that patient.

*A patient presents in the clinic with mild fever, nasal congestion, discomfort, and cough. A swab of his throat is taken. Instead of culture to identify abnormal microorganisms by their pattern of growth, the sample is analyzed by recombinant DNA techniques. The cotton throat swab is mixed with a cocktail of DNA probes. Enzymes that digest and release the DNA from both host cells and invading bacteria make the* DNA in the sample immediately available for hybridization to the probes. The swab is swirled in the liquid mix of the prepackaged test kit for 1 minute. The liquid is then poured through a column that separates hybridized DNA molecules (bacterial target DNA sequences bound to probe DNA) from all other debris (taking several minutes). A chemiluminescence detection system for the probes shows two of several possible colors indicating mixed infection. The diagnostic result, available in 10 minutes, indicates a Rhinovirus of a strain known to be epidemic in the geographic area. A significant superinfection with a penicillin-resistant streptococcus is also identified. With a definitive diagnosis, the patient is started on the appropriate antibiotic.

**Inexpensive biosensor devices capable of detecting Salmonella contamination in meat and poultry were already commercially available in 1998.**
In the 20th century, diagnoses frequently involved a high degree of uncertainty, largely due to the general lack of comprehensive molecular diagnostic tools. Thus diagnosis would be guided by statistical analyses; one branch of decision analysis, called utility analysis, even allowed the patient to participate in the decisionmaking process. When the correct decision is unclear, urges one textbook, it is well to remember time-honored Hippocratic aphorisms such as "first, do no harm" and "common things occur commonly." The eminent Canadian physician Sir William Osler (1849-1919) lamented that "errors of judgement must occur in the practice of an art which consists largely in balancing probabilities." Most doctors would prefer to understand the root cause of medical problems rather than adopt mere statistical approaches.

Nanomedicine tools will vastly reduce diagnostic uncertainty. Using nanomedical instrumentalities, doctors will gain access to unprecedented amounts of information about their patients including in-office comprehensive genotyping and real-time whole-body scans for particular bacterial coat markers, tumor cell antigens, mineral deposits, suspected toxins, hormone imbalances of genetic or lifestyle origin, and other specified molecules, producing three-dimensional maps of desired targets with submillimeter spatial resolution. Embedded in vivo nanomedical data archives (Section 10.2.5, Chapter 19) can provide onboard storage of regularly updated self-diagnostic scans, reducing to a minimum the need for symptomatic interview data from patients who may be unconscious, inarticulate, or verbose, who may have limited powers of self-analysis or self-observation and who may have forgotten, suppressed, or amplified descriptions of symptoms.

Of course, physicians do not require an exhaustive survey of the entire body of each patient to molecular detail to make a valid diagnosis. In any particular case, it is the function of the trained medical mind to quickly ascertain where and where not to look in molecular detail. But in the nanomedical era, powerful tools will be available to allow the practitioner to examine almost any portion of a patient in as much detail as desired, right down to the molecular level, with results available in seconds or minutes, and at reasonable cost (e.g. Section 2.4.2).

1.2.3.3 Prognosis and Treatment

Prognosis is a judgement or forecast, based upon a correct diagnosis, of the future course of a disease or injury, and of the patient’s prospects for partial or full recovery. Guttenagel identifies prognosis as "the predicted course of the reduced state of the patient’s psychosomatic freedom of action," and treatment as "the physician’s ability to intervene."

But prognosis is a function of treatment as well as disease. From the post-Hippocratic era through the 18th century, treatments were almost purely empirical and often did more harm than good. During the 19th and early 20th centuries, treatments were scientific but largely homeostatic — the medical intervention was rational but served mainly to assist the body in healing itself. Throughout the remainder of the 20th century, truly curative treatments began to rescue some patients from conditions from which their unaided bodies would not have been able to recover. Although conventional biotechnology will enable some important tissue and cellular replacement treatments by the early 21st century, nanomedicine will enable major reconstructive and restorative procedures at the tissue, cellular and molecular levels and will employ active antibiotic devices. The prognosis will almost always be good, except in cases of severe neural damage and a few other specialized circumstances. Therapeutic treatments will be selected to reverse all pathological effects of disease or injury, with a minimum of pain, discomfort, side-effects, intrusiveness and time, and with a maximum of effectiveness, efficiency, and likelihood of success, though of course some tradeoffs will always exist. Nanomedicine also will excel in the correction of molecular defects of a kind which Nature has no predesigned tools — such as the breakdown and removal of intracellular lipofuscin (for which there appear to be no natural enzymes) and the removal of indigestible waste products which interfere with neuronal axon transport.

As in Section 1.2.3.1, we may compare the therapeutic response to a simple infection at several different levels of technological competence. Consider a patient who has been diagnosed with eastern equine encephalitis, a mosquito- or tick-borne arbovirus. In the 20th century, there was no specific treatment for this disease. Care was generally supportive, with the doctor attempting to maintain the patient’s heart and lung function while the infection ran its course. The prognosis was poor. There was a 50%-75% mortality rate with frequent sequelae including seizures and paralysis, especially in children.

Biotechnologists proposed a molecular approach to therapeutics using recombinant DNA technology that was not yet possible in 1996 but was regarded as likely by the early 21st century: A patient enters the hospital with high fever and intense headaches. A spinal fluid tap is submitted to the molecular microbiology lab. After screening for several viruses, a species of equine encephalitis virus is identified that is endemic to a location recently visited by the patient. A call to the Centers for Disease Control results in the emergency delivery of a new antiviral agent. Antisense oligonucleotides are injected into the cerebral spinal fluid. These small DNA pieces bind directly to the virus and block its further proliferation. A temporary reservoir giving access to the cerebral spinal fluid is placed and infusion of this therapeutic molecular inhibitor of the virus continues for 5 days until signs of encephalitis have passed.”

The nanomedical therapy? As before, a nanomedical cure for eastern equine encephalitis may be far simpler, less painful and a great deal quicker. A single therapeutic dose consisting of ~0.1 cm³ of isotonic saline fluid containing ~10 billion active micron-size virucidal nanodevices, a 10% volumetric nanodevice suspension, is injected into the cerebrospinal fluid. Each therapeutic nanorobot has chemical sensors that can unambiguously recognize fluidborne or in cyto arbovirus particles and, once recognition has occurred, destroy them and also reverse the cellular damage. A nanorobot population of this size should be able to destroy all viral particles and effect needed repairs in at most an hour (Chapter 19), after which the devices are programmed and equipped either to eliminate themselves from the body (Chapter 16) or to be manually exfused (e.g., nanapheresis; Section 10.3.6).

1.2.3.4 Validation and Prophylaxis

A proper therapeutic protocol will include a procedure for follow-up to ensure that the prescribed treatment was correctly executed with good results. This step is often neglected in order to save costs and may be considered unimportant by some practitioners because approximately 80%-90% of all illnesses which take patients to the doctor are self-curing or self-limiting. For example, the common cold, most infectious diseases and many minor injuries are problems that usually will resolve on their own even with no treatment. In these cases the purpose of treatment is not to provide a cure, but rather to speed the healing process, improve comfort, and avoid...
Basic Capabilities • The Prospect of Nanomedicine

complications. Many nanomedical treatments will require supervision and will run quickly to completion, thus follow-up may come back into vogue. Validation may also be viewed as a post-treatment re-diagnosis to ensure that no disease remains present in the patient.

Prophylaxis is the prevention of disease, typically including patient education, immunization programs, amelioration of occupational hazards, and other preventive and public health measures. In a treatment environment that is rich in effective antibacterial instrumentalities, those microbes which survive will evolve to produce only modest or negligible symptoms that are insufficiently annoying to motivate a patient to seek professional therapeutic relief. It is well-known that bacteria can modify their behavior over time. For example, syphilis had a much more fulminating course in the Middle Ages than it has in the 20th century. Some future strain of the syphilitic microbe might produce negligible symptoms, but we should still insist on its eradication because of its potential to revert to its earlier virulence if allowed to spread unchecked in a more benign form. Preventative procedures may also be needed to discover, diagnose, and treat apparently symptomless diseases, and a variety of molecular-based physiological malfunctions and structural micropathologies may require nanoscale tools in order to detect them.

With medical conditions that require ongoing supervision and adjustment, such as maintaining optimum hormone balance and minimal accumulation of molecular debris (e.g. anti-aging medicine), nanoscale monitoring stations may act as onboard cellular guidance systems, stimulating or suppressing endocrine secretion as necessary to preserve an ideal state of equilibrium. In some cases, direct manufacture of compounds not easily produced by ribosomes or other biological organelles may be required.

1.2.4 Evolution of Bedside Practice

The relationship between physician and patient has been evolving in response to the rapidly changing medical environment. Two of the most important trends are the decline of traditional holism (with a concomitant increase in specialization) and the rise of therapeutic customization in medical practice.

1.2.4.1 Specialization and Holistic Medicine

Holism is a philosophy which holds that individuals function as complete units that cannot be reduced merely to the sum of their parts.2223 It is unarguable that a simplistic reductionist view of the patient which ignores the complex interactions among the many cells, tissues, organs, and systems constituting the human body is deeply flawed. For example, an understanding of the molecular basis of the contractile proteins of heart muscle will not alone tell us how heart muscle cells will look or act; knowing only about the parts of something is not sufficient to predict the behavior of the whole.2229 However, traditional concepts of holistic medicine go well beyond such basic systems philosophy — incorporating requirements for a consideration of all physical, emotional, social, spiritual, environmental and economic needs of the patient.

N. Jewson2203 and others2205 have decried the modern shift away from holistic medicine, which is asserted to have taken place in three historical phases in the West. The initial phase is identified as the practice of “bedside medicine,” where wealthy fee-paying clients in the 17th and 18th centuries helped shape their own diagnosis and treatment by medical practitioners in an holistic manner. Aspects of the patient’s emotional and spiritual life were seen as central by the practitioner in making a diagnosis, since most physical treatments were only palliative and so the doctor had to focus on nonphysical supportive measures. This frame of reference was progressively replaced during the 19th century with the trend toward “hospital medicine,” wherein physicians concentrated on generic classifications of diseases that were manifested in the patient, moving doctors away from the earlier focus on the individual as a whole person. The 20th century saw the development of “laboratory medicine,” which moved diagnosis and therapy even further away from the whole patient, “who came to be medically conceived as little more than a depersonalized object, comprised of a complex of cells.”

A more charitable view is that physicians have increasingly specialized in treating those physical diseases for which effective treatments may be readily specified, leaving nonphysical and nontreatable issues for psychiatrists, social workers, fitness coaches, priests, lawyers, or other professionals to deal with. In the 21st century this operational specialization may become complete, since nanomedicine phenomenologically regards the human body as an intricately structured machine with trillions of complex interacting parts, with each part (and each subsystem of parts) subject to individual scrutiny, repair, and possibly replacement by artificial technological means. In this new medical cosmology, the concept of the whole patient almost completely dissolves into a data-intensive whirlwind of molecular detail at the cellular, tissue, organ, and systemic levels.

And yet, as in a bygone era, patients once again will help to shape their own diagnosis and treatment at the hands of medical practitioners who begin to apply the volitional normative model of disease (Section 1.2.2) in their practices. This may breathe new life into the age-old medical school dictum to “treat the patient as a person” and to “focus on the sick person” rather than exclusively on the body.2230

The availability of extremely powerful and transforming molecular technologies also argues for a return to the romantic and perhaps quaint concept of a single doctor taking care of a single patient. The potential for interactions among highly potent nanorobotic instrumentalities argues for diagnostic and therapeutic “gatekeeping” by a single trusted practitioner in whom strategic treatment responsibility is vested — in partnership, of course, with the patient.

1.2.4.2 Customized Diagnosis and Therapeutics

In science the objective is to understand the individual occurrence by means of a general law; in medical practice, knowledge of what is generally the case does not tell the physician how to treat a particular patient. Thus the problem of how to proceed from the prototypic case to the individual instance remains to be solved in a systematic manner by the practitioner.2230

The practitioner of “bedside medicine” in the 17th and 18th centuries had few curative and almost no customized tools at his disposal — perhaps a few dozen basic surgical techniques, performed septically, hazardous, and without anesthesia, and a drug/herbal formulary consisting of a few hundred substances. For instance, the Edinburgh Pharmacopeia of 1803 listed only 222 simples while the London Pharmacopeia of 1809 listed fewer than 200 items, most of which had variable, uncertain, minimal, or nonspecific potency. A few more options became available to the 19th century medical practitioner, including complex surgical techniques for specific conditions that could be performed aseptically with anesthesia and a good chance for success, a somewhat broader and more efficacious pharmacopeia, vaccines targeted to several diseases, and an improving diagnostic ability. By the 20th century, the physician could prescribe
from among tens of thousands of specific drugs to target specific bacterial, viral, fungal, or parasitic infections; select from among a very precise array of anesthetic agents, chosen to avoid allergic responses in particular patients; perform a wide variety of noninvasive tests and scans for diagnostic purposes to identify very specific conditions; and perform minimally invasive surgeries directed at many arbitrary tissue masses as small as 1 mm³ in volume. The first glimmerings of personalized genetic therapies also began to appear.

With the arrival of nanomedicine in the 21st century, the treatment paradigm will complete its transition from coarse-grained, one-size-fits-all, slow-acting methods to molecularly-precise, completely customized, speedy and highly-efficacious procedures and instrumentalities. The irregular gunshot pattern of 18th century palliative will evolve during the early 21st century into a penetrating and perfectly tailored hail of “magic bullets” each targeting an individual cell or group of cells unique to the individual patient. The 19th century herbalist John Ayrton Paris could have been describing the nanomedical future in his popular textbook Pharmacologia, 1840 edition, when he wrote:

“If [a physician] prescribes upon truly scientific principles, he will rarely in the course of his practice compose two formulae that shall, in every respect, be perfectly similar, for the plain reason that he will never meet with two cases exactly alike. Now let me ask what constitutes the essential difference between the true physician and his counterfeit — between the philosopher and the empiric? Simply this — that the latter exhibits the same medicine in every disease, however widely each may differ from the other in its symptoms and character; while the former examines, in the spirit of philosophic analysis, all the existing peculiarities of his patient, and of his discord...and then adapts with a sound discretion and with a correct judgement of his medicinal agents, such means as may best be calculated to control and correct the patient’s morbid condition.”

1.2.4.3 The Physician-Patient Relationship

Many other aspects of the physician-patient relationship, especially as this relationship may evolve in the coming era of nanomedicine, are important and worthy of extensive discussion. One such issue is the obligation of both parties in the partnership to tell the truth. The patient as a fellow human being has every right to know the truth about his or her biological condition, but other considerations may enter into the fulfillment of this obligation. Patients have a quite natural anxiety about their own possible death, and it has been claimed that this anxiety implies that no one can be truly objective toward his or her own body. Guttentag observes that "tellling an unwellcome truth to the unprepared is as ill-conceived as trying to hide the truth from the prepared."

In the nanomedical era, the sheer number of "truths" that may become available for disclosure will increase enormously even as the terminal prognosis becomes rare. For example, each human being is believed to possess at least 4-10 potentially serious genetic defects; up to 1% of human DNA is of exogenous viral origin, and as much as 10% of the genome consists of transposons, discrete sequences that are positionally mobile among the chromosomes (Chapter 20). Should something be done about this, or not? What should the average patient make of the news that his physician has discovered exactly 57 submicron-scale lamellar defects scattered throughout the compact bone of the caudal epiphysis of the patient’s right humerus? In the nanomedical era, people will gain the ability to personalize their own physical structure to minute detail, but many patients will not be ready, willing, or able to assume responsibility for this knowledge. Thus there is no ideal substitute for the doctor’s interpretative abilities and judgements on the patient’s behalf as to the personal significance of specific diagnostic information. As the great clinician Thomas Addis observed in another context: “Honesty with patients requires thought and discipline and effort. Perhaps the single most important aspect of the physician-patient relationship, in any century, is the humanistic quality of the good doctor. The patient seeks a physician who cares about him as a person and will diagnose and prescribe in a sensitive and compassionate manner, accepting some degree of obligation to the patient. Speaking to medical students, J.C. Bennett describes the implicit social contract between doctor and patient that will still apply in the nanomedical era, as it does today:

“To receive medical care, patients must trust their bodies and their very lives to physicians, and so to be in an honest position to give medical care, physicians must earn such radical trust. Mere technical treatment of disease does not suffice. Patients must be able reasonably to believe that their physicians care about them in an extraordinarily personal way. This exchange of care for trust, while not identical to friendship or love, is equally binding. From it develops an interdependence that is far from unwholesome: rather, it potentiates care and promotes healing. Our late twentieth century sophistication and technologic orientation have too often cost us warmth, humor, and humanity, leaving us in social isolation. We do far better as professionals to err on the side of being human with our patients, than to try to play dux ex machina, the god from the machine."

1.2.5 Changing View of the Human Body

How does a patient regard his or her own body, and how might this most intimate of all relationships change in the nanomedical era? The so-called dualist theory of the human compound, as originally developed by Descartes and widely accepted today by the ordinary person, holds that the human being consists of two separate kinds of thing: the body and the mind or soul. The body acts as a host or receptacle for the mind. The mind, often called “the ghost in the machine,” is manifested by the brain, which it uses (via the bodily senses) to acquire and store information about the world, and to integrate this with its genetically-driven imperative to live, thus resolving internal conflicts among action-choices and expressing these in what we (in our consciousness) experience as decisive action.

Scientific medicine has concentrated primarily on the body. The ancient Roman physician Galen first dissected and vivisected a variety of animals to increase his knowledge of anatomy and physiology, and dissection became increasingly important in the training of physicians and surgeons, and in painting and sculpture, during the Renaissance. By the late 20th century, dissection had reached the molecular level, with the insides of the human cell and nucleus being taken apart and examined by molecular biologists, literally receptor by receptor. Dissection and the mechanistic understanding it provides have led some to decree what they regard as the modern “soulless” view of the human body as a mere machine.

In the nanomedical era, even the most diehard reductionist must come to see the human body not merely as a heap of parts but rather as a finely tuned vehicle that is owned and piloted by a single human mind. As with automobiles, some body-owners will be more diligent about maintenance, regular tuneups, and paint jobs than other body-owners. Some will crave the latest upgrades, while others may prefer a more conservative model that reliably gets them around town. At either extreme, all human vices and virtues will be on full display, though one may perhaps anticipate an increasing pride of corporeal ownership if for no other reason than because maintenance and repair will become quick, convenient, and inexpensive.
From this simple analogy of body-and-mind to car-and-driver, it might at first appear that the advent of nanomedical technology will confirm and strengthen the traditional dualist conception of the body. But closer inspection reveals that the analogy is at best incomplete, and at worst deeply flawed. This is because mind, first being necessarily embedded in physical structure and relying upon that structure for its faithful execution, and second, this physical structure now being manipulable at the molecular level, enters also into the purview of our mechanic. Both car and driver may be modified in the shop. Speaking allegorically, it is as if the driver, after getting his car a tuneup, emerges from the shop no longer favoring chocolate but enjoying vanilla instead, or now preferring jazz over classical, the opposite of before. Such psychological changes may be either volitional or emergent.

Until the late 20th century, human progress was measured almost exclusively in terms of externalities. Food was gathered, then sown, then manufactured. Shelters had no running water, then gainedouthouses, then indoor plumbing. Natural lighting and campfires gave way to candles, then oil lamps, then electric illumination. Finger-counting yielded first to the abacus, then the mechanical adding machine, and finally to the digital computer. But throughout all of history, the human body itself has remained largely untouched by progress. We have always regarded our bodies, evolved by natural selection, as fundamentally inviolate and degradationless, but rarely to any significant intrinsic improvement on the timescale of human civilization.

Now we are set to embark upon an era in which our natural physiological equipment may for the first time in history become capable of being altered, improved, augmented, or rendered more comfortable or convenient, due to advances in medical technology. The physical human body may be one of the last bastions of “naturalness” (Section 1.3.4). It will also be one of the last elements in our original natural senses such as hearing, sight, and smell. In the nanomedical era, machine-mediated sensory modalities may permit direct perception of physical phenomena well removed from our bodies in both time and space, or which are qualitatively or quantitatively inaccessible to our original natural senses. Perception will gradually expand to incorporate nonphysical phenomena including abstract models of mental software, purely artificial constructs of simulated or enhanced realities, and even the mental states of others. Such new perceptions will inevitably alter the way our minds process information.

But the winds of change will sweep deeper still, into our very souls. Like ants oblivious to the collective purpose of their colony, the billions of neurons in the human brain are all busily buzzing, wholly ignorant of the emergent plan. This is the physical, mechanical world of our electrochemical hardware. People also have thoughts, feelings, emotions, and volitions, a higher level in the data processing hierarchy which in turn is equally oblivious of the brain cells. We can happily think while being totally unaware of any help from our neurons. But nanomedicine will give us unprecedented systemic multilevel access to our internal physical and mental states, including real-time operating parameters of our own organs, tissues, and cells, and, if desired, the activities of small groups of (or even individual) neurons. Diverse parts of our selves previously closed to our attention may slowly conjoin and enter our conscious awareness.

Will this access promote an integrated identity or lead to hopeless confusion, or worse? Marvin Minsky, in his collection of essays _The Society of Mind_, persuasively argues that our selves or identities are in fact networks of semi-autonomous neurological “agencies” which sometimes cooperate and sometimes compete with one another. We think of ourselves as singular “persons,” but we also experience “conflicting desires” and “differing viewpoints” within our minds that are, in Minsky’s view, a direct experience of the multiplicity of our brain’s neurostructures. Other models of the human mind suggest that our internal mental states, prospectively transparent via nanomedical augmentation, are diverse and intricate; Julian Jaynes is one of many writers who have drawn attention to profound dichotomies between the two cerebral hemispheres. The component-oriented personality models of Freud (e.g. ego/id/superego), Jung (e.g. archetypes), and Rank (e.g. will/counterwill) and the identification of 451 distinct personality traits by Allport and Odbert warn us that full access to our brain’s architecture could be perilous.

More seriously, most of us suppose that we are endowed with free will. But if choices by free will are simply the resolution of conflicts of neurological subsystems, and we become consciously aware of those subsystems and are able to intervene in their processes, do we run the risk of runaway instabilities at the deepest levels of what we presently call our “minds”? Will we find that these instabilities are profound counterparts to the maladies we currently designate as epilepsy, or psychosomatic illnesses? In any redesigns of our brains which would involve opening doors to, quite literally, the ultrastructure of our thoughts, we could become “naked to ourselves” in ways that we can only vaguely speculate about at present. Along with any other dangers we might encounter, this will raise entirely new issues of the proper role of psychotherapy and the sanctity of personal privacy.

Repairs to the brain may be carefully monitored to ensure quality control and to verify intended results, as already proposed in another context. Major modifications might be strictly regulated, both to prevent abuse by unscrupulous third parties and also to forestall accidental or volitional alterations that could render the patient a significant threat to society. Nanomedical alterations to the brain and other physical systems may give us vastly expanded freedom to be who we choose to be (Section 1.3.4), along with increased responsibility to make wise and informed choices. The ethical and legal aspects of these questions, as well as the scientific and psychological ones, are extremely important and should be thoroughly debated in the years and decades that lie ahead.

1.3 The Nanomedical Perspective

1.3.1 Nanomedicine and Molecular Nanotechnology

A mature nanomedicine will require the ability to build structures and devices to atomic precision, hence molecular nanotechnology and molecular manufacturing are key enabling technologies for nanomedicine. The prefix “nano-” (from the Greek root nanos, or dwarf) means one-billionth (10^{-9}) of something. The term “nanotechnology” refers most generally to technology on the scale of a billionth of a meter, or a nanometer (a nanometer is ~6 carbon atoms wide). Similarly, the words “nanomachine,” “nanorobot,” “nanomotor” and “nanocomputer” may refer to complex engineered objects fabricated by positioning matter with molecular control.

Molecular engineering was discussed as an extension of bulk technologies in the 1960s and 1970s by von Hippel, von Foerster, and Zingheim. The phrase “Nano-Technology” was first used in print in 1974 by N. Taniguchi to refer to the increasingly precise machining and finishing of materials, progressing from larger to smaller scales and ultimately to nanoscale tolerances,
following in the path of Feynman’s proposed “top-down” approach, a schema which persisted in Taniguchi’s thinking throughout the 1980s and 1990s. In 1981, K.E. Drexler described a new “bottom-up” approach involving molecular manipulation and molecular engineering in the context of building molecular machines and molecular devices with atomic precision, a fundamentally different method. Drexler again described molecular technology in 1982 and molecular mechanical devices in 1983, first using the word “nanotechnology” in 1985 and 1986 as synonymous with molecular technology, finally settling upon “molecular nanotechnology” in 1991 and “molecular machine systems” in 1992 to clarify that his concept involved working devices constructed with atomic precision, as distinguished from nanostructured bulk materials, micromachinery, polymeric self-assembly, pure biotechnology, nanolithography, Langmuir-Blodgett thin films, and the like. Drexler’s definition — molecular nanotechnology as the three-dimensional positional control of atomic and molecular structure to create materials and devices with molecular precision — is the usage adopted in this book. The first known use of the term “nanomedicine” was in 1991 by Drexler, Peterson, and Pergamit in their popular book Unbounding the Future.

Is molecular nanotechnology possible? This question is explicitly addressed in Chapter 2, but the bottom line is that molecular nanotechnology violates no physical laws and there exist many possible technical paths leading to useful results. Even by 1985, for example, G. Yamamoto had reported on molecular gears, describing “compounds that exist in conformations which are regarded as static meshed gears with two-toothed and three-toothed wheels and some of them behave as dynamic gears.” H. Iwamura prepared a system that formed a chain of beveled molecular gears with GHz rotation rates. In 1998, it was generally accepted that molecular nanotechnology would be developed, although there was still some disagreement about how long it would take.

It is often noted that molecular biological systems are themselves nanomachines, constituting an existence proof for molecular nanotechnology. Indeed, biotechnology is one possible implementation pathway for molecular nanotechnology that is being pursued (Section 2.3.1). Table 1.3 reveals the close functional correspondence between the macroscale components of everyday machines and the molecular components of natural biological systems. Such comparisons have a long history. For example, Marcello Malpighi (1628-1694), a professor of medicine in Pisa who discovered the fine structure of the lungs and the capillaries using the microscope, once observed that our bodies are composed of strings, thread, beams, levers, cloth, flowing fluids, cisterns, ducts, filters, sieves, and other similar mechanisms.

Parallels to living systems as molecular machines were drawn by Changeau, McClaire, Laing, Drexler, and Mitchell, inspiring early thinking in molecular manufacturing. For example, in 1991 Drexler observed:

“Technology-as-we-know-it is a product of industry, of manufacturing and chemical engineering. Industry-as-we-know-it takes things from nature — ore from mountains, trees from forests — and coerces them into forms that someone considers useful. Trees become lumber, then houses. Mountains become rubble, then molten iron, then steel, then cars. Sand becomes a purified gas, then silicon, then chips. And so it goes. Each process is crude, based on cutting, stirring, baking, spraying, etching, grinding, and the like.”

“Trees, though, are not crude. To make wood and leaves, they neither cut, stir, bake, spray, etch, nor grind. Instead, they gather solar energy using molecular electronic devices, the photosynthetic reaction centers of chloroplasts. They use that energy to drive molecular machines — active devices with moving parts of precise, molecular structure — which process carbon dioxide and water into oxygen and molecular building blocks. They use other molecular machines to join these molecular building blocks to form roots, trunks, branches, twigs, solar collectors, and more molecular machinery. Every tree makes leaves, and each leaf is more sophisticated than a spacecraft, more finely patterned than the latest chip from Silicon Valley. They do all this without noise, heat, toxic fumes, or human labor, and they consume pollutants as they go. Viewed this way, trees are high technology. Chips and rockets are not.”

Contemplating applications of nanotechnology to medicine, Brian Wowk concluded that 20th century physicians were in a predicament similar to that which would be faced by 18th-century engineers trying to maintain a 20th-century automobile — repairs would be crude at best, and breakdowns inevitable:

“Like primitive engineers faced with advanced technology, medicine must catch up with the technology level of the human body before it can become really effective. What is the technology level? Since the human body is basically an extremely complex system of interacting molecules (i.e., a molecular machine), the technology required to truly understand and repair the body is molecular machine technology — nanotechnology. A natural consequence of [our achieving] this level of technology will be the ability to analyze and repair the human body...”

| Table 1.3. Comparison of Macroscale and Biomolecular Components and Functions |
|-----------------------------------------------|-----------------|-----------------|
| **Macroscale Device** | **Device Function** | **Biomolecular Examples** |
| Struts, beams, casings | Transmit force, hold positions | Microtubules, cellulose |
| Cables | Transmit tension | Collagen |
| Fasteners, glue | Connect parts | Intermolecular forces |
| Solenoids, actuators | Move things | Conformation-changing proteins, actin/myosin |
| Motors | Turn shafts | Flagellar motor |
| Drive shafts | Transmit torque | Bacterial flagella |
| Bearings | Support moving parts | Sigma bonds |
| Containers | Hold fluids | Vesicles |
| Pumps | Move fluids | Flagella, membrane proteins |
| Conveyor belts | Move components | RNA moved by fixed ribosome (partial analog) |
| Clamps | Hold workpieces | Enzymatic binding sites |
| Tools | Modify workpieces | Metallic complexes, functional groups |
| Production lines | Construct devices | Enzyme systems, ribosomes |
| Numerical control systems | Store and read programs | Genetic system |

(continued on the next page)
as completely and effectively as we can repair any conventional machine today.”

In Engines of Creation, Drexler drew inspiration from the cell’s eye view to recognize that nanotechnology could bring a fundamental breakthrough in medicine. Noting that 20th century physicians relied chiefly on surgery and drugs to treat illness, Drexler explained:

“Surgeons have advanced from stitching wounds and amputating limbs to repairing hearts and reattaching limbs. Using microscopes and fine tools, they join delicate blood vessels and nerves. Yet even the best microsurgeon cannot cut and stitch finer tissue structures. Modern scalpels and sutures are simply too coarse for repairing capillaries, cells, and molecules. Consider ‘delicate’ surgery from a cell’s perspective. A huge blade sweeps down, chopping blindly past and through the molecular machinery of a crowd of cells, slaughtering thousands. Later, a great obelisk plumbs through the divided crowd, dragging a cable as wide as a freight train behind it to rope the crowd together again. From a cell’s perspective, even the most delicate surgery, performed with exquisite knives and great skill, is still a butcher job. Only the ability of cells to abandon their dead, regroup, and multiply makes healing possible.”

“Drug therapy, unlike surgery, deals with the finest structures in cells. Drug molecules are simple molecular devices. Many affect specific molecules in cells. Morphine molecules, for example, bind to certain receptor molecules in brain cells, affecting the neural impulses that signal pain. Insulin, beta blockers, and other drugs fit other receptors. But drug molecules work without direction. Once dumped into the body, they tumble and bump around in solution haphazardly until they bump a target molecule, fit, and stick, affecting its function. Drug molecules affect tissues at the molecular level, but they are too simple to sense, plan, and act. Molecular machines directed by nanocomputers will offer physicians another choice. They will combine sensors, programs, and molecular tools to form systems able to examine and repair the ultimate components of individual cells. They will bring surgical control to the molecular domain.”

By the end of the 20th century, mainstream military (DoD), NIH, NSF, and other U.S. government and international groups had begun to seriously consider the potential future applications of molecular nanotechnology in medicine. For example, in 1997 a panel of U.S. Department of Defense health science experts known as Military Health Service Systems (MHSS) 2020 concluded in its final report:

“If a breakthrough to a [molecular] assembler occurs within ten to fifteen years, an entirely new field of nanomedicine will emerge by 2020. Initial applications will be focused outside the body in areas such as diagnostics and pharmaceutical manufacturing. The most powerful uses would eventually be within the body. Possible applications include: programmable immune machines that travel through the bloodstream, supplementing the natural immune system; cell healing machines to stimulate rapid healing and tissue reconstruction; and cell repair machines to perform genetic surgery.”

The present book takes as its starting point the assumption that the mass production of nanomachines at modest cost is technically feasible (Chapter 2), and then explores the medical implications of this assumption. Proposed systems presented in this trilogy are intended not as final engineering blueprints but merely as points of departure for further analysis and refinement. All designs and projected capabilities are, for the most part, conservatively drawn with generous safety margins, leaving a considerable volume of design space yet to be explored by more intrepid future investigators.

1.3.2 Nanomedicine: History of the Idea

Conclusive proof of the existence of atoms was not obtained until the close of the 19th century. This may explain why the idea of nanomedicine is an exclusively 20th century phenomenon. The first hint of it may be found in a famous 1929 essay written by J.D. Bernal:

“The discoveries of the twentieth century, particularly the micro-mechanics of the Quantum Theory which touch on the nature of matter itself, are far more fundamental and must in time produce far more important results. The first step will be the development of new materials and new processes in which physics, chemistry and mechanics will be inextricably fused. The stage should soon be reached when materials can be produced which are not merely modifications of what nature has given us in the way of stones, metals, woods and fibers, but are made to specifications of a molecular architecture. Already we know all the varieties of atoms; we are beginning to know the forces that bind them together; soon we shall be doing this in a way to suit our own purposes. The result — not so very distant — will probably be the passing of the age of metals and all that it implies — mines, furnaces, and engines of massive construction. Instead we should have a world of fabric materials, light and elastic, strong only for the purposes for which they are being used, a world which will imitate the balanced perfection of a living body.”

Development of the concept of nanomedicine has followed two principal paths, which Richard Smalley has termed “wet nanotechnology” in the biological tradition, and “dry nanotechnology” in the mechanical tradition. Both approaches were presaged in speculative fiction. The following abbreviated history focuses on nanomedicine largely to the exclusion of broader issues in molecular engineering, manufacturing and nanoscopy, and includes a number of inspirational, speculative, or fictional early references from non-refereed sources.

We will first consider the biotechnology approach (Section 1.3.2.1), followed by the molecular nanotechnology approach (Sections 1.3.2.2-3), and then these two approaches will be contrasted and compared in Section 1.3.3.

1.3.2.1 The Biological Tradition

The general idea of biological engineering stretches back at least to the mid-19th century, but the first science fiction story involving actual genetic engineering was “Proteus Island” (1936), written by the chemical engineer Stanley Weinbaum. Artificial engineered organisms subsequently appeared in minor roles in several stories, a notable example being the familiaries employed by the fake witches in Fritz Leiber’s “Gather, Darkness!” (1943), and A.E. van Vogt used “gene transformation” to create the superman in “Slan” (1940). The first artificially evolved creatures appeared in Theodore Sturgeon’s “Microcosmic God” (1941), wherein a biochemist established conditions allowing accelerated artificial evolution, creating the Neoters, a submillimeter-sized race of intelligent hypermetabolic creatures which could accomplish tasks very rapidly. The first artificially designed microcreatures appeared in James Blish’s “Surface Tension” (1952). In this story, crash-landed dying human astronauts create a completely new form of humanity — tiny men and women reduced to protozoan size — who are seeded in the pools and puddles at the surface of the new planet, and who go on to master the biotechnology necessary to travel from one water puddle to another.

The scientific tradition of biological nanomachines for medical purposes began in 1964 when Robert Ettinger, an early cryonics pioneer, suggested that cellular-level or even molecular-level repair might be developed for life extension. Ettinger speculated that “…surgeon machines, working 24 hours a day for decades or even
centuries, will tenderly restore the frozen brains, cell by cell, or even molecule by molecule in critical areas.

In 1965, the synthesis of artificial life was publicly proposed as a national goal by the president of the American Chemical Society, Professor Charles Price, who pointed out that many new types of life might be made, not “mere imitations” of biology as we know it.153

In 1967, Isaac Asimov2257 suggested the future possibility of “factories...where the working machinery consists of submicroscopic nucleic acids” and that a “repertoire of hundreds or thousands of complex enzymes” could be used to “bring about chemical reactions more conveniently than any methods now used” and also for “helping to construct life.”

In 1968, G.R. Taylor2258 cited the possibilities for genetic engineering and genetic surgery: “The microsurgery of DNA may possibly be achieved by physical methods: fine beams of radiation (probably laser light or pulsed X-rays) may be used to slice through the DNA molecule at desired points.” He also cited predictions that bacteria would soon be programmed.

In 1969, J. White2259 suggested that a modified virus could be used as a cell repair machine: “It has been proposed that appropriate genetic information be introduced by means of artificially constructed virus particles into a congenitally defective cell for remedy; similar means may be used for the more general case of repair. The repair program must use means such as protein synthesis and metabolic pathways to diagnose and repair any damage...[Information] can be preserved by specifying that the repair program incorporate appropriate RNA tapes into itself...”

In 1970, Jeon et al158 carried out “the reassembly of Amoeba proteus from its major components: namely nucleus, cytoplasm, and cell membrane,” taken from three different cells.

In 1972, Ettinger160 proposed using genetic engineering to make microscopic biorobots: “Genetic engineering’s most sensational impact will concern the modification of humans, but it will have other uses as well. Some of the “robots” that will serve us will need to be nanominiaturized.” Existing organisms could be modified to make biologically-based programmable biorobots for medical applications: “If we can design sufficiently complex behavior patterns into microscopically small organisms, there are obvious and endless possibilities, some of the most important in the medical area. Perhaps we can carry guardian and scavenger organisms in the blood, superior to the leukocytes and other agents of our human heritage, that will efficiently hunt down and clean out a wide variety of hostile or damaging invaders.” Computerized cellular repair machines “must use means such as protein synthesis and metabolic pathways to diagnose and repair any damage...[Information] can be preserved by specifying that the repair program incorporate appropriate RNA tapes into itself...” Also in 1972, Daniell [2260] described various possibilities for generating new life forms via “life-synthesis” and genetic engineering, noting that “macromolecular engineering” might enable the development of very powerful and compact macromolecular computer systems.

In 1974, Halacy2261 noted “some rather inglorious ways” to use “the miracle of artificial life,” including potential capabilities for growing diverse items ranging from computers to airplanes. Morowitz153 suggested cooling cells to cryogenic temperatures in order to analyze and determine their structure. Artificial cells could also be assembled at such temperatures and then be set in motion by thawing. Morowitz further reported that microsurgery experiments on amoebas “have been most dramatic. Cell fractions from four different animals can be injected into the eviscerated ghost of a fifth amoeba, and a living functioning organism results.”

In 1975, Richard Laing2260 described the theoretical possibility of molecular machine self-replicators using molecular (data) tapes based on the idea of universal Turing machines, examining several ways that such “artificial organisms” might replicate themselves as “a vehicle for the exploration of broad biological possibilities.”

In 1976, Donaldson2262 presented the first detailed (and quite ambitious) list of biotechnological techniques that appeared necessary to achieve cell repair and might prove feasible, writing at a time that predated many current capabilities such as automated protein/DNA sequencing and synthesis, and most knowledge of restrictions on cellular developmental pathways and genetic programs and networks. For repair at the level of the cell, Donaldson’s techniques would have included:

1. The ability to design enzymes to produce specific repair functions such as renaturing denatured proteins, joining broken lipoprotein complexes, annealing broken strands of DNA or RNA, reading proteins of existing or special types onto RNA and replicating them, and giving a cell the ability to metabolize new substrates, use novel cofactors, or construct essential amino acids;

2. Specially constructed bacteria or macrophages able to replicate themselves, spread throughout a specific target tissue, and carry out specific repairs according to the programs designed into their DNA/RNA; these could be designed to operate at unnatural temperatures or to utilize metabolic pathways not presently found in nature;

3. The abilities to re-introduce lost DNA or lost organelles such as mitochondria into a cell, to introduce entirely new forms of organelles perhaps to perform specific repair functions, and to introduce new metabolic capacities into a cell;

4. The ability to modify at will the developmental program of a cell, as for instance to induce postmitotic cells such as neurons to divide,* according to a specific program, forming daughter cells with specific properties; and

5. Several different types of repair bacteria able to work together in an integrated fashion, and linked together by chemical means (e.g., hormones), so as to apply optimal repairs to every body cell in order to (A) diagnose the precise nature of the damage, call other repair bacteria to its location, and report to the attending doctor that new types of repair bacteria other than those already introduced are needed, and (B) identify structures which must be preserved (e.g., memory) and reconstruct them if necessary.

For repairs at the level of the whole organism, Donaldson offered the following rather aggressive biological nanotechnology techniques:

1. Understanding the physiology of aging combined with the ability to reverse it;

* It is now known that genetic programs to allow postmitotic cell division do not exist for many cells, although in theory artificial programs could possibly be devised. Stem cells may differentiate, and in some cases cells may be induced to de-differentiate back to stem cells thus allowing subsequent division and re-differentiation — but losing, in the case of neurons, the existing neural connections.
2. Control over growth and development, including the ability to program types of growth and development which do not naturally occur, such as growth of new eyes or other organs which have been lost or damaged, growth of an entire and well-formed body from a head alone, and regrowth of injured or lost brain tissue; and

3. Nonpermanent “substitute organs” which will take over from others which have been lost, including (A) the ability to keep a given tissue alive and healthy in vitro for an indefinite time and similar abilities for a body part, and (B) temporary replacements for any body organ which may have been lost. These would be used to support the body while new organs were growing, e.g. as “metabolic crutches.” This capacity specifically includes the ability to make temporary replacements for diffuse systems such as the vascular or nervous systems. For instance, a specially created “plant” would grow an entire vascular system into the patient, down to replacement for the capillaries and venules from a single seed, always introducing its fibrils between cells and destroying none or very little of the original structure.

The last of these is a description of a (clearly speculative) “repair net,” a concept which Donaldson may have been the first to propose in 1976. A whole-body repair net was later termed a “chrysalis,” and twelve years later, in 1988, Donaldson provided an artist’s conception and an additional description of these proposed biotechnological instrumentalities:161,1724

“Severe crushing or mangle injuries require us to provide a new vascular system. The repair device might resemble a fungus, growing mycelia into the injured tissue. A repair net would grow into a [crushed] limb, guided by recognition of the injured cells and a plan for how the limb should look after repair...[In the most severe injuries,] a chrysalis first envelops the patient, then enters in between all his cells. It disassembles the patient, surrounding each cell with its own repair machinery and vascular system. The geometry already preserves information about locations of the patient’s cells. If necessary, morphogen chemical gradients could also retain this information. A patient would [locally] swell up to 10 times original diameter. After repair, the chrysalis withdraws the same way it entered.”

In 1977, Darwin2263 further developed the theme of tissue repair and cellular repair biomachines, independently proposing a modified white blood cell to perform repair functions. In 1981, Asimov2264 suggested that we “consider the bacteria. These are tiny living things made up of single cells far smaller than the cells in plants and animals...[We can, by properly designing these tiniest slaves of ours, use them to reshape the world itself and make it close to our hearts’ desire.” More importantly, Donaldson161 elaborated on his earlier discussion of how cryonically suspended human beings might be repaired. He extended the earlier concepts of Ettinger60 and White,2259 concluding that “with such hybrid technology as micro-miniature biological-mechanical machines the size of viruses” and related technology, “it seems unlikely that (to repair a single cell) there would be any difficulty at all in principle to carrying out any imaginable repair.” He estimated that about 10 programmable cell-repair biomachines could be introduced into a cell that was under repair without causing too much mechanical disruption. In 1988, Donaldson1724 described artificial macrophages that “can carry control machinery to recognize target cells, responding only to them, or responding differently depending upon cell type or cell conditions. They can still work even if the target cell isn’t functioning (unlike viruses), rebuild target cell machinery other than the genes, and transfer many more genes up to an entire copy of the patient’s genome.” They can also “communicate with one another [and] release diffusible chemicals to guide one another’s behavior.”

By the 1990s, bioengineered viruses of various types and certain other vectors were routinely being used in experimental genetic therapies3001-3011 as a means to target and penetrate certain cell populations, with the objective of inserting therapeutic DNA sequences into the nucleus of human target cells in vivo. Retrovirally-altered lymphocytes (T cells) began to be injected into humans for therapeutic purposes. Another example of an engineered cell in therapeutics was the use of genetically modified cerebral endothelial cell vectors to attack glioblastoma, which was being pursued by Neurotech (in Paris) in 1998. Engineered bacteria were being pursued by Vion Pharmaceuticals in collaboration with Yale University:3038 In their “Tumor Amplified Protein Expression Therapy” program, antibiotic-sensitive Salmonella typhimurium (food poisoning) bacteria were attenuated by removing the genes that produce purines vital to bacterial growth. The tamed strain (cell line VNP20009) could not long survive in healthy tissue, but quickly multiplied -1000-fold inside tumors, which are rich in purines. The next step would be to add genes to the bacterium to produce anticancer proteins that can shrink tumors, or to modify the bacteria to deliver various enzymes, genes, or prodrugs for tumor cell growth regulation. The engineered bacteria were available in multiple serotypes (Section 8.5.2) to avoid potential immune response in the host. Phase I human clinical trials were expected to begin in 1999 using clinical dosages produced in 50-liter fermenters, and other possible bacterial vector species were being examined.

Micro-biorobotics was still regarded by many in the biotechnology community as a highly speculative topic in 1998. Glen A. Evans8104 described the possible construction of synthetic genomes and artificial organisms. His proposed strategy involves determining or designing the DNA sequence for the genome, synthesizing and assembling the genome, introducing the synthetic DNA into an enucleated pluripotent host cell, then introducing the host cell into an organism. Evans foresees the high throughput “read-out” of gene products discovered through genomic sequencing, rapid construction of designer genes and genomes, automated designer vector synthesis for gene replacement/insertional mutagenesis, and finally the construction of synthetic genomes and artificial organisms. Robert Bradbury8105 has considered requirements and costs for genome synthesis and replacement. For example, Bradbury estimates genome synthesis costs by assuming a projected ~$0.03/base for DNA synthesis (compared to ~$0.80-$1.00/base-micromole in 1999), with ~20,000 expressed genes per biorobotic cell with an average gene size of ~3000 bases, giving a ~60 megabase expressed genome per biorobotic cell (1 chromosome) and a raw synthesis cost of ~$3 million per designer cell line (excluding design costs). There has also been discussion of “chemical reaction automata” as precursors for synthetic organisms,3016 synthetic lifeforms2265 and “nanobiology”3017,3079,3081 “birobotics”3017,3079,3081 “cell rovers”3341, microfluidic and artificial engineering,1962,3248,3543,3544 lymphocyte engineering,365 and artificial chromosomes (Chapter 20), and cell engineering (Chapter 21) is rapidly gaining popularity.

1.3.2.2 The Mechanical Tradition

The late science fiction author Robert A. Heinlein155 nearly invented the concept of molecular nanotechnology in 1942 when he suggested a process for manipulating microscopic structures. Heinlein envisioned the extensive use of size-life teleoperator hands, called “waldoes,” complete with sensory feedback for full, remote-controlled telepresence. His fictional hero, Waldo, used a collection of these mechanical teleoperated hands for building and operating a series...
of ever-smaller sets of such mechanical hands. The smallest mechanical hands, “hardly an eighth of an inch across,” were equipped with micro-surgical instruments and stereo “scanners,” and were used to “manipulate living nerve tissue, [to examine] its performance in situ,” and to perform neurosurgery. Eric Frank Russell’s 1947 story “Hobbyist” described a fabrication process with “atom fed to atom like brick after brick to build a house.” In 1955, Russell’s serial “Call Him Dead,” also published in *Astounding Science Fiction*, had a virus-based alien intelligence that spread through contact with blood or saliva; the story features a “microforger,” a man who makes “surgical and manipulatory instruments so tiny they can be used to operate on a bacillus.” Also in the mechanical tradition, Isaac Asimov’s *Fantastic Voyage* in *Astounding Science Fiction* included producing vast quantities of ultrasmall computers and various micro- and nano-robots.

Feynman was clearly aware of the potential medical applications of the new technology he was proposing. After discussing his ideas with a colleague, Feynman offered the first known proposal for a nanomedical procedure to cure heart disease: “A friend of mine (Albert R. Hibbs) suggests a very interesting possibility for relatively small machines. He says that, although it is a very wild idea, it would be interesting in surgery if you could swallow the surgeon. You put in some of the mechanisms involved in growth and aging.”

In 1961, K.R. Shoulders rejected the use of biological building blocks, even though biological “processes do work, and they can do so in a garbage can without supervision.” He saw them as too limited environmentally and too difficult to control with available technology. Instead, he sought to directly produce much simpler, more powerful and rugged nanostructure arrays, operating at video frequency rates, which in turn could ultimately aid in their own replication. In 1965, Shoulders reported the actual operation of micromanipulators able to position tiny items with 10 nm accuracy while under direct observation by field ion microscopy.

In 1970, Volkenstein noted that “the creation of a nonmacromolecular system which would act as a model for living organisms is definitely possible” but could not arise by itself, and that the macromolecularity of present organisms is not essential, but due to their evolutionary origins. Taking some poetic license, he added: “Consequently, the cybernetic nonmacromolecular machine, which simulates life, could have been and can be created on earth only by man. Then it could perfect itself without limits.” T. Nemes discussed artificial self-replicating machines and described how to construct “an automatic lathe able to reproduce itself,” a concept apparently developed before von Neumann’s work on machine replication.

In 1981, Drexler suggested the construction of mechanically deterministic nanodevices using biological parts; these devices could inspect cells at the molecular level and also repair cellular tissues that had been damaged during cryonic suspension. In 1982, Drexler described cell repair machines even more clearly in the mechanical tradition, in a popular publication.

By 1983, Drexler began privately circulating a draft technical paper entitled “Cell Repair Machines” which investigated for the first time, in some detail, whether an advanced mechanical-based nanotechnology would “permit construction of systems of molecular-scale sensors, computers, and manipulators able to enter and repair cells; the nature of the computational algorithms and magnitude of the computational resources needed to guide repairs; and the physical capabilities and constraints important to the repair process [in order to] sketch the conceptual design of a cell repair system based on a mature molecular technology.”

Peterson notes that “medical applications were explored by Drexler during the early 1980s but the medical community was not ready for the concept.” A technical paper by Drexler, invited by an editor at the *Journal of the American Medical Association*, was dismissed by a referee as “science fiction.”

In 1985, G. Feinberg proposed using short-wavelength coherent laser energy to power and communicate with “nanosensors that could be implanted into the human body. They could...[monitor] various physiological functions from subcellular molecules up through tissues and organs...essential in determining some of the mechanisms involved in growth and aging.”

In a 1985 book entitled *Robotics*, edited by Marvin Minsky (a well-known computer scientist and artificial intelligence pioneer), Minsky briefly described how fully-automatic cellular repair machines might work, following Drexler’s vision:

“Suppose that we could design a repair machine so small that it could repair an artery from the inside! The first such machines might be the size of fleas. (There is room for a great deal of machinery in something the size of a flea — as the body of the flea itself testifies.) These micromachines could crawl into all but the smallest blood vessels, clean out debris, and reline the walls with suitable materials yet to be invented. Later generations of micromachines could be even smaller, perhaps no larger than the body cells they repair. These minuscule machines would be mass-produced by the billions, either by larger machines or by techniques of making them reproduce themselves. Perhaps these biological janitors would even be implanted in our bodies to remain there as permanent maintenance workers, just like many biological cells that already serve such purposes.”

“Today the idea of such a technology may seem fantastic, yet many of the circuits in our computers are already smaller than many of our body’s cells. Let’s try, for a moment, to look ahead to: mass production of highly intelligent machines; gigantic advances in miniaturization; a technology so advanced that these machines reproduce themselves without our help. Fantastic? Not at all. Even the simplest algae and bacteria can do that. True, they’re not intelligent, but each of them contains enough computerlike machinery and memory to do those things. So, to build bacteria-size
computers should be perfectly feasible, once we have the necessary microtechnology. Some day we’ll have the means to build artificial, cell-like machines with all those capabilities.”

R.A. Freitas Jr. also contributed a chapter to Minsky’s 1985 book, offering the following suggestion for remote-controlled incisionless nanosurgery: “One possibility is the concept of remote-controlled medical mites made feasible by modern micromachinery technology. Some medical mites would be like microminiature submarines, released inside the human body for internal sensing. Other mites could float, crawl, or swim through major arteries in the human body and perform on-site repairs from within, controlled by radio link under direction of a skilled telemicrosurgeon.”


In 1988, A.K. Dewdney reported an early nanomedical concept of an artery-cleaning nanorobot that he attributed to Drexler, accompanied by an artist’s conception with the caption “a nanomachine swimming through a capillary attacks a fat deposit.” (A preferable nanorobotic design for this purpose, the endotheliocyte, is presented in Chapter 22.)

1.3.2.3 Nanomedicine in the 1990s

Despite the tremendous importance of molecular nanotechnology, a published 1993 literature review of robotics in health care included not a single reference to nanotechnology or to nanomedicine. At this writing in 1998, the list of original post-1989 technical and popular nonfiction works that deal with nanomedical topics in the mechanical tradition is short enough to permit virtually an exhaustive listing, including Beardsley, Bova, Coombs and Robinson, Crawford, Drexler, DuCharme, Emanuelson, Fahy, Fiedler and Reynolds, Freitas, Kaehler, Klatz and Kahn, Kurzweil, Lampton, Merkle, Merrill, Minsky, More, Ostman, Reifman, and Wowk and Darwin. The first nanomedical device design technical paper was published in 1998 by Freitas in the biotechnology journal Artificial Cells, and the present trilogy (Nanomedicine) is the first book-length technical treatment of the medical implications of molecular nanotechnology.

The author again must caution enthusiasts that the full capabilities of these machines will not be realized without a great deal of sweat and toil by legions of well-funded and dedicated researchers working for many decades to develop the technology. One purpose of this book is to help motivate this future work.

1.3.3 Biotechnology and Molecular Nanotechnology

Some pragmatic readers may be wondering what is so special about molecular nanotechnology, that existing or anticipated biotechnology could not accomplish just as well? After all, biotechnology is already an established medical capability. It has real applications and real products already on the market. Reflecting upon the future possibility of sophisticated medical nanorobots equipped with powerful nanocomputers, in 1989, one well-known cryobiologist mused that “what is not clear is just what need we have for such devices.” The simplest answer, suggested by Figure 1.5, is that each of the three contemporary branches of “nanotechnology” offers something of unique value to the practice of medicine.

Nanoscale materials technology has already found widespread use in medicine, including biocompatible materials and analytical techniques, surgical and dental practice, nerve cell research using intracellular electrodes, bionanostructures research and biomolecular research using near-field optical microscopy, scanning-probe microscopy and optical tweezers, and vaccine design. In recent times, the field has expanded to include genetic engineering and now takes as its ultimate goal no less than the engineering of all biological systems, even completely artificial organic living systems, using biological instrumentalities.

The third branch, molecular nanotechnology, takes as its purview the engineering of all complex mechanical systems constructed from the molecular level — potentially offering new tools for medical practice, the principal subject of this book. Observes G.M. Fahy, “the difference between nanotechnologists and biotechnologists is that the former do not restrict themselves to the biological limitations of the latter, and they are much more ambitious about the kinds of accomplishments that they want to achieve.”

Doctors can utilize solutions to medical problems from all three approaches. As noted earlier, 80-90% of medical complaints resolve themselves via natural homeostatic processes, or without the necessary involvement of active biotechnological or molecular-nanotechnological agents. But by employing biotechnology, the range and efficacy of treatment options greatly increases. With molecular nanotechnology, the range, efficacy, comfort and speed of possible medical treatments again expands enormously. Molecular nanotechnology is essential when the damage to the human body is extremely subtle, highly selective, or time-critical (as in head traumas, burns, or fast-spreading diseases), or when the damage is very...
massive, overwhelming the body’s natural defenses and repair mechanisms.

Table 1.4, inspired by an earlier discussion from G.M. Fahy, makes these options more explicit. At every difficulty level, most classes of medical problems may be resolved with varying efficacy within the homeostatic/nanomaterials, biotechnological, or molecular-nanotechnological approaches. As the chosen technology becomes more precise, active, and controllable, the range of options broadens and the quality of the options improves. Thus the question is not whether molecular nanotechnology is required to accomplish a given medical objective. In many cases, it is not — though of course there are some things that only biotechnology and molecular nanotechnology can do, and some other things that only molecular nanotechnology can do. Rather, the important question is which approach offers a superior solution to a given medical problem, using any reasonable metric of treatment efficacy. For virtually every class of medical challenge, a mature molecular nanotechnology offers a wider and more effective range of treatment options than any other approach.

It is quite possible to imagine an advanced biotechnology that uses an engineered white cell, fibroblast, or macrophage chassis, energized by native oxygen and glucose (Section 6.3.4) and modified mitochondrial powerplants, driven by pseudopodia, cilia or flagella (Section 9.4), communicating and navigating via biochemical signals (Sections 7.2.1 and 8.4.3), and even incorporating onboard digital biocomputers (Section 10.2.3) to make microscopic biorobots. Principal arguments favoring the biotechnology approach for medical purposes are: (1) that we are already somewhat familiar with such systems, after half a century of intensive molecular biology research; (2) that we have already “built” precursor systems, such as a whole living amoeba constructed from five distinct parts, viruses as DNA insertion devices, and natural replication stimulated in genetically engineered starter microbes; (3) that biocompatibility will not be a major issue since fibroblasts (which express no HLA Class II antigens, hence stimulate no rejection response; Section 8.5.2.1) could be used as the starter material; (4) that both engineered viruses and bacteria are already in wide commercial and research use; and (5) the greater complexity of self-repair in mechanical systems, should it be needed. Many believe that the development pathway to early biorobots may be considerably shorter than for the mechanical nanorobots of the molecular nanotechnology approach, for which in 1998 not a single working prototype yet existed, even in research laboratories.

It is also possible to imagine a molecular nanotechnology that uses mechanical nanorobotic systems. Such systems will have many constitutional differences from biological-based systems. For instance, mechanical systems will transport parts, materials, energy and instructions via fixed channels, whereas most (but not all) biological systems operate by diffusion. Mechanical systems will have structures constrained by specific geometries, whereas biological systems have structures defined by patterns of containment and interconnection — the shape of a membrane compartment in a cell matters less than its continuity and the contents of the volume it defines. Mechanical systems will be deterministically manufactured by operations analogous to manual construction, whereas engineered ribosomes will self-assemble via diffusion and stochastic matching of complementary parts; in other words, biology uses recipes, while mechanical systems use blueprints. Cells grow, with their

<table>
<thead>
<tr>
<th>Level of Difficulty</th>
<th>Specific Medical Challenge</th>
<th>Homeostatic or Nanomaterials Technology</th>
<th>Biotechnology Approach</th>
<th>Molecular Nanotechnology Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Minor Symptoms and Minor Physical Trauma</td>
<td>Herbal preparation, minor surgery, passage of time</td>
<td>Symptomatic suppression pharmacology</td>
<td>Pharmaceuticals, zicropes, respirocytes, etc.</td>
</tr>
<tr>
<td>II</td>
<td>Infectious Agents</td>
<td>Natural immune system</td>
<td>Antibiotics and vaccines</td>
<td>Nanobiotics</td>
</tr>
<tr>
<td>III</td>
<td>Mutation and Cell Disease</td>
<td>Natural immune system Major surgery</td>
<td>Molecular diagnostics and treatments; gene therapies</td>
<td>Nanobiotics Cell repair machines</td>
</tr>
<tr>
<td>IV</td>
<td>Health Maintenance and Aging</td>
<td>Vitamins, good diet, exercise, moderation and “clean living”</td>
<td>Gene therapy to enhance self-repair and immune function</td>
<td>Whole-body cyto-assay Cell repair machines</td>
</tr>
<tr>
<td>V</td>
<td>Major Organ Replacement or Repair</td>
<td>Organ transplantation Major surgery</td>
<td>Stem cell &amp; tissue engineering; embryonic gene reactivation</td>
<td>Cell mills, tissue mills, organ mills, nanosurgery</td>
</tr>
<tr>
<td>VI</td>
<td>Morpho-Engineering</td>
<td>Cosmetic surgery</td>
<td>Control of natural morphogenetic systems</td>
<td>Nanosurgery Chromatin editors</td>
</tr>
<tr>
<td>VII</td>
<td>Emergency Care; Restoration of Nonhomeostatic Tissue</td>
<td>Emergency major surgery Prosthesis or bionics</td>
<td>Biological repair nets Bioengineered immune cells</td>
<td>Traumapods Warm biostasis</td>
</tr>
<tr>
<td>VIII</td>
<td>Augmentation of Natural Structure or Function</td>
<td>Prosthesis or bionics</td>
<td>Cross-species and artificial morphogenetic supplementation</td>
<td>Cell engineering devices Autogenous control</td>
</tr>
</tbody>
</table>
parts adapting to one another; mechanical nanorobots may be constructed from parts of fixed structure.\textsuperscript{2244} Biology uses self-repair; mechanical systems generally do not\textsuperscript{2222} — largely because self-repair (by component exchange) will not be needed in molecular mechanical systems, whose designs may be made more simple by relying upon high component redundancy\textsuperscript{10} (Chapter 13). These and other differences imply a number of important advantages that mechanical-based medical systems will enjoy over biological-based medical systems, which, taken together, strongly suggest that predominantly mechanical nanosystems may be the approach of choice for a mature medical nanotechnology. The advantages of molecular nanotechnology (e.g. the mechanical tradition) are many:

1. **Speed of Medical Treatment** — Doctors may be surprised by the incredible quickness of nanorobotic action when compared to the speeds available from fibroblasts or leukocytes. Normal homeostatic processes such as dermal wound repair via natural fibroblasts may require weeks to run to completion. Typical fibroblast movements occur at 0.1-1 microns/sec (Section 9.4), but mechanical nanomanipulators can operate at 1-10 cm/sec speeds (Section 9.3.1) or faster, a speed advantage of 4-5 orders of magnitude. Even the strongest biological fibers (e.g. intermediate filaments) have a failure strength 3 orders of magnitude below the strongest mechanical fibers (e.g. fullerene nanotubes; Table 9.3). Biological cilia beat at ~20 MHz while mechanical nanocilia may cycle up to ~20 MHz, though practical power restrictions and other considerations may limit them to the ~10 KHz range for most of the time (Section 9.3.1). Flexible mechanical surfaces can complete a morphing motion in ~0.1 milliseconds, compared to the ~100 milliseconds snapback time for pinched red cell membrane (Section 5.3.1.4), again a thousandfold advantage in speed. Thus we expect that mechanical therapeutic systems can reach their targets up to ~10,000 times faster, all else equal, and treatments which require ~10^3 sec for a biological system may need only ~10^2 sec for a mechanical system.\textsuperscript{3233} Tachyatria improves both patient and physician comfort. In either biological or mechanical systems, large numbers of devices of comparable physical size (e.g. -microns) may be employed to do the work, so numbers alone cannot offset the mechanical speed advantage.

2. **Power Density and Transduction** — Biological cells typically employ power densities of 10^3-10^4 W/m^3, with maximum densities of ~10^6 W/m^3 in honeybee flight muscle cells and bacterial flagellar motors (Table 6.8). By contrast, nanomechanical power systems can produce power densities of 10^2-10^12 W/m^3 (Section 6.3), an advantage of 10^3-10^8 for mechanical over biological systems. By 1998, conducting polymer-based actuators generated 20-100 times the force for a given cross-sectional area as mammalian skeletal muscle.\textsuperscript{238} Additionally, amoebic locomotion in motile cells requires diffusion-limited cytoskeletal disassembly and reassembly to achieve movement; mechanical motility systems may employ simple cable-pulling, winches, or ratchets, which are faster and more direct. Some biological energy transducers are reversible, but muscle contraction is irreversible — not only cannot muscle actively re-expand, but stretching it doesn’t make it produce much useful chemical energy. In contrast, electric motors can be run backwards to generate electricity, forcing pistons makes them pump, and loudspeakers can be used as microphones.\textsuperscript{2022}

3. **Superior Building Materials** — Typical biological materials have tensile failure strengths in the 10^5-10^6 N/m^2 range, with the strongest biological materials such as wet compact bone having a failure strength of ~10^8 N/m^2, all of which compare poorly to ~10^10 N/m^2 for good steel, ~10^11 N/m^2 for diamond and carbon fullerenes (Table 9.3), again showing a 10^7-10^9 advantage for mechanical systems that use nonbiological materials. Nonbiological materials can be much stiffer, permitting the application of higher forces with greater precision of movement, and they also tend to remain stable over a larger range of temperature, pressure, salinity and pH. Proteins are heat sensitive in part because much of the functionality of their structure is due to noncovalent bonds involved in folding, which are broken more easily at higher temperatures; in diamond, sapphire, and many other rigid materials, structural shape is covalently fixed, hence is far more temperature-stable. Most proteins tend to become dysfunctional at cryogenic temperatures, unlike diamond-based mechanical structures (Section 10.5). Biomaterials are not ruled out for all nanomechanical systems, but represent only a small subset of the materials that can be used in nanorobots. Mechanical systems can employ a wider variety of atoms and molecular structures in their design and construction, with novel functional forms that might be difficult to implement in a biological system such as steam engines (Section 6.3.1) or nuclear power (Section 6.3.7). As another example, an application requiring the most effective bulk thermal conduction possible should use diamond, the best conductor available, not some biomaterial with inferior thermal performance.

4. **Nondegradation of Treatment Agents** — Diagnostic and therapeutic agents constructed of biomaterials generally are biodegradable in vivo, although there is a major branch of pharmacology devoted to designing drugs that are moderately non-biodegradable — anti-sense DNA analogues with unusual backbone linkages and peptide nucleic acids (PNAs) are difficult to break down. However, suitably designed nanorobotic agents constructed of nonbiological materials are not biodegradable. An engineered fibroblast may not stimulate an immune response when transplanted into a foreign host, but its biomolecules are subject to chemical attack in vivo by free radicals, acids, and enzymes. Even “mirror” biomolecules or “Doppelganger proteins”\textsuperscript{2285} comprised exclusively of unnatural D-amino acids have a lifetime of only ~5 days inside the human body.\textsuperscript{2286} Nonbiological materials such as diamond and sapphire are highly resistant to chemical breakdown or leukocytic degradation in vivo, and pathogenic biological entities cannot easily evolve useful attack strategies against these materials (Section 9.3.5.3.6).

5. **Control of Nanomedical Treatment** — Present-day biotechnological entities are not programmable and cannot be switched on and off conditionally during task execution. A digital biocomputer, while possible in theory, represents a considerable conceptual departure from the usual biological paradigm. Even assuming that a digital biocomputer could be installed in a fibroblast, and that appropriate effector mechanisms could be attached, such a system would necessarily have slower clock cycles, less capacious memory per unit volume, and longer data access times, implying less diversity of action, poorer control, and less complex executable programs than would be available in nanoscale electromechanical computer systems (Section 10.2). The mechanical approach emphasizes precise control of action, including control of physical placement, timing, structure, and interactions with other (especially biological) entities. The biological approach emphasizes the use of poorly controlled natural structures, needlessly sacrificing huge blocks of the available functionality and design space.

6. **Nanodevice Versatility** — Mechanical systems can readily incorporate biological elements if necessary, but artificial biological
systems can incorporate nonbiological materials such as carbon nanotubes or diamond/sapphire structural elements only with difficulty, in part because biology has a more limited repertoire of “effector” mechanisms. Artificial biological systems cannot easily incorporate nonbiological materials where desired because natural biological assembly methods make no provisions for these materials either in the coded instructions in DNA or in the attachment chemistries. Rebuilding or reconstructing the human body with nonbiological components (e.g. fullerenic encased bone for bone damaged in an accident; Chapters 24 and 30), or augmentation of human body function with unnatural abilities or features (e.g. autogenous paracrine control; Chapter 12), will be very difficult or impossible to achieve using purely biotechnological means.

7. Avoiding Overspecialization — M. Krummenacker notes that one of the most glaring shortcomings of bacteria and other naturally occurring molecular machinery — when viewed as systems subject to further engineering — is the rather limited range of molecular substrates they can utilize. For example, bacterial enzymes are highly specialized devices, with very narrow substrate specificities. Thousands of different enzymes are needed in each organism, and the substance classes capable of digestion are limited to some sugars, various amino acids including proteins that can be degraded by excreted proteases, lipids, and a few other smallish oxygen-functionalized carbon molecules such as glycerol and ethanol. Some bacteria can metabolize CO₂ and a handful of aromatic compounds, but there is a vast range of organic chemicals that most bacteria cannot degrade or manufacture. A much smaller set of substantially more general molecular tools can probably be designed using the mechanical approach; mechanosynthesis can fabricate and assemble, or disassemble, a far wider range of molecular structures than are available to the cellular machinery of life.

8. Faster and More Precise Diagnosis — The analytic function of medical diagnosis requires rapid communication between the injected devices and the attending physician. If limited to chemical messaging, biotechnology devices will require minutes or hours to complete each diagnostic loop. Nanomachines, with their more diverse set of input-output mechanisms (Chapter 7), can outmessage the results of in vivo reconnaissance or testing literally in seconds. Such nanomachines can also run more tests of greater variety in less time. Mechanical nanoinstruments, including molecule-by-molecule disassemblers (Chapter 19), will make comprehensive cell mapping and cell interaction analysis possible. Bacterial resistance can be assayed at the molecular level, allowing new treatment agents to more easily be composed, manufactured and immediately deployed (Chapter 19).

9. More Sensitive Response Threshold for High-Speed Action — Unlike natural systems, an entire population of nanobiotic devices can be triggered globally by just a single local detection of the target antigen or pathogen. The natural immune system takes >10⁵ sec to become fully engaged after exposure to a systemic pathogen or other antigen-presenting intruder. A biotechnologically enhanced immune system that can employ the fastest natural unit replication time (-10³ sec for some bacteria) will require ~10⁴ sec full deployment post-exposure. By contrast, a nanobiotic immune system (Chapter 19) can probably be fully engaged (though not finished) in at most two blood circulation times, or ~10⁻² sec.

10. More Reliable Operation — Engineered macrophages would probably individually operate less reliably than mechanical nanorobots. Many pathogens, such as *Listeria monocytogenes* and *Trypanosoma cruzi*, are known to be able to escape from phagocytic vacuoles into the cytoplasm while biotech drugs or cell manufactured proteins could be developed to prevent this (e.g. cold therapy drugs are entry-point blockers), nanorobotic trapping mechanisms can be more secure (Section 10.4.2). Proteins assembled by natural ribosomes typically incorporate one error per ~10⁴ amino acids placed; current gene and protein synthesizing machines utilizing biotechnological processes have similar error rates. A molecular nanotechnology approach will improve error rates by at least a millionfold (Chapter 20). Mechanical systems can also incorporate sensors to determine if and when a particular task needs to be done, or when a task has been completed. Finally, it is unlikely that natural organisms will be able to infiltrate mechanical nanorobots or to co-opt their functions. By contrast, a biological-based robot could be diverted or defeated by microbes that can piggyback on its metabolism, interfere with its normal workings, or even incorporate the device wholesale into their own structures, causing the engineered biomachine to perform some new or different function than was originally intended. There are many examples of such co-option among natural biological systems, including the protozoan mixotrichs found in the termite gut that have assimilated bacteria into their bodies for use as motive engines, and nudibranch mollusks (marine snails without shells) that steal nematocysts (stinging cells) away from coelenterates such as jellyfish (i.e. a Portuguese man-of-war) and incorporate the stingers as defensive armaments in their own skins, a process which S. Vogel has called “stealing loaded guns from the army.”

11. Verification of Progress and Treatment — Using a variety of communication modalities, nanorobots can report back to the attending physician, with digital precision, a summary of diagnostics and therapy. The receiving computer can then forward this information to the attending physician, with digital precision, a summary of diagnosis and therapy. Nanorobots can communicate with the attending physician, with digital precision, a summary of diagnosis and therapy. Nanorobots can communicate with the attending physician, allowing her to modify treatment as needed. A biological-based biomachine can’t perform in this fashion. The natural immune system takes minutes or hours to complete each diagnostic loop. Nanomachines, with their more diverse set of input-output mechanisms (Chapter 7), can outmessage the results of in vivo reconnaissance or testing literally in seconds. Such nanomachines can also run more tests of greater variety in less time. Mechanical nanoinstruments, including molecule-by-molecule disassemblers (Chapter 19), will make comprehensive cell mapping and cell interaction analysis possible. Bacterial resistance can be assayed at the molecular level, allowing new treatment agents to more easily be composed, manufactured and immediately deployed (Chapter 19).

12. Minimum Side Effects — Almost all drugs have side effects, such as conventional cancer chemotherapy which causes hair loss and vomiting, although computer-designed drugs (Chapter 18) have high specificity and relatively few side effects. Carefully tailored cancer vaccines under development in the late 1990s were expected unavoidably to affect some healthy cells. Even well-targeted drugs are distributed to unintended tissues and organs in low concentrations, although some bacteria can target certain organs fairly reliably without being able to distinguish individual cells. By contrast, mechanical nanorobots may be targeted with virtually 100% accuracy to specific organs, tissues, or even individual cellular addresses within the human body (Chapter 8). Such nanorobots should have few if any side effects, and will remain safe even in large dosages because their actions can be digitally self-regulated using rigorous control protocols (Chapter 12) that affirmatively prohibit device activation unless all necessary preconditions have been, and continuously remain, satisfied. G.M. Fahy has noted that these possibilities could transform “drugs” into “programmable machines with a range of sensory, decision-making, and effector capabilities [that] might avoid side effects and allergic reactions...attaining almost complete specificity of...
action....Designed smart pharmaceuticals might activate themselves only when, where, and if needed. Additionally, nanorobots may be programmed to excuse themselves from the site of action, or even from the body, after a treatment is completed; by contrast, spent biorobotic elements containing ingested foreign materials may have more limited post-treatment mobility, thus lingering at the worksite causing inflammation when naturally degraded or removed.*

13. Reduced Replicator Danger — Drexler points out that living systems are evolved systems, while nanomechanical replicators would be designed: “The former are shaped to serve the goal of their own survival and replication in a natural environment, whereas the latter will be shaped (whether well or poorly) to serve human goals, perhaps in an artificial environment.” Genetic engineering involves not design of replicators from scratch, but tinkering with the molecular machinery of existing bioreplicators. Since bioreplicators were not designed, they are not necessarily structured in ways that lend themselves to complete understanding, and processes based on diffusion and matching allow complex nonlocal interactions that can be hard to trace. Bioreplicators can be crippled, but having evolved in nature, they resemble systems that can survive in nature. Typically, they are able to exchange genetic information with wild organisms, raising the possibility of the introduction of new, unconstrained replicators in the natural environment. Having evolved to evolve, they have a capacity for further evolution — to serve their own survival, not human goals. For example, even when stripped of key pieces of DNA to interfere with its replication powers, a live attenuated AIDS vaccine can slowly recover its virulence when stripped of key pieces of DNA to interfere with its replication. There is no requirement for replication in vivo; such replication would be needlessly dangerous, and adding this capability would reduce effectiveness in carrying out the primary medical task. Analogously, viral vectors employed in genetic therapies are modified to be “incapable” of replication.

14. Assured Patentability — Microscopic biorobots, unavoidably derived from natural biological material, may someday be deemed unpatentable under a general prohibition on “genetic colonialism” or other emerging legal doctrines. In contrast, mechanical nanorobots, being fully-artificial and designed machines, should always be patentable provided they satisfy the customary legal criteria.

1.3.4 Naturophilia

As already noted (Section 1.3.1), living things in general and the human body in particular are awesome examples of a powerful and intricately woven natural molecular technology to which human engineers, in 1998, still aspire. But embracing Nature is not the same as finding her perfect. Today, the word “natural” has acquired a strong connotation of rightness, even of sanctity. For most of human history, notes biologist Steven Vogel, “the natural and human worlds stood opposed. Nature was something to be tamed and utilized; we had the ordinary attitude of organisms toward other species. Nowadays the natural world intrudes far less but gets venerated far more. And why not? When one’s meat is bought in a store, when locusts don’t threaten one’s corn crop, when central heating and plumbing are the norm, the aesthetics of nature hold greater appeal.” And so we embrace the natural rectitude or moral superiority of nature’s ways, a kind of pantheism which may be called ethical naturalism, or biophilia, or naturophilia.

Many great minds have fallen prey to naturophilia. In the 4th century BC, Aristotle wrote that “if one way be better than another, that you may be sure is Nature’s way.” In the 15th century, we have from Leonardo da Vinci: “Human ingenuity may make various inventions, but it will never devise any inventions more beautiful, nor more simple, nor more to the purpose than Nature does; because in her inventions nothing is wanting and nothing is superfluous.”

However, as Virginia Postrel notes in The Future and Its Enemies, “If nature is itself a dynamic process rather than a static end, then there is no single form of the natural.” An evolving, open-ended nature may impose practical constraints, but it cannot dictate eternal standards. It cannot determine what is good. The distinction between the artificial and the natural must lie not in their source — human or not — but in their characteristics, in the way they relate to the world around them.”

According to the dictionary, “artificial” usually means “made by man, rather than occurring in nature.” More usefully, Herbert Simon defines the artificial as that which is designed, expresses goals, and possesses external purposes. The artificial is controlled and serves its creator’s purposes, subject to the universal laws of physics. Kevin Kelly defines the natural as “out of control.” Nature is evolved, not designed, and serves no goal or external purpose save its own survival. Nature, lacking intent, is amoral — it simply is. By building the artificial, observes Postrel, “we do not overthrow nature, but cooperate with it, using nature’s own art to create new

*R. Bradbury notes that artificial eukaryotic biorobots may possess an apoptotic pathway (Section 10.4.1.1) which could be activated to permit clean and natural self-destruction that avoids inflammation in surrounding tissues. Artificial prokaryotic biorobots could also be designed for human biocompatibility; by replacing bacterial genes with generic human genes, such a device may be undetectable to the immune system.
natural forms. Our artifice alters the path of nature, but it does not end it, for nature has no stopping point, no final shape. It is a process, not an end."

Some naturophilist writers have decried the increasing "medicalization of society" in which formerly natural functions have come to be regarded as medical conditions requiring intervention or treatment. However, history suggests that naturophilia is usually undermined by any new medical technology that offers clear, safe, and immediate benefits to patients. For example, prior to 1842, intense pain was viewed as the natural outcome of being cut with a scalpel during surgery. It had always been so — how could it ever be otherwise? The invention of anesthesia in 1842 (Section 1.2.1.9.2) suddenly altered this natural outcome and replaced it with a less painful artificial outcome, despite anguished cries from naturophiles within the medical community that eliminating pain might somehow diminish the human character.

Another example of a widely-accepted medicalization of normal function is childbirth, a quite natural activity that can nevertheless be very dangerous to the mother's health. Precise prehistoric death rates are unknown, although archeological evidence shows that Neandertal females tended to die before the age of 30 due to hazards of childbirth, In the worst 19th century maternity hospitals the natural death rate from childbirth was 9–10%, falling to a very artificial 0.4% rate in England by 1930 and to less than 0.01% in the U.S. during the 1990s. As a result, it now seems "natural" for a woman always to survive childbirth, even though the reverse may have been true for most of human history. Warns historian Roy Porter, "We should certainly not hanker after some mythic golden woman always to survive childbirth, even though the reverse may have been true for most of human history."

A disease seems "natural" to those who suffer from it when no treatment exists. But once a treatment is discovered and is widely employed, the disease becomes rare and its absence now becomes "natural." To those in the past, writes K. E. Drexler, "the idea of cutting people open with knives painlessly would have seemed miraculous, but surgical anesthesia is now routine. Likewise with bacterial infections and antibiotics, with the eradication of smallpox, and the vaccine for polio: each tamed a deadly terror, and each is now half-forgotten history. What amazes one generation seems obvious and even boring to the next. The first baby born after each breakthrough grows up wondering what all the excitement was about." In the next century, says Charles Sheffield, "our descendants will look on angiograms, upper and lower GIs, and biopsies the way we regard the prospect of surgery without anesthetics."

Future generations who take for granted an all-pervasive nanomedicine in their lives may look aghast upon the 20th century, wondering among other things how we managed to retain productive focus given the constant annoyance of our numerous undiagnosed minor disease states. Most of these diseases are not yet recognized as such, and many are still regarded as "natural" and not worthy of treatment. In a few decades, this may change. Some examples:

1. Addictions — In 1998, many people laugh off seemingly harmless addictions to chocolate (chocoholics), fats or sugar (sweet tooth), food (gluttony), nicotine (smokers), caffeine (coffee and cola drinkers), work (workaholics), exercise (runner's high), telling falsehoods (pathological liars), gambling (wagerphilia), stealing (kleptomania), medical treatments (hypochondria), marriage (polygamy), power-seeking (domination), skydiving or bungee-jumping (thrillseeking), superstition (astrology), shopping (spendoholics), driving cars that kill 40,000 Americans per year (mobilophilia), unusual sexual preferences (bestiality), sexual activity (nymphomania, satyriasis), or pregnancy (gravidiophilia). Without making any value judgements, it is highly likely that most or all of these addictions have genetic or physiological components which, once properly modified, can greatly reduce or eliminate the addiction if so desired. Many on the list are already suspected to have genetic components, much like schizophrenia, drug abuse, bulimia, and alcoholism (dipsomania).

2. Allergies and Intolerances — A food allergy is an allergic reaction to a particular food, although true food allergies are much rarer than is generally believed. In the cases of milk, eggs, shellfish, nuts, wheat, soybeans, and chocolate, sufferers may lack an enzyme necessary for digesting the substance. In other cases, dust particles, plant pollens, pet danders, drugs, or foods may be allergens for natural IgE-mediated immunosensitivity. Intolerance, a much more common condition, is any undesirable effect of eating a particular food, including gastrointestinal distress, gas, nausea, diarrhea, or other problems. Urticaria (hives), angioedema and even mild anaphylaxis are common reactions to various drugs, insect stings or bites, allergy shots, or certain foods, particularly eggs, shellfish, nuts and fruits. Physical allergies to ordinary stimuli such as cold, sunlight, heat, pollen, pet dander, or minor injury can produce itching, skin blotsches, pimples, and hives.

3. Minor Physical Annoyances — In a world where most major medical maladies are readily treated, numerous minor medical conditions which today escape our notice will rise up from obscurity and present themselves annoyingly to our conscious minds, demanding attention. These conditions may be of several kinds. First is cosmetics, including small moles, freckles and blemishes on the skin; broken fingernails or unevenly-growing cuticles; minor skin reddenings or pimples; old childhood scars, wrinkled skin, birthmarks or stretch marks; unwanted hair growth in unusual places, or differential hair color or texture growing in patches; fingerprint patterns that are aesthetically unappealing; and mismatched leg lengths, hands with different left/right ring sizes, an asymmetrical face, or lopsided breasts. Second is minor aches and pains, which may include headaches; eyebrow hairs trapped in the eyeball conjunctiva; bent-hair pain (folliculalgia); dyspepsia; cracking limb joints and stomach growling; ingrown nails and hairs; carwash plugs and temporary tinnitus; chapped lips, canker sores and heat rashes; stuffy nose or gitty eyes upon rising in the morning; dermal chafing marks from elastic bands in clothing; minor flatulence; PMS (premenstrual syndrome); a leg or arm "falling asleep" in certain postures; nervous ticks, itches, and twitches; uncracked knuckles, stiff neck, or backache; blocked middle ear following descent from high altitude; restless leg syndrome (akathisia); rotationally-induced dizziness, as on an amusement park ride; or rock-and-roll neck, wherein active musical performers or listeners bob their heads violently, rupturing small blood vessels in the neck. Third is minor physical or functional flaws, such as poor stream during male urination, female papillary leakage, colorblindness, snoring, unpleasant body odors, nosebleeds, declining visual or aural acuity, handedness (currently ~90% dextromanual, ~10% sinistromanualmild strabismus (eyeball misalignment), bad moods (neurotransmitter imbalances), or post-intoxication hangover.

4. Undiscovered Infectious Agents — Peptic ulcers once were thought to result from a stressful life, a purely natural response to a lifestyle choice. Then it was found that the major cause of ulcers is the
presence of Helicobacter pylori bacteria in the stomach. Bacteria have been implicated in some cases of atherosclerosis\textsuperscript{2970} and Alzheimer's disease,\textsuperscript{2971} and nanobacteria have been proposed as possible nucleation sites for kidney stones.\textsuperscript{2409} Other seemingly natural but undesirable conditions may also be due to undiscovered microbial agents,\textsuperscript{3237} especially since bacteria outnumber tissue cells in our highly infested 20th century bodies by more than 10:1 (Section 8.5.1).

5. \textit{Unwanted Syndromes} — Syndromes are groups of related symptoms and signs of disordered function that define a disease whose cause remains unknown, that is, idiopathic. A good example is irritable bowel syndrome (IBS), which affects up to 20\% of the adult U.S. population and includes symptoms of abdominal distention and pain, with more frequent and looser stools. Many are unaware they are afflicted. In 1998 there was no known cause or simple complete treatment for this still "natural" disease.\textsuperscript{3713-3717} Even more mysterious than IBS is our general activity level — some people seem to have high-energy personalities, while others have more phlegmatic low-energy personalities. Either may be regarded as "natural," but nanomedicine can probably bring this ill-defined neurophysiological variable under human control. The need to sleep is another imperfectly understood syndrome. It is experienced by everyone and thus was universally regarded as "natural" in the 20th century. Physiological short sleepers\textsuperscript{2121} were unusual, insomnia or asomnia\textsuperscript{2275} was thought of as an abnormal state, and there were a few anecdotal but medically undocumented instances of total nonsomnia, such as the celebrated case of Al Herpin.\textsuperscript{2312}

6. \textit{Psychological Traits} — Psychological traits which, if identified by a patient as undesirable, might be subject to genetic or physiological modification could include: sexual preference (6-10\% of the adult population is homosexual),\textsuperscript{1604} shyness or boldness,\textsuperscript{2332} acquisitive modification could include: sexual preference (6-10\% of the adult population is homosexual),\textsuperscript{1604} shyness or boldness,\textsuperscript{2332} acquisitive modification could include: sexual preference (6-10\% of the adult population is homosexual),\textsuperscript{1604} shyness or boldness,\textsuperscript{2332} such as antisocial, paranoid, schizotypal, histrionic, narcissistic, avoidant, dependent, obsessive-compulsive, and passive-aggressive disorders; panic attacks (experienced at least once by ~33\% of all adults each year);\textsuperscript{1604} and phobic disorders such as social phobias (~13\% of the population), specific phobias including fear of large animals (zoophobia), needles (belonephobia, \textsuperscript{3272} ~10\%), the dark (necrophobia) or scotophobia), or strangers (xenophobia) (total ~5.7\%), the fear of blood or hemophobia (~5\%), agoraphobia (~2.8\%),\textsuperscript{1604} and other unusual phobias\textsuperscript{2223} such as the fears of certain colors (chromophobia), daylight (phengophobia), girls (parthenophobia), men (androphobia), stars in the sky (siderophobia), the number thirteen (triskaidekaphobia), and even the fear of developing a phobia (phobophobia).

The above sampling of minor afflictions, almost all considered "natural" in 1998, may come to be regarded as commonplace correctable medical conditions in the nanomedical era. By the time such petty annoyances are deemed worthy of immediate treatment, biotechnology and nanomedicine already will have defeated the most fearsome illnesses of the late 20th century\textsuperscript{2318} and will have moved on to other challenges.\textsuperscript{2511,2864,2973} Naturophiles may dissent, but the emerging trend from medical biotechnology is to characterize health, not as a static standard, but rather as a condition defined by the lives that people want to lead. Affirming the volitional normative model of disease (Section 1.2.2), Virginia Postrel concludes:\textsuperscript{2259}

"Different goals will produce different choices about trade-offs and standards. What makes a condition unhealthy is not that it is unnatural but that it interferes with human purposes. Revering nature [would mean] sacrificing the purposes of individuals to preserve the world as given. It [would require] that we force people to live with biological conditions that trouble them, whether diseases such as cystic fibrosis or schizophrenia, disabilities such as myopia or crooked teeth, or simply lack of beauty, intelligence, happiness, or grace than could be achieved through artifice. In a world where it's no big deal to take hormone therapy, Viagra, or Prozac, to have a face lift, or to know a child's sex before birth, a world in which even such radical interventions as exchange operations and heart transplants have failed to turn society upside down, it is extremely difficult to argue that medical innovations are dangerous simply because they fool Mother Nature."

1.4 Background and Brief Overview of This Book

The publication of \textit{Nanosystems}\textsuperscript{10} in 1992 laid a strong technical foundation for the field of molecular manufacturing. Such technology was already known to have major implications for medicine.\textsuperscript{8,9} Still lacking, however, was an overview and synthesis of the interactions between molecular machine systems and living systems, particularly human living systems.

With \textit{Nanosystems} as the intellectual cornerstone, research on \textit{Nanomedicine}, originally intended to be a single volume, began in 1994. It quickly became clear that the field of nanomedicine would be even more interdisciplinary than molecular manufacturing, in part because of the many essential interfaces between mechanical nanosystems and living systems. This realization prompted ramification of the book into multiple volumes, allowing the field to be defined from diverse technical orientations. Additionally, the need to consider all major aspects of nanomedical device design and operations demanded a substantial increase in the intended length of the work, a regrettable but necessary circumstance for which the author apologizes in advance.

As a result, \textit{Nanomedicine} has become a three-volume technical work with 31 chapters. Its intended audience is technical and professional people who are seriously interested in the future of medical technology. The three Volumes build upon each other cumulatively. The first Volume, now complete, describes basic capabilities common to all medical nanodevices, and the physical, chemical, thermodynamic, mechanical, and biological limits of such devices. Its primary audience is physical scientists, chemists, biochemists, and biomedical engineers engaged in basic research. The second Volume, still in progress, deals with aspects of device control and configuration, biocompatibility and safety issues, and basic nanomedical components and simple systems. Its primary audience is systems and control engineers, research physiologists, clinical laboratory analysts, biotechnologists, and biomedical engineers doing applied research. The third Volume, also in progress, discusses specific treatments for specific conditions and injuries, using nanomedical technology in the context of clinical (e.g. doctor-patient) situations. Its primary audience is clinical specialists and research physicians, and interested general practitioners.

The first molecular assemblers (Chapter 2) may be able to build only very simple nanomechanical systems, possibly only from very highly ordered substrates with significant rate-limiting intermediate steps, and only in very small numbers, so the earliest functional nanomachines may be laboratory curiosities. As assembler technology slowly improves, progressively more complex and capable nanomachines will be manufactured in vastly larger numbers. This book primarily investigates the rational design and operation of these more complex and capable nanomachines, and assumes that cubic centimeter quantities (e.g. \texttimes 10\textsuperscript{12} micron-sized nanorobots) will not
be unreasonably expensive to manufacture (Section 2.4.2) and thus can be therapeutically deployed routinely by future doctors.

Volume I of Nanomedicine, Basic Capabilities, describes the set of basic capabilities of molecular machine systems that may be required by many, if not most, medical nanorobotic devices. These include the abilities to recognize, sort and transport important molecules (Chapter 3); sense the environment (Chapter 4); alter shape or surface texture (Chapter 5); generate onboard energy to power effective robotic functions (Chapter 6); communicate with doctors, patients, and other nanorobots (Chapter 7); navigate throughout the human body, i.e. determining somatographic or cytographic location within vessels, organs, tissues, or cells (Chapter 8); manipulate microscopic objects and move about inside a human body (Chapter 9); and timekeep, perform computations, disable living cells and viruses, and operate at various pressures and temperatures (Chapter 10).

Volume II of Nanomedicine, Systems and Operations, considers system-level technical requirements in the design and operation of medical nanodevices. Part 1 describes aspects of nanomedical operations and configurations, including scaling factors and general design principles (Chapter 11); control issues including teleoperation and haptic controllers, swarm motions, autogenous control systems, and various operational protocols (Chapter 12); repair, replacement, and reliability issues (Chapter 13); and molecular machine systems design issues such as tradeoffs between special-purpose and general purpose architectures, and deployment configurations such as nano-organs, medical utility fogs, and replicators (Chapter 14). Part 2 deals with a multitude of issues involving clinical safety and performance, specifically medical nanorobot biocompatibility including immunoreactivity and thrombogenicity (Chapter 15); methods of nanorobotic ingress and egress from the human body (Chapter 16); and possible nanodevice failure modes, environmental interactions, side effects of nanomedical treatments, iatrogenic factors, nanodevice software bugs, and other safety issues (Chapter 17). Part 3 summarizes various classes of medical nanosystems, including instruments, tools, and diagnostic systems (Chapter 18); specific medical nanorobot devices (Chapter 19); rapid mechanical reading and editing of chromatin and protein macromolecules (Chapter 20); and various complex nanorobotic systems that will make possible advanced cytopathology and cell repair, tissue and organ manufacturing, and personal defensive systems (Chapter 21).

Volume III of Nanomedicine, Applications, describes the full range of nanomedical applications which employ molecular nanotechnology inside the human body, from the perspective of a future practitioner in an era of widely available nanomedicine. Proof-of-concept designs for whole nanodevices, artificial nano-organs, and nanomedical treatments include rapid cardiovascular repair (Chapter 22); treatments for pathogenic disease and cancer, with epidemiological considerations (Chapter 23); responses to various physical traumas, burns and radiation exposures, and new methods of first aid, surgery, and emergency or critical care (Chapter 24); neurography, spinal restoration and brain repair (Chapter 25); improved nutrition and digestion (Chapter 26); sex, reproduction, and population issues (Chapter 27); cosmetics, recreation, veterinary and space medicine (Chapter 28); the control of aging processes, eliminating most causes of death prevalent in the 20th century, and strategies for biostasis (Chapter 29); and human augmentation systems (Chapter 30). The Volume concludes with a discussion of the sociology of nanomedicine, regulatory issues, nanotechnology implementation timelines, and some speculations on the future of hospitals, pharmaceutical companies, and the medical profession (Chapter 31).
2.1 Is Molecular Manufacturing Possible?

Most contemporary industrial fabrication processes are based on “top-down” technologies, wherein small objects are sawn or machined from larger objects, or small features are imposed on larger objects, in either case by removing unwanted matter. The results of such processes may be small, such as micron-featured integrated circuits, or very large, such as jet aircraft, but in most cases the material is being processed in chunks far larger than molecular scale.

Molecular manufacturing, on the other hand, represents a “bottom-up” technology. Desired products will be built directly by “assembler” machines, molecule by molecule, making larger and larger objects with atomic precision. The results of such processes may also be very small or very large, much as biology builds both micron-sized bacteria and 100-meter tall sequoia trees. However, since assemblers add matter only where it is intended, little need be removed and hence there may be minimal waste during the process. By guiding with precision the assembly of molecules and supramolecular structures, such a manufacturing system could construct an extraordinarily wide range of products of unprecedented quality and performance.

In 1959, the Nobel physicist Richard P. Feynman observed:

“The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done; but, in practice, it has not been done because we are too big . . . Ultimately, we can do chemical synthesis.”

“A chemist comes to us and says, “Look, I want a molecule that has the atoms arranged thus and so, make me that molecule.” The chemist does a mysterious thing when he wants to make a molecule. He sees that it has got that ring, so he mixes this and that, and he shakes it, and he fiddles around. And, at the end of a difficult process, he usually does succeed in synthesizing what he wants.

“But it is interesting that it would be, in principle, possible (I think) for a physicist to synthesize any chemical substance that the chemist writes down. Give the orders and the physicist synthesizes it. How? Put the atoms down where the chemist says, and so you make the substance. The problems of chemistry and biology can be greatly helped if our ability to see what we are doing, and to do things on an atomic level, is ultimately developed—a development which I think cannot be avoided.”

Nearly 40 years after Feynman’s famous “Plenty of Room at the Bottom” speech, and a decade after Drexler’s original proposal for a bottom-up approach to machine-building using molecular assemblers, Nobel chemist Richard Smalley also largely agreed that this objective should prove feasible. Noted Smalley: “On a length scale of more than one nanometer, the mechanical robot assembler metaphor envisioned by Drexler almost certainly will work…”

Many skeptical questions arise when one first encounters the ideas of molecular nanotechnology and molecular assemblers (Section 2.4.2). It is useful to keep in mind the proven feasibility of such systems in the biological tradition (Section 1.3.2.1) over billions of years of natural evolution. In one sense, molecular nanotechnology will be a refinement and expansion upon how nature works at the molecular scale. Nature’s examples, such as human beings, certainly answer the most basic skeptical questions, such as: Can macroscopic objects be built from molecular scale processes? (Yes, thanks to cellular replication.) Are molecular objects stable? (Of course; the human population alone contains \(>10^{29}\) reasonably stable and quite functional ribosomal “nanomachines”). Observes Nobel chemist Jean-Marie Lehn: “The chemist finds illustration, inspiration and stimulation in natural processes, as well as confidence and reassurance since they are proof that such highly complex systems can indeed be achieved on the basis of molecular components.”

What about quantum effects? The uncertainty principle makes electron positions somewhat fuzzy, but the atom as a whole has a comparatively definite position set by the relatively great mass of the atomic nucleus. The quantum probability function of electrons in atoms tends to drop off exponentially with distance outside the atom, giving atoms a moderately sharp “edge”. Mathematically, the positional uncertainty of a single carbon atom of mass \(m_C = 2 \times 10^{-26}\) kg bound in a single C-C bond of stiffness \(k_C = 440\) N/m may be crudely estimated from the classical vibrational frequency \(\nu_C = (k_C/m_C)^{1/2} = 1.5 \times 10^{14}\) Hz. This sets the zero-point vibrational bond energy \(E_C = h \nu_C / 2 = 4.9 \times 10^{-20}\) J = \(k_C \Delta x_C^2 / 2\) where \(h = 6.63 \times 10^{-34}\) J-sec (Planck’s constant) and \(\Delta x_C ~ 0.015\) nm is the maximum classical amplitude of the bound carbon atom (roughly the same as the 3 dB point for the gaussian wavefunction, notes J. Soreff). Thus \(\Delta x_C\) is just \(-5\%\) of the typical atomic electron cloud diameter of \(-0.3\) nm, imposing only a modest additional constraint on the fabrication and stability of nanomechanical structures. (Even in most liquids at their boiling points, each molecule is free to move only \(-0.07\) nm from its average position.)

How about the effects of Brownian bombardment on nanomachines? Describing a nanomechanical component as a harmonic oscillator embedded in a gas, Drexler notes: “At equilibrium, an impinging gas molecule is as likely to absorb energy as to deliver it, and so molecular bombardment has no net effect on the amplitude of vibration. How a system is coupled to a thermal bath can affect its detailed dynamics, but not the statistical distribution of dynamical quantities.” Nanomachines must also obey the laws...
of thermodynamics. Nanorobots cannot be used as a “Maxwell’s demon” — energy must be expended to do useful work; there can be no free lunch.

Will high-energy radiation damage nanomachines? Radiation can break chemical bonds and disrupt molecular machines. The annual failure rate for a properly designed ~1 micron molecular machine (containing many billions of atoms) can be as low as several percent. This estimate is derived from a consideration of all known thermal, photochemical, radiation, and other damage mechanisms, but the dominant error mechanism that appears difficult to substantially reduce is damage caused by background radiation. Failure rates are not zero, but are nonetheless remarkable by today’s standards (Chapter 13), especially given error detection and correction at the module level.

During nanodevice fabrication, the problems of crossbonding and reactive intermediates pose additional constraints that must be dealt with using appropriate assembler designs. Specifically, assemblers with rigid positioning components operating in vacuo can restrict reactive intermediates and molecular tool tips to desired locations with a precision of at least ~0.1 nm (Section 9.3.1.4), or slightly less than an atomic diameter. Within the confines of an evacuated or “eutactic” assembler workspace, the location of every structural atom may be known to within the uncertainty created by thermal noise. Internal structural components may possess relatively high stiffness, so positional uncertainty caused by thermal noise may be small. In the classical analysis, positional variance $\Delta X^2 = k T / k_s$, where $k = 1.381 \times 10^{-21}$ J/kelvin (Boltzmann’s constant), $T$ is temperature in kelvins, and $k_s$ is component stiffness, typically $\geq 10$ N/m for nanometer-scale diamondoid components, hence the standard deviation of component position is ~0.02 nm at room temperature, or ~0.01 nm at cryogenic (e.g., LN$_2$) temperatures. By employing sufficiently stiff mechanisms the position of every structural atom in the system can be known to within a fraction of an atomic diameter with high reliability and without the need for explicit positional sensing.* Steric tool hindrance—in molecular manufacturing, the conical volume occupied by a tool tip or manipulator platform—is a technical design problem that may be overcome by using sufficiently stiff tools or multi-tip tools, working in vacuo. Additionally, while some synthetic methods might involve the manipulation of individual atoms by appropriate tools, such as the hydrogen abstraction tool which removes a single selected hydrogen atom from a diamondoid surface (Section 2.3.3), other reactions will involve small clusters of atoms or larger molecular components (Section 2.4.1).

What about friction and wear among nanomechanical components? Properly designed molecular machines lack wear mechanisms, although other damage mechanisms remain. Dry bearings with atomically precise surfaces can have negligible static friction and wear, and very low dynamic friction or drag. Nanomechanical components are better viewed as moving smoothly in a force field than as sliding subject to friction. For example, total drag power dissipated by an isolated molecular sleeve bearing of radius 2 nm spinning at 1 MHz is dominated by band-stiffness scattering, amounting to $\sim 0.000004$ picowatts (pW) (Eqn. 6.4), very small compared to the 1-1000 pW power draw that will be typical of micron-size medical nanorobots (Section 6.5.3). The strong precise surfaces of nanomachines experience no change during a typical operational cycle, hence zero wear. Within the single-point failure model (Chapter 13), the first step in a wear process (e.g., a dislocated atom) is regarded as fatal, hence cumulative wear plays no role in determining device lifetimes. In eutactic nanomachines, contaminants of all kinds are rigorously excluded. The equivalent of wear particles cannot appear until the device has failed. On this scale, repulsive fields provide “lubrication” and an oil molecule would be a contaminant object, not a lubricant.

The general conclusion is that molecular nanotechnology violates no physical laws. In 1998, it was generally accepted that molecular nanotechnology will be developed, although there remains some disagreement about how long it will take. To progress from today’s limited capabilities to the complex nanorobots described later in this trilogy will clearly require a great deal of research and development effort. Still, the proper question is no longer “if” but “when.” This Chapter presents a few of the many possible technical paths leading to a working molecular assembler that were being discussed in 1998. The reader is strongly urged to consult the most current literature for the latest results. Section 2.2 briefly summarizes the conventional top-down approaches to nanotechnology. Section 2.3 describes several possible direct bottom-up pathways to molecular manufacturing including biotechnology, supramolecular chemistry, and scanning probes. Gimzewski and Joachim note that “bottom-up approaches to nanofabrication may one day compete with conventional top-down approaches in providing nanotechnologies for the next millennium. Top-down refers to increasing miniaturization through extension of existing microfabrication schemes, [whereas] the bottom-up scenario is one of ever-increasing complexity at the molecular level while maintaining control on an atom-by-atom basis”.

2.2 Top-Down Approaches to Nanotechnology

Nanotechnology was first proposed by Nobel physicist Richard Feynman in December 1959, in a talk in which he also issued a seemingly “impossible” challenge to build a working electric motor no larger than a 1/64th-inch (400-micron) cube, backed by a $1000 prize to spur interest in the new field. Just 11 months later, engineer William McLellan had constructed a 250-microgram 2000-rpm motor out of 13 separate parts and collected his reward. (McLellan’s entire motor is only as big as the period at the end of this sentence.) In 1995, a $250,000 Feynman Grand Prize (Section 2.4.2) was made available, this time sponsored by the Foresight Institute, for any engineer who could build the first programmable nanometer-scale robotic arm. How long will it take for history to repeat?

In his famous 1959 talk, Feynman proposed the prototypical top-down strategy for building complex nanomachinery—essentially a completely teleoperated machine shop, including mills, lathes, drills, presses, cutters, and the like, plus master-slave grippers to allow the human operator to move parts and materials around the workshop.

To build a nanomachine using Feynman’s scheme, the operator first directs a macroscale machine shop to fabricate an exact copy of itself, but four times smaller in size. After this work is done, and all machines are verified to be working properly and as expected, the reduced-scale machine shop would be used to build a copy of itself, another factor of four smaller but a factor of 16 tinier than the original machine shop. This process of fabricating progressively smaller machine shops proceeds until a machine shop capable of manipulations at the nanoscale is produced. The final result is a nanomachine shop capable of reconstructing itself, or of producing...
any other useful nanoscale output product stream that is physically possible to manufacture, using molecular feedstock.

As Feynman describes the process:

“When I make my first set of slave hands at one-fourth scale, I am going to make ten sets. I make ten sets of hands, and I wire them to my original levers so they each do exactly the same thing at the same time in parallel. Now, when I am making my new devices one-quarter again as small, I let each one manufacture ten copies, so that I would have a hundred hands at the 1/16th size... If I made a billion little lathes, each 1/4000th the scale of a regular lathe, there are plenty of materials and space available because in the billion little ones there is less than two percent of the materials in one big lathe. It doesn’t cost anything for materials, you see. So I want to build a billion tiny factories, models of each other, which are manufacturing simultaneously, drilling holes, stamping parts, and so on.”

“As we go down in size, there are a number of interesting problems that arise. All things do not simply scale down in proportion. There is the problem that materials stick together by the molecular (van der Waals) attractions. There will be several problems of this nature that we will have to be ready to design for...[But] if we go down far enough, all of our devices can be mass produced so that they are absolutely perfect copies of one another. We cannot build two large machines so that the dimensions are exactly the same. But if your machine is only 100 atoms high, you only have to get it correct to one-half of one percent to make sure the other machine is exactly the same size—namely, 100 atoms high!”

Progress is being made on Feynman’s top-down approach in a relatively new engineering field known as Micro Electro-Mechanical Systems or MEMS, originally an extension of chip-etch technology rather than micromanipulation. Conventional purely-electronic devices fabricated on silicon chips with ~0.2 micron features must be coupled to external systems to produce mechanical effects, but MEMS devices integrate mechanical components directly with electrical circuitry. Thus while microelectronics chips merely route electrons, microelectromechanical systems give the electronics immediate access to control applications, enabling single microdevices to interact directly with the physical world.

Generic microsystems research has proven to be a rewarding activity for two decades, with each successively smaller motor or machine widely publicized.2685 The field gathered momentum after the first working micromotors were demonstrated in the late 1980s by groups at Berkeley and MIT. By 1990, tiny electrostatic motors with 0.2 micron features must be coupled to external systems to produce mechanical effects, but MEMS devices integrate mechanical components directly with electrical circuitry. Thus while microelectronics chips merely route electrons, microelectromechanical systems give the electronics immediate access to control applications, enabling single microdevices to interact directly with the physical world.

Progress is being made on Feynman’s top-down approach in a relatively new engineering field known as Micro Electro-Mechanical Systems or MEMS, originally an extension of chip-etch technology rather than micromanipulation. Conventional purely-electronic devices fabricated on silicon chips with ~0.2 micron features must be coupled to external systems to produce mechanical effects, but MEMS devices integrate mechanical components directly with electrical circuitry. Thus while microelectronics chips merely route electrons, microelectromechanical systems give the electronics immediate access to control applications, enabling single microdevices to interact directly with the physical world.

Generic microsystems research has proven to be a rewarding activity for two decades, with each successively smaller motor or machine widely publicized.2685 The field gathered momentum after the first working micromotors were demonstrated in the late 1980s by groups at Berkeley and MIT. By 1990, tiny electrostatic motors with 0.2 micron features must be coupled to external systems to produce mechanical effects, but MEMS devices integrate mechanical components directly with electrical circuitry. Thus while microelectronics chips merely route electrons, microelectromechanical systems give the electronics immediate access to control applications, enabling single microdevices to interact directly with the physical world.

Some MEMS research followed surprising directions. For example, in the mid-1990s MEMS engineers at the University of Tokyo Department of Mechanical Engineering fabricated an 8-hinge, 8-plate wing microstructure capable of independent flight through the air.1573-1576.2353 The structure was fabricated by traditional silicon micromachining techniques and was then released from the wafer surface, with the initially flat sections lifted into flight position using a micromanipulation system with glass probes several microns in diameter. Each wing was ~500 microns square; applying 300 VAC at 10 KHz caused resonant vibration and the flapping amplitude of the wings exceeded 30°.2353 In another 4-wing configuration,1574 the flapping resonance frequency was ~150 Hz and the device freely flew up and down a thin glass pole when placed in an alternating electromagnetic field. Another bi-wing design made of magnetized nickel coated silicon measuring 2000 x 10,000 microns successfully
flew without power supply cables and guides in the manner of a butterfly or mosquito, when an alternating magnetic field strength of \(\sim 400\) oersted and \(12\) Hz was applied;\(^{1577}\) a similar device with \(320 \times 800\) micron wings also flew in a \(300\) Hz magnetic field.\(^{1575}\) “If we release the silicon mosquito from the silicon chip, it flies off and we cannot find it again,” complained researcher H. Miura. “It’s very small, like dust.” MEMS ‘gnat robots’ have also been described.\(^{1250}\)

Using electron-beam (e-beam) lithography, in 1997 researchers at the Cornell University Nanofabrication Facility built what they believed were the world’s smallest mechanical devices, including a 4-micron-wide Fabry-Perot interferometer and, just for fun, the world’s smallest guitar, carved out of crystalline silicon and no larger than a single human cell.\(^{2354}\) The “nanoguitar” was 10 microns long and 2 microns wide, with six strings each ~50 nm (~200 atoms) wide that would probably resonate at \(>10-100\) MHz if plucked. Presentation of the nanoguitar enthralled the public, surprising at least one Cornell researcher (in another lab) who remarked that “anyone with a sufficiently advanced e-beam lithography system could make it. I make 50 nanometer silicon objects routinely. Doesn’t seem like a big deal to me, but evidently most people have no idea what we’re capable of these days.”

Direct-write e-beam lithography uses a tightly-focused beam of electrons, steered by electromagnetic deflectors, to trace out patterns on a wafer resist surface. The electrons chemically alter the resist, which is then etched away, leaving the desired pattern. By 1998, e-beam spot sizes had been focused to less than 1 nm, and 5-nm beam widths were routine, although forward scattering of the electrons imposed a limit of resist exposure resolution of ~10 nm for e-beams\(^{2358}\) and ~5 nm for focused ion beams.\(^{1259}\) E-beam writing times were very slow—for example, 80 hours were required to carve a single photolithography mask for a 1G DRAM, though much of this time was required for proximity calculations by the AI (artificial intelligence) software [E.A. Rietman, personal communication, 1999] (a better approach is to use a neural net).\(^{3012}\) By 1998, an electron interferometer containing atomically smooth mirrors spaced a few atomic layers apart had been fabricated,\(^{3183}\) and in 1999, careful electron-energy-loss spectroscopy experiments revealed that silicon dioxide must be at least 4 atoms (~0.7 nm) thick to act as a conventional electrical insulator.\(^{3295}\)

Atom lithography, employing a low-energy neutral atomic beam (usually Na or Cr) cooled and collimated by optical molasses, then concentrated at the nodes of the standing wave formed by the superposition of a laser beam and its reflection, was predicted also eventually to allow surface deposition of 5-10 nm features.\(^{2355,2370,2838-2841}\) The first directional atomic beam (of sodium) also eventually to allow surface deposition of 5-10 nm features.\(^{2355,2370,2838-2841}\) Atom lithography, employing a low-energy neutral atomic beam (usually Na or Cr) cooled and collimated by optical molasses, then concentrated at the nodes of the standing wave formed by the superposition of a laser beam and its reflection, was predicted also eventually to allow surface deposition of 5-10 nm features.\(^{2355,2370,2838-2841}\) The first directional atomic beam (of sodium) was demonstrated in early 1999.\(^{2925}\) By 1998, an electron interferometer containing atomically smooth mirrors spaced a few atomic layers apart had been fabricated,\(^{3183}\) and in 1999, careful electron-energy-loss spectroscopy experiments revealed that silicon dioxide must be at least 4 atoms (~0.7 nm) thick to act as a conventional electrical insulator.\(^{3295}\)

At the same time, another group achieved a continuous stream of rubidium atoms lasting up to ~0.1 sec in duration, with a beam radius potentially as small as ~1 nm.\(^{3205}\) Current-carrying wires have been used to guide cold lithium atoms “just as optical fibers guide light”;\(^{3182}\) the first curved focusing atom mirror also was demonstrated in early 1999.\(^{2925}\) By 1998, the speculative possibility of quantum-mechanical micromanipulation also had been described.\(^{6}\)

Other approaches offered comparable results. For instance, engineers at the University of Minnesota NanoStructure Laboratory used nanoimprint lithography (NIL) in 1996 to etch out patterns of lines, grooves and circles as small as 25 nanometers in a polymer,\(^{2357}\) and by 1998 NIL had been demonstrated in nanoscale quantum-wire, quantum-dot, and ring transistors with feature sizes below 10 nm.\(^{2358}\) Nanochannel array glasses containing 33-nm capillaries arranged in a 2-D hexagonal close packing configuration at a number density of \(\sim 300/\text{micron}^2\) had been available since 1992.\(^{2356}\)

Thus by 1998, top-down MEMS techniques could in theory already make nanoscale (though not atomically precise) parts. Might these techniques also produce assemblies of such parts, creating complex machines? Very simple mobile robots of ~1 cm\(^3\) volume were commonplace,\(^{2361,2562}\) so for a more challenging demonstration of MEMS’ ability to manufacture complete working microrobots, in 1994 Japanese researchers at Nippondenso Co., Ltd. fabricated a 1/1000th-scale working electric car.\(^{2351,2352}\) As small as a grain of rice, the micro-car was a 1/1000-scale replica of the Toyota Motor Corp’s first automobile, the 1936 Model AA sedan (Fig. 2.1). The tiny vehicle incorporated 24 parts, including tires, wheels, axles, headlights and taillights, bumpers, a spare tire, and hubcaps carrying the company name inscribed in microscopic letters, all manually assembled using a mechanical micromanipulator of the type generally used for cell handling in biological research (Chapter 21). In part because of this handcrafting, each microcar cost more to build than a full-size modern luxury automobile.

The Nippondenso microcar was 4800 microns long, 1800 microns wide, and 1800 microns high, consisting of a chassis, a shell body, and a 5-part electromagnetic step motor measuring 700 microns in diameter with a ~0.07-tesla magnet penetrated by an axle 150 microns thick and 1900 microns long. Power was supplied through thin (18 micron) copper wires, carrying 20 mA at 3 volts. The motor developed a peak torque of \(1.3 \times 10^{-6}\) N-m (mean \(7 \times 10^{-7}\) N-m) at a mean frequency of \(\sim 100\) Hz (peak frequency \(\sim 700\) Hz), propelling the car forward across a level surface at a top speed of 10 cm/sec. Some internal wear of the rotating parts was visible after ~2000 sec of continuous operation; the addition of ~0.1 microgram of lubricant to the wheel microbearings caused the mechanism to seize due to lubricant viscosity. The microcar body was a 30-micron thick 20-milligram shell, fabricated with features as small as ~2 microns using modeling and casting, N/C machine cutting, mold etching, submicron diamond-powder polishing, and nickel and gold plating processes. Measured average roughness of machined and final polished surfaces was 130 nm and 26 nm, respectively. The shell captured all features as small as 2 mm on the original full-size automobile body. Each tire was 690 microns in
diameter and 170 microns wide. The license plate was 10 microns thick, 380 microns wide and 190 microns high. Nippondenso subsequently used similar manufacturing techniques to build a prototype of a capsule intended to crawl through tiny pipes in a power plant or chemical plant like an inchworm, hunting for cracks. In 1999, three Japanese electronics companies announced the creation of a 0.42-gram, 5-mm long “ant size” robot reportedly able to lift 0.8-gram loads and to move at ~2 mm/sec, as part of the government’s ongoing Micro Machine Project.3258

By 1998, micromanipulators with high placement precision had been demonstrated in various laboratories around the world. For example, the Spider-II micromanipulation robot employed bimorphic piezoactuators within a 260-micron cubic work volume and a three-axis gripper placement accuracy of 8 nm.2359 A microassembly robot using stick-and-slip actuators had 5-nm resolution over a 200-nm scanning range with a maximum speed of 4 mm/sec and 0.4% repeatability at 10 KHz, delivering a maximum driving force of 155 mN and rigidity of 6.3 N/micron.2360

In terms of system size, micromachined MEMS devices (typically ~10-13 m³ in volume) lie exactly intermediate between the macroscale world (~10⁻⁴ m³) and the nanoscale world (~10⁻²² m³). The medical nanorobots described in this book generally will have dimensions, precisions, and component part sizes ~1000-fold smaller than the Nippondenso microcar, which itself was, in turn, ~1000-fold smaller than a macroscale automobile. It is difficult to see how MEMS techniques involving statistical materials deposition and removal could fabricate parts to better than ~10 nm feature sizes or tolerances, or could position parts during assembly to much better than ~10 nm spatial resolution, nor does it appear likely that MEMS techniques, on their own, can manufacture or position any structure to atomic precision. However, this does not rule out the possible use of MEMS techniques to fabricate crude useful parts for subsequent “finishing” to atomic precision by other techniques, or to facilitate the assembly of nanoscale components (of atomic precision or otherwise) that have been fabricated or subassembled by other means (Section 2.3).

2.3 Bottom-Up Pathways to Molecular Manufacturing

To manufacture machines of any kind generally requires two primary capabilities—fabrication of parts, and assembly of parts. As will be shown below, by 1998 at least a primary parts fabrication and parts assembly capability had been demonstrated at the molecular level using three different enabling technologies:2889 biotechnology (Section 2.3.1), supramolecular chemistry (Section 2.3.2), and scanning probes (Section 2.3.3). This Section reviews the most preliminary steps and basic approaches to molecular manufacturing. Consideration of more complex nanoscale components and assemblies is deferred to Section 2.4. The primary focus here is on bridging technologies that may lie between what can already be done and what is desired to be done—specifically, the manufacture of mechanical nanorobots. The reader is again cautioned that in 1998 this field was experiencing an explosion of interest and pathbreaking research efforts, so the current literature should be consulted for the latest results.

2.3.1 Biotechnology

The first broad category of enabling technology for molecular manufacturing in the mechanical tradition is biotechnology.10,322 Molecular biologists study and modify systems of molecular machines, and genetic engineers reprogram these systems, sometimes to build novel molecular objects having complex functions. By 1998, biotechnologists could make almost any DNA sequence or polypeptide chain desired (though not extremely long aperiodic ones), assemble “devices” such as artificial chromosomes and viruses, and were actively pursuing gene therapy, all of which are examples of molecular engineering. While earlier writers had suggested that biomolecules could be used as mechanical components,2414 Drexler182 evidently was the first to point out, in 1981, that complex devices resembling biomolecular motors, actuators, bearings, and structural components could be combined to build versatile molecular machine systems analogous to machine systems in the macroscopic world (Table 1.3). “Development of the ability to design protein molecules,” Drexler wrote, “will open a path to the fabrication of devices to complex atomic specifications, thus side-stepping obstacles facing conventional microtechnology. This path will involve construction of molecular machinery able to position reactive groups to atomic precision.”

Perhaps the best-known biological example of such molecular machinery is the ribosome (Fig. 2.2; see also Section 8.5.3.4), the only freely programmable nanoscale assembler already in existence.* In nature, ~8000 nm³ ribosomes act as general-purpose factories building diverse varieties of proteins by bonding amino acids together in precise sequences under instructions provided by a strand of messenger RNA (mRNA) copied from the host DNA, powered by the decomposition of adenosine triphosphate (ATP) to adenosine diphosphate (ADP). Each ribosome is a compact ribonucleoprotein particle consisting of two subunits, with each subunit consisting of

![Fig. 2.2. The ribosome acts as a programmable nanoscale assembler of protein nanoparticles (redrawn from Ball182).](image)

---

*D. Heidel notes that there are three distinct types of ribosome—eubacterial, archaeabacterial and eukaryotic—and multiple subvarieties among the ~10⁸ different species on Earth; additionally, other bioassembler units such as poly-ketide synthetases also exist but these enzymes generally only make one type of product per assembly.
several proteins associated with a long RNA molecule (tRNA). Biotechnology can employ the ribosomal machinery of bacteria to produce novel proteins, which proteins then might serve as components of larger molecular structures.\(^{182}\)

Of course, such biological mechanisms have been disparaged as “slip-shod devices working in spite of haphazard design.” Dan Heidel, a University of Washington biomimetics engineer, agrees that biological systems are sloppy but observes that they are designed to tolerate it, and offers high praise for the lowly ribosome:

“The average ribosome sits in an aqueous environment surrounded by thousands of cosolutes at concentrations just a hair’s breadth from precipitating out of solution. The ribosome itself consists of three RNA molecules that spontaneously fold into flappy spaghetti noodle piles of chemically active groups surrounded by >50 proteins of similarly dubious structure. The other solutes are bombarding this motley assemblage at tens to hundreds of miles an hour and the solvating water molecules are a constantly shifting and unpredictable network of hydrogen bonds and polar ionic charges constantly impinging upon the ribosome. The ribosome must recognize one particular cognate transfer RNA (tRNA) for the codon in its A-site out of ~60 other nearly identical tRNAs. This particular tRNA must be discriminated from all its tRNA brethren along with whatever other solutes can fit into the A-site pocket using only three base pairs for discrimination. In fact, most codons use only two of the three base pairs for recognition so some tRNAs are recognized by as few as 4 hydrogen bonds. All of this requires sub-Angstrom precision in the active sites but since the ribosome is a massive complex of subunits held together by only ionic, hydrogen, hydrophobic and Van der Waal bonds, it’s like getting sub-millimeter accuracy from points on a bunch of taped-together water balloons being rolled down a staircase.”

“This ridiculous machine can assemble proteins at a 20 Hz subunit-incorporation frequency with a 10\(^{-3}\) error rate. Ribosomes are even more impressive when one considers that the number of amino acids in each ribosome’s own protein subunits combined with the 10\(^{-3}\) protein assembly error rate ensures that each ribosome has several errors in it and that each cell almost certainly has no two ribosomes exactly alike. I dare any engineer on this planet to nationally design an assembler at any scale that is able to operate at a comparable level of performance under similarly adverse operating conditions.”

Thus ribosomes can manufacture, with >99.9% fidelity to arbitrary specifications, linear strings of amino acids of great length which may then fold up to produce three-dimensional protein structures performing a broad range of highly differentiated functions, providing good analogs for working nanomachines capable of forming self-assembling systems. Much of the great diversity of protein activity flows from the selectivity with which they bind to specific receptor sites on other molecules. Some of these binding mechanisms might be copied to construct molecular robotic arms, while other devices could be made by adapting naturally-occurring proteins, or by synthesizing new, custom-built ones.\(^{2393}\)

Protein engineering might be used to create proteins with desired qualities as building blocks for constructing the first primitive assemblers.

The field of protein engineering has its own journal of the same name, and has already achieved remarkable results in the synthesis of novel structures. Drexler\(^{10}\) cites the creation of the first de novo structure by DeGrado,\(^{2394}\) the development of synthetic branched protein-like structures that depart substantially from protein models,\(^{2395}\) and the engineering of a branched, nonbiological protein with enzymatic activity.\(^{2396}\) Since then, numerous alpha-helical peptides\(^*\) have been designed de novo,\(^{2410,2411}\) a 20-residue artificial peptide exhibiting designed beta-sheet secondary structure has been created,\(^{2398}\) and various peptide analogs have also been synthesized.\(^{2399}\) By 1998, advances in computational techniques allowed the design of precise side-chain packing in proteins with naturally occurring backbone structures, and the study of de novo backbone structures had begun.\(^{2400}\)

The catalytic task space\(^{266}\) was being laboriously investigated at the molecular level—for example, one study of comparative enzyme structures revealed that changing as few as four amino acids converted an oleate 12-desaturase to an oleate hydroxylase, and as few as six substitutions could convert an oleate hydroxylation into an oleate desaturase.\(^{2401}\) To further expand the toolkit, Mills\(^{2431}\) and Fahy\(^{322}\) proposed exploiting the degeneracy of the genetic code (wherein amino acids are specified by up to 6 different codons; Chapter 20) by assigning some codons to unnatural amino acids, allowing the creation of artificial proteins with unprecedented structural or catalytic properties. With 61 usable codons available, there is room for up to 41 novel amino acids in addition to the natural 20 amino acids; unnatural amino acids have been introduced into beta-lactamase and T4 lysozyme enzymes in a site-directed fashion, using artificial tRNAs and mRNAs in an in vitro translation system,\(^{2402,2403}\) and into an engineered peptide as a reporter group.\(^{2412}\) Additionally, artificial novel base pairs can be incorporated enzymatically into DNA\(^{2404,2405,2406}\) and other schemes are possible.\(^{2407}\)

By incorporating one additional base pair, the number of possible codons is expanded from 4\(^3\) = 64 codons to 6\(^3\) = 216 codons, potentially allowing up to 193 novel amino acids to be incorporated into artificial protein structures—or at least 48 new amino acids assuming that present levels of redundancy are retained and the existing 4-letter code is kept as a subset for “upward compatibility.”\(^{2431}\)

If the genetic alphabet is increased to eight bases (8\(^3\) = 512 codons), the number of available codons for novel amino acid-like molecules increases to as many as 489,\(^{322}\)

One of the greatest challenges in protein engineering (where one objective is to create functional, atomically-precise 3-D aperiodic structures) has been the difficulty of protein fold design, because individual amino acids have no strong, natural complementarity.\(^{2392}\) In 1998, Duan and Kollmann\(^{2321}\) successfully folded a solvated 36-residue (~12,000-atom) protein fragment (i.e., a peptide) by molecular dynamics simulation into a structure that resembles an intermediate native state. During a 1-microsec simulation, the chain folded during 150 nanosec into a compact structure resembling the intermediate native state, as known by NMR, then unfolded and refolded again for a shorter period. In 1998, whole-protein folding to the final stable native state had not yet been computed deterministically using molecular dynamics—the real protein will fold and refold hundreds of times before it stumbles into the stable conformation with the lowest free energy—\(^{2405}\)—but research pursuing this objective was active and ongoing.\(^{2406,2407}\) Structural changes that occur during protein function (e.g., enzymatic action) were also being avidly studied.\(^{2408}\)

A different protein-based approach to near-atomically-precise 3-D structures avoids the folding problem by making use of semirigid protein nanoshells of deterministic size and shape to guide the ordered assembly of inorganic particles. For example, the capsid of the cowpea chlorotic mottle virus, a protein container with a 28-nm external diameter and an 18-nm internal cavity diameter, is composed of 180 identical coat protein subunits arranged on an icosahedral lattice. Each subunit presents at least nine basic residues (arginine and lysine) to the interior of the cavity, creating a positively charged interior surface which provides an interface for inorganic crystal nucleation and growth, while the outer capsid surface is not highly charged. In one experiment,\(^{2391}\) the empty capsid shell was found to act as a spatially selective nucleation catalyst in paratungstate mineralization, in addition to its role as a size- and shape-constrained

---

\(^*\)Chains of amino acids shorter than ~100 residues are customarily called peptides; longer chains are proteins.
difficulties were surmounted. Horn and Urdea reported branched field accelerated in the 1990s after numerous experimental materials; Alivisatos, Schultz and colleagues used DNA to organize spacings, with the goal of assembling nanoparticles into macroscopic could be formed with these molecules. Mirkin's group attached DNA to rigid organic linkers, showing that cyclical forms of various sizes create a new class of hybrid nanomechanical devices.” Other natural 9.3.4). Montemagno, genetically engineered the 12-nm wide over hand down the line in the manner of a ciliary array (Section 9.2.5) and the motors were activated, passing 25-nm wide microtubules hand over hand in the line in the manner of a ciliary array (Section 9.3.4). Montemagno explained that his laboratory’s “long-term goal is the integration of the F1-ATPase biological motor with nano-electromechanical systems (NEMS) to create a new class of hybrid nanomechanical devices.” Other natural nanomotors such as the self-assembling bacterial flagellar motor and the hexagonal “packaging RNA” or prRNA pumping mechanism that packs DNA into the capsid shell of the bacteriophage Phi 29 are also being intensively studied.

Yet another material which allows moderately well-defined three-dimensional nanoassemblies is DNA. The ideas behind DNA nanotechnology have been around since 1980, but activity in this field accelerated in the 1990s after numerous experimental difficulties were surmounted. Horn and Ude reported branched and forked DNA polymers. Niemeyer and Smith exploited DNA specificity to generate regular protein arrays, and both suggested using self-assembled DNA as an early material for molecular nanotechnology. Shi and Bergstrom attached DNA single strands to rigid organic linkers, showing that cyclical forms of various sizes could be formed with these molecules. Mirkin's group attached DNA molecules to 13-nm colloidal gold particles with -6 nm particle spacings, with the goal of assembling nanopolyhedra into macroscopic materials; Alivisatos, Schultz and colleagues used DNA to organize 1.4-nm gold nanocrystals in arrays with 2-10 nm spacings. In either case, the specificity of DNA pairing should allow the construction of complex geometries since the use of colloidal particles can potentially add structural elements (which are more rigid than any polymer strand) to the set of building blocks for nanometer-scale structures—which are also floppy. In 1993 Seeman discovered the more rigid anti-parallel DNA “double crossover” motif, which in 1996 he used to design and build a stiff double junction to keep his structures from sagging. The next goal was to bring together a large number of stick figures to form large arrays or cage-shaped DNA crystals that could then be used as frameworks for the assembly of other molecules into pre-established patterns. These DNA molecules would serve as the scaffolding upon which new materials having precise molecular structure could be assembled.

In 1998 the most intensive work on three-dimensional engineered DNA structures was taking place in Nadrian Seeman's laboratory in the New York University Department of Chemistry. Seeman originally conceived the idea of rigid 3-D DNA structures in the early 1980s, while examining DNA strands that had arranged themselves into unusual four-armed Holliday junctions. Seeman recognized that DNA had many advantages as a construction material for nanomechanical structures. First, each double-strand DNA with a single-strand overhang has a “sticky end,” so the intermolecular interaction between two strands with sticky ends is readily programmed (due to base-pair specificity) and reliably predicted, and the local structure at the interface is known (sticky ends associate to form B-DNA). Second, arbitrary sequences are readily manufactured using conventional biotechnological techniques. Third, DNA can be manipulated and modified by a large variety of enzymes, including DNA ligase, restriction endonucleases, kinesins and exo-nucleases. Fourth, DNA is a stiff polymer in 1-3 turn lengths and has an external code that can be read by proteins and nucleic acids.

During the 1980s, Seeman worked to develop strands of DNA that would zip themselves up into more and more complex shapes. Seeman made junctions with five and six arms, then squares, stick-figure cubes comprised of 480 nucleotides, and a truncated octahedron containing 2550 nucleotides and a molecular weight of ~790,000 daltons. The cubes (Fig. 2.3) were synthesized in solution, but Seeman switched to a solid-support-based methodology in 1992, greatly improving control by allowing construction of one edge at a time and isolating the growing objects from one another, allowing massively parallel construction of objects with far greater control of the synthesis sequence. By the mid-1990s, most Platonic (tetrahedron, cube, octahedron, dodecahedron, and icosahedron), Archimedean (e.g., truncated Platonic, semiregular prisms and prisms, cuboctahedron, etc.), Catalan (linked rings and complex knots), and irregular polyhedra could be constructed as nanoscale DNA stick figures.

Seeman's DNA strands that formed the frame figures were strong enough to serve as girders in a molecular framework, but the junctions were too floppy. In 1993 Seeman discovered the more rigid anti-parallel DNA “double crossover” motif, which in 1996 he used to design and build a stiff double junction to keep his structures from sagging. The next goal was to bring together a large number of stick figures to form large arrays or cage-shaped DNA crystals that could then be used as frameworks for the assembly of other molecules into pre-established patterns. These DNA molecules would serve as the scaffolding upon which new materials having precise molecular structure could be assembled.

In 1998, Erik Winfree and colleagues in Seeman's laboratory reported the design and construction of two-dimensional DNA crystals using self-assembling double crossover molecules. Repeating array units were approximately 2 nm x 4 nm x 16 nm in size, and examination of the array with an Atomic Force Microscope (AFM) revealed domains of up to 500,000 interconnected units, showing that the self-assembly process (Section 2.3.2) could be very reliable under ideal conditions. While the work described the construction of two- and four-unit lattices, the number of component tiles that could be used in the repeat unit did not appear to be limited to such small numbers, suggesting that complex patterns could be assembled into periodic arrays, which might also serve as templates for nanomechanical assembly. Noted the authors: “Because oligonucleotide synthesis can readily incorporate modified bases at arbitrary positions, it should be possible to control the structure within the periodic group by decoration with chemical groups, catalysts, enzymes and other proteins, metallic nanostructures,
to be converted from the normal B-DNA structure to the unusual Z-DNA structure. The free ends of the arms shift position by ~2.0 nm during the structural conversion. Explained Seeman: “Using synthetic DNA as a building material, we have constructed a controllable molecular mechanical system. In the long-term, the work will have implications for the development of nanoscale robots and for molecular manufacturing.”

DNA can also serve as an assembly jig in solution phase. Bruce Smith and colleagues are devising a method for the assembly and covalent linkage of proteins into specific orientations and arrangements as determined by the hybridization of DNA attached to the proteins, called DNA-Guided Assembly of Proteins (DGAP). In this method, multiple DNA sequences would be attached to specific positions on the surface of each protein, and complementary sequences would bind, forcing protein building blocks (possibly including biomolecular motors, structural protein fibers, antibodies, enzymes, or other existing functional proteins) together in specific desired combinations and configurations, which would then be stabilized by covalent interprotein linkages. This technique could also be applied to non-protein components that can be functionalized at multiple sites with site-specific DNA sequences, although proteins, at least initially, may be more convenient building blocks due to their size, their surface chemistry, the wide variety of functions and mechanical properties they can confer on the resulting assemblies, and the many existing techniques for introducing designed or artificially evolved modifications into natural proteins of known structure. (In 1998, custom DNA and peptide sequences could be ordered online.).

Methods for covalently attaching functional proteins to a DNA backbone in a specified manner at ~8.5 nm (25 base-pair) intervals, addressable protein targeting in macromolecular assembly, and “protein stitchery” were being explored by others. Drexler notes that evolution has not maximized the stability of natural proteins, and that substantially greater stability may be engineered by various means (e.g., increasing folding stability by >100 Kcal/mole).

2.3.2 Molecular and Supramolecular Chemistry

A second broad category of enabling technology for molecular manufacturing includes molecular chemistry—the conscious design of completely artificial, nonbiological chemical structures that could potentially serve as molecular “parts” or which have specific device-like functionality. For example, catalysts may very loosely “be thought of as rudimentary assemblers that are slightly programmable through changes in the reaction milieu (i.e., changes in pH, temperature, etc.).” The number of natural molecular “parts” is immense—perhaps a few hundred thousand bio-organic compounds have been separated, purified and identified—but by 15 February 1999 chemists had already registered 19,245,458 well-characterized artificial molecules or other “substances” in the CAS Registry. Synthesis of natural molecules containing ~100 atoms (e.g., a ~1 nm³ molecular “part”), using methods of classical organic synthesis often requiring ~1 synthesis step per atom with ~90% yield per step, was state-of-the-art in 1998.

“What is exciting about modern nanotechnology,” says Nobel chemist Roald Hoffmann, “is (a) the marriage of chemical synthetic talent with a direction provided by ‘device-driven’ ingenuity coming from engineering, and (b) a certain kind of courage provided by those incentives, to make arrays of atoms and molecules that ordinary, no, extraordinary chemists just wouldn’t have thought of trying. Now they’re pushed to do so. And of course they will. They can do anything.”
Molecular chemists study how small numbers of atoms combine covalently with each other to produce molecules with a wide variety of different properties. Supramolecular chemists are primarily concerned with the much weaker noncovalent forces that can arise between molecules—such as hydrogen bonds and van der Waals interactions—that can be sufficient to bind collections of molecules tightly enough to form functional nanometer-sized structures. Supramolecular chemistry studies "chemistry beyond the molecule" or "the chemistry of the noncovalent bond". Its subject matter includes atomically precise supermolecules (well-defined discrete oligomolecular species resulting from the intermolecular association of a few components, such as a receptor and its substrate) and polymer-like supramolecular assemblies (polymolecular entities that result from the spontaneous association of a large undefined number of components).

Self-assembly is a key concept in supramolecular chemistry as it is in nature, allowing the manufacture of large numbers of compound objects simultaneously and in parallel, rather than sequentially. Molecular self-assembly is a strategy for nanofabrication that involves designing molecules and supramolecular entities so that complementarity causes them to aggregate into desired structures. Self-assembly of atomically precise supermolecules thus demands well-defined adhesion between selected molecules, which may require steric (shape and size) complementarity, interactional complementarity, large contact areas, multiple interaction sites, and strong overall binding.

G.M. Whitesides explains that self-assembly has a number of advantages as a strategy. First, it carries out many of the most difficult steps in nanofabrication—those involving the smallest atomic-level modifications of structure—using the very highly developed techniques of synthetic chemistry. Second, it draws from the enormous wealth of examples in biology for inspiration; self-assembly is one of the most important strategies used in biology for the development of complex, functional structures, such as the well-studied examples of complex biomachine self-assembly of whole bacteriophage virions and flagellar rotor motors from their smaller molecular "parts". Third, self-assembly can incorporate biological structures directly as components in the final systems. Fourth, self-assembly requires target structures to be among the most thermodynamically stable available to the system, thus tends to produce structures that are relatively defect-free and self-healing. As Lehmann points out: "By increasing the size of its entities, nanochemistry works its way upward towards micro lithography and microphysical engineering, which, by further and further miniaturization, strive to produce ever smaller elements". Hierarchical self-assembly of many billions of micron-scale spherical components into a periodic honeycomb-structured photonic crystal optical device -30 microns thick and -1 cm^2 in area was demonstrated in 1998.

There is a wide range of different molecular systems that can self-assemble, including those which form ordered monomolecular structures by the coordination of molecules to surfaces, called self-assembled monolayers (SAMs), self-assembling thin films, and Langmuir-Blodgett films, and self-organizing nanostreams. In these systems, a single layer of molecules affixed to a surface allows both thickness and composition in the vertical axis to be adjusted to 0.1 nm by controlling the structure of the molecules comprising the monolayer, although control of in-plane dimensions to <100 nm is very difficult. Fluidic self-assembly of microscale parts and the dynamics of Brownian self-assembly have also been described, and the theory of designable self-assembling molecular machine structures is beginning to be addressed.

Another approach using solid-phase peptide synthesis and a massively convergent self assembly process was employed by M.R. Ghadiri and colleagues, who designed and synthesized a number of self-assembling peptide-based nanotubes using cyclic peptides with an even number of alternating D- and L-amino acids for the building blocks of the nanotubes. The alternating stereochemistry of the cyclic peptides allowed all the side chains of the amino acids to be pointing outwards which would not be possible in an ordinary all L-cyclic peptide. In this conformation, the amide backbone can H-bond in a direction perpendicular to the plane of the cyclic peptide. The stacking of two cyclic peptides forms an H-bonding network resembling an anti-parallel beta-sheet as is commonly found in natural proteins. The H-bonding lattice quickly propagates perpendicular to the plane of the cyclic peptide, forming a tubular microcrystalline structure with 0.75-nm pores, or, in another experiment, 1.3-nm pores. Another group of nanotubes was designed with a highly hydrophobic outer surface and a hydrophilic inner pore. These nanotubes were easily inserted into a lipid bilayer and have been shown to be highly efficient ion channels.

Self-assembled crystalline solids such as zeolites incompletely fill space, leaving substantial voids that may be occupied by solvent molecules or other guest molecules, producing solid-state host-guest complexes known as clathrates (from the Latin clathratus, meaning “enclosed by the bars of a grating”). MacNicol and co-workers reported the first rationally designed clathrate host in 1978, but the first true de novo design of an organic clathrate was in 1991 and involved the formation of an extensive three-dimensional diamond-like porous lattice built from a single Tinkertoy-type subunit containing four tetrahedrally arrayed pyridone groups that acted as connectors to assemble the units together in a well-defined geometry. However, the crystal was held together only by weak hydrogen bonds and the links had a rotational degree of freedom, rendering the exact crystal molecular arrangement unpredictable. Engineered nanoporous molecular crystals can provide molecular-scale voids with controlled sizes, shapes, and embedded chemical environments at resolutions of 0.3-4.0 nm. Other self-assembling crystalline structures have been described including the crystal engineering of diamondoid networks, template-directed colloidal crystallization or colloidal epitaxy, organic templating to form crystalline zeolite-type structures with ordering lengths <3 nm, 3-D polymer channels with chemically functionalizable channel linings, and micromolding combined with templating and cooperative assembly of block copolymers (Section 3.5.7.1) to produce hierarchical ordering over discrete and tunable characteristic length scales of ~10 nm, ~100 nm, and ~1000 nm in a single body.

Dendrimers, also known as arborols or fractal polymers, are a well-known set of self-assembling structures that possibly could be used as tools to assist in the assembly of early nanomachines. Dendrimers are large regularly-branching macro molecules resembling fractal patterns, made by an iterative process in which small linear molecules are allowed to bind to each other at a certain number of sites along their length, building up branches upon branches with each iteration working outward from a core...
molecule having at least two chemically reactive arms, somewhat mimicking tree growth (e.g., Fig. 2.20C). Different core molecules or building-block chains produce macromolecules with different shapes, and the outer surface can be terminated with chemical groups of specific functionality, producing hundreds of differently-surfaced branched molecules\textsuperscript{2671} that have already found use in medical research.\textsuperscript{2671} Growth is regulated, so size can be accurately controlled—dendrimers are typically a few nanometers wide but have been constructed with masses exceeding a million protons (\(\approx 10^5\) atoms) and with diameters larger than 30 nm. In 1999, convergent assembly of specified 6-mer dendrimers, via trimerization of precursor dimer “parts,” was demonstrated.\textsuperscript{3275}

Autocatalysis or self-replication (Section 2.4.2 and Chapter 14) is yet another approach to the chemical self-assembly of subunits. In 1989, J. Rebek\textsuperscript{131,2450-2452} reported the synthesis of supermolecules that could generate copies of themselves when placed in a sea of simpler molecular parts, with each component part consisting of up to a few dozen atoms. Molecular self-replication has continued to be studied by others.\textsuperscript{2524-2526,2602} Ghadiri\textsuperscript{2457} reported a self-replicating peptide in 1996, and by 1998 his laboratory had moved on to studies of more complex autocatalytic cycle networks.\textsuperscript{2458-2660} It is unknown how complex such assembly cycles may be made, but Ghadiri was investigating a self-replicating 256-component molecular ecosystem incorporating ~32,000 possible binary interactions\textsuperscript{2601} in 1998.

By 1998, chemists had already synthesized a vast variety of simple molecular parts of which only a tiny sampling can be mentioned here. For instance, carbon rings have been linked edge-on or by corners to create propeller-shaped molecules called propellanes\textsuperscript{2497} (Fig. 2.4) or rotanes (Fig. 2.5), or stacked with short hydrocarbon links to make gear-looking molecules such as cyclophane\textsuperscript{2493} and superphane (Fig. 2.6). Hydrocarbon molecular polyhedra in the shape of triangular (e.g., triangulane\textsuperscript{2571}), cubic (e.g., cubane, first made by P. Eaton in 1964\textsuperscript{2570}), pentagonal and hexagonal prisms, collectively known as prismanes, have been created (Fig. 2.7), along with more unusual geometrical forms such as paddlane, housane, basketane, churchnane, pagodane, and bivalvane\textsuperscript{382} (Fig. 2.8; one carbon atom at each vertex). Cryptands,\textsuperscript{2479} spherands,\textsuperscript{2492} cryptaspherands and carcerands,\textsuperscript{410} corannulene baskets,\textsuperscript{2639} and calixarenes\textsuperscript{2601} are bowl-shaped molecules whose benzene ring walls hold their cavities rigid. The bowl rim of a calixarene (Fig. 2.9) can be lined with chemical groups to determine which guests it will accept, or two bowls can be bonded together,\textsuperscript{2547} making a hollow nanoscale reaction vessel, and various molecular cages and capsules with precise sizes and characteristics have been assembled.\textsuperscript{1262,2549-2551} A variety of nonchiral and chiral square molecules have been synthesized,\textsuperscript{2473,2543} as well as molecular “bottlebrushes”;\textsuperscript{2475} DNA-polymer constant-force springs;\textsuperscript{2476,2477} the ferric wheel (Fig. 2.10);\textsuperscript{2532} molecular barrels\textsuperscript{2558} and molecular tennis balls;\textsuperscript{2551} various helical,\textsuperscript{2554-2556,2569,2605,2638} boxlike,\textsuperscript{2559} and other hollow organic\textsuperscript{2572} structures; molecular “turnstiles”\textsuperscript{2573} and “molecular tweezers”;\textsuperscript{2577} simple and complex trefoil knot molecules;\textsuperscript{2562} and stable carbonye* molecular rods with chains lengths up to ~300 carbon atoms having alternating single and triple covalent bonds.\textsuperscript{317} Chemists are justifiably awed by the creation of structures like the ferric wheel; “for me,” observed Nobel chemist Roald Hoffmann,\textsuperscript{125} “this molecule provides a spiritual high....The molecule is beautiful because its symmetry reaches directly into the soul, [playing] a note on a Platonic ideal.”

*In chemical nomenclature, hydrocarbons with single carbon-carbon bonds are “-anes,” double bonds make “-enes,” and triple bonds make “-ynes.”
gave numerous examples of molecular gear systems that he had synthesized using the methods of traditional solution chemistry, which "resemble to an astonishing degree the coupled rotations of macroscopic mechanical gears." He added: "It is possible to imagine a role for these and similar mechanical devices, molecules with tiny gears, motors, levers, etc., in the nanotechnology of the future."

In 1998, Gimzewski and coworkers synthesized and observed a 1.5-nm diameter gearlike single molecule (hexa-tert-butyl decacyclene) rotating within a supramolecular bearing (a void in a 2-D lattice) at room temperature on the surface of a Cu{100} lattice. The authors suggest that added non-thermal noise (e.g., temperature differences produced by tunnel current heating of the rotor) could be rectified by asymmetries in the rotor/neighbor potential energy curve, allowing the rotor to turn unidirectionally. Says Gimzewski: "Our wheel is frictionless in the conventional sense of the word, and it doesn't wear out." Gimzewski’s experimental work confirms Drexler’s calculations that van der Waals bearings can operate in molecular systems, that multi-atomic bearing surfaces with higher load capacities than sigma bond bearings can be built, that these bearings can run with no lubricants, and that such bearings have sufficiently low energy barriers to allow turning by thermal vibrations alone.

By 1998, limited control of the mechanical action of molecular parts had been demonstrated. For example, in 1994 T. Ross Kelly’s group created a “paddlewheel” molecule (a spinning propeller-shaped wheel) with a built-in chemically-controlled brake. In solution, the wheel spins freely, but when Hg++ ions are added, a blocking ligand physically rotates into a new position (tripping the brake) and stopping the rotation; removal of the mercury causes the wheel to resume spinning. In 1997, Kelly’s group used similar techniques to synthesize the first molecular ratchet, a propeller with three benzene-ring “blades” that serve as gear teeth. A row of four rings (the pawl) sits between two of the blades, such that the propeller cannot turn without pushing it aside; because of a twist in the pawl, it is easier to turn clockwise than counterclockwise under thermal agitation, providing mechanical rectification without net motion or energy extraction. Shinkai and colleagues fabricated a pair of molecular tongs for grasping metal ions that uses two crown ether rings as the jaws, linked by two benzene rings joined through two double-bonded nitrogen atoms (Fig. 2.11). Irradiation with ultraviolet light induces photoisomerization of the double bond between the nitrogen atoms, closing the jaws; heat reverses the process, opening the jaws. In related systems, collectively called “butterfly molecules,” irradiation with visible light or pH changes can also induce opening. Seeman’s chemically-operated mechanical DNA system provides similar control of molecular mechanical movement (Section 2.3.1). Numerous controllable molecular electronics devices and possible nanocomputer components are reviewed in Section 10.2.

Noncovalent self-assembly of molecular parts has also been demonstrated. For example, Fraser Stoddart and colleagues have extensively investigated the rotaxanes, molecules in which one part is threaded through a hole or loop in another, and kept from unthreading by end-groups, like a ring trapped on a barbell. In one system (Fig. 10.8), a ring-shaped molecule slides freely along a shaft-like chain molecule, moving back and forth between

---

**Fig. 2.8.** Other unusual molecular “parts.”

**Fig. 2.9.** Calixarene.

**Fig. 2.10.** Ferric wheel (Fe, O, C atoms shown; H, Cl atoms not shown).

**Fig. 2.11.** “Butterfly Molecule” used as molecular tongs.
stations at either end with a frequency of ~500 Hz, making an oscillating "molecular shuttle."\textsuperscript{2483} This shuttling behavior can be controlled by a series of different chemical, electrochemical, or photochemical external stimuli.\textsuperscript{2335,2354} Using a different approach, G. Wenz\textsuperscript{2487} threaded ~120 molecular beads onto a single, long, polyimino oligomethylene polymer chain, making a "molecular necklace." Other groups are pursuing related research.\textsuperscript{2322,2329,2350} regarding \[n\]-rotaxanes\textsuperscript{2538,2539} and pseudorotaxanes (mechanically-threaded molecules),\textsuperscript{2541,2542} while still others are using the necklace technique as a means, for example, of self-assembly - 2 nm wide cycloextrin nanotubes.\textsuperscript{2488}

Another interesting molecular system that can be made entirely by self-assembly is the catenanes,\textsuperscript{2485} which have two or more closed rings joined like the links of a chain. The rings are mechanically linked—there are no covalent bonds between separate links. The first [2]-catenane (2 linked rings) was constructed by Edel Wasserman of AT\&T Bell Laboratories in 1960, using simple hydrocarbon rings; the first [3]-catenane was synthesized in 1977. In 1994, Fraser Stoddart and David Ambibiano constructed the first [5]-catenane,\textsuperscript{2485} having ~372 atoms, dubbed "olympiadane" because of its resemblance to the Olympic rings, and in 1997 the same group announced the first deterministically-ordered 7-ring heptacatenane,\textsuperscript{2535} another group later reported the spontaneous self-assembly of a 10-component catenane.\textsuperscript{3280} By 1998, much research on larger polymeric chains of linked rings (e.g., oligocatenanes,\textsuperscript{2532}) supramolecular "daisy chains",\textsuperscript{2533,2534} fullerene-containing catenanes,\textsuperscript{2536} and supramolecular weaving\textsuperscript{2537} was in progress. The smallest [2]-catenane constructed to date has dimensions 0.4 nm x 0.6 nm,\textsuperscript{2528} self-assembled mechanically-interlocked 2-dimensional\textsuperscript{2595,2597} and 3-dimensional\textsuperscript{2598,2599} "infinite" arrays of single molecular ringlike species are known. Interestingly, in 1998 it was discovered that many viral coats (which also self-assemble) appear to be a 2-dimensional "hairpin" composed of mechanically-interlocked protein rings—"hairs" of the bacteriophage HK97 capsid shell consist of exactly 72 interlinked protein rings, specifically 60 hexamers and 12 pentamers.\textsuperscript{2486}

Complex molecular parts also can be built up from simpler molecular parts by covalent bonding as low-symmetry shape-invariant molecular object polymers,\textsuperscript{2602} or via processes variously known as modular chemistry,\textsuperscript{2498,2499} heterosupramolecular chemistry,\textsuperscript{3220} chemical assembly,\textsuperscript{2541} or structure-directed synthesis\textsuperscript{2489,2490}—including the well-known Diels-Alder reactions\textsuperscript{2557} (e.g., see Fig. 2.20F) which may be employed repetitively to fabricate a series of atomically-precise molecular rods, rings, and nanocages (e.g., "beltenas" and "collarenes")\textsuperscript{2489} composed of porphycene units with nanometer-scale dimensions. Such efforts have been described as steps toward "molecular LEGO,"\textsuperscript{2489} "molecular Meccano,"\textsuperscript{2480} "molecular Tinkertoys"\textsuperscript{2509,2516,2842} or "molecular building blocks."\textsuperscript{2850} Some structure-directed assembly has been achieved with simple DNA forms.\textsuperscript{1916} Other simple techniques for synthesizing nanowires,\textsuperscript{643} nanorods,\textsuperscript{2672,2716} nanotubes,\textsuperscript{1411,2673,2674,2862} and nanocages,\textsuperscript{2640,2675,2862} or for selecting polymers for length,\textsuperscript{2686} are well-known.

For instance, Jean-Marie Lehn\textsuperscript{2679} has created rectangular grid complexes by mixing rodlike ligands that make use of different numbers of binding sites—one example is a 4 x 5 array consisting of nine ligands and 20 metal ions, nanoscale grids which he claims might one day find use "as components within a futuristic information storage and processing nanotechnology". C. M. Drain\textsuperscript{2680,2681} has described a 21-component (5 nm)\textsuperscript{2} square array of nine porphyrins tethered together by 12 palladium ions shaped like a four-pane window. The structure is built from three different kinds of porphyrins—(1) an X-shaped unit that coordinates to four metal ions, forming the center of the array, (2) a T-shaped unit that coordinates to three metals and forms each side of the array, and (3) an L-shaped unit that coordinates to only two metals and forms each corner of the array. When these components are placed in solution in the correct ratio, they form the square array within a half hour at room temperature with about 90% yield. The same porphyrin units can be combined in different ratios and induced to form wires or tapes. The properties of such arrays can be fine-tuned by choosing the appropriate metal ion linker and functionalized porphyrin unit.\textsuperscript{2679}

Jeffrey Moore has investigated a three-dimensional nanoscaffolding comprised of a molecular lattice that could serve as a framework for catalysts, photosynthetic molecules, or more complex molecular devices. Moore's original objective was to synthesize modular building blocks with physical and chemical properties that would dictate a "programmed assembly" protocol for "nanoarchitectures."\textsuperscript{2574-2578} The characteristics of each modular unit shaped the weak intermolecular attractions—electrostatic, van der Waals, and hydrogen-bonding forces—between it and its neighbors so that units would array themselves in a larger structure with the desired spacing and geometry. Moore's group first linked phenylacetylene subunits into oligomers of various lengths which bend and twist like wire sculptures until their ends meet to form closed polyhedral molecules—the basic scaffold components (of which there were at least half a dozen shapes). These building blocks then self-assembled or folded\textsuperscript{2555} into geometries at least partly controlled by designed hydrogen-bonding interactions. The molecular polyhedra each possessed several phenylene groups, "chemical handles" upon which other molecular groups could be hung. Moore then focused on ~10-nm dendrimeric forms\textsuperscript{2580-2585} and later began attaching photosensitive antenna molecules to these structures.\textsuperscript{2586,2591}

By 1997, Moore had begun investigating norbornadiene on Si[100] in an effort to develop ways to covalently attach molecules onto surfaces with sub-nanometer spatial precision,\textsuperscript{2592,2593} and had suggested a packing model for interpenetrated diamondoid structures.\textsuperscript{2594} Robson's group\textsuperscript{2846} has also investigated possible scaffolding molecules.

Josef Michl is attempting to build a molecular construction kit using molecular rods and connectors, pursuing an explicit vision of "molecular Tinkertoys"\textsuperscript{2509-2516} by working with simple molecular structures that form stiff, flexible rods. Michl has assembled rods from a mixture of carbon-boron molecules (e.g., 10-vertex or 12-vertex carboranes)\textsuperscript{2495} and carbon-hydrogen molecules, providing fine control over the total rod length. The rods are built up from more primitive molecular parts, such as propellane (a strained form of \textit{C}_5\textit{H}_4\textit{C}) and cubane (a strained form of \textit{C}_8\textit{H}_8), to make "staffanes"\textsuperscript{2500-2510} which are a series of cage units in a linear series (Fig. 2.12). (Strained molecules are constructed with bonds that are forced out of their normal angles, as for example 90° in the case of cubane, compared with carbon's normal orientation of 109.5°.) Michl has fabricated rods whose lengths vary from 0.5 nm to 2.5 nm, in precise 0.1-nm steps. There are other ways of making rodlike molecules, but Michl's are highly inert, do not absorb visible or UV light, are stable up to at least 200°C, and do not react with the oxygen in air even at high temperatures.\textsuperscript{2496}

Related work is underway in Michl's laboratory on connectors to join the rods together.\textsuperscript{2496} Metal atoms would be the simplest
solution, offering the useful quality of strong joints that can be easily disassembled. Different metals give different binding geometries—square, octahedral, and so on. Many of the connectors are metal-containing species; some are symmetrically trisubstituted or hexa substituted benzenes or tetrasubstituted cyclobutadiene complexes. Michl must build the right chemical groups onto the rod termini to make them adhere to the connectors as desired. In one example, tiny crosses were built by attaching carboxylate groups to one end of the staffanes, allowing them to bind to a connector made of two rhodium atoms. The groups at the other end of the rods were converted to an ester group, which doesn’t bind to rhodium, and so the crosses self-assembled correctly. Michl has also fabricated “star connectors” in which the rods are built into the molecules as covalently bonded arms of a star. For instance, one 3-arm star connector uses three large carboranes coupled to a central benzene ring, allowing a pedestal to be attached vertically to the benzene ring via a ruthenium “sandwich” bond. Additional devices could be installed on this molecular scaffolding, such as optically active molecules or even molecular mechanical “windmills.” Extensive ab initio quantum mechanical calculations have been fabricated. Extensive ab initio quantum mechanical calculations accompany the experimental work and aid in the interpretation of the results. Michl’s ultimate goal is “the production of thin layers of solids of completely controlled aperiodic structure consisting of an inert covalent scaffolding carrying selected active groups.”

G. Leach and colleagues used computers to simulate various diamond and graphite nanostruts using molecular dynamics calculations, analyzing solid rectangular struts with varying aspect ratios, cross-sections, terminating atoms (for diamond), potential energy functions, and temperatures. The most dramatic differences were seen between struts with 100:1 aspect ratios and struts with lower (10:1 or 1:1) aspect ratios—the diamond strut with a 100:1 aspect ratio and a unit cell (0.4 nm)2 cross-section begins to curl (with end-to-end distance decreasing 1.2 nm after 20 picosec) due to thermal vibration (both at 150 K and 300 K), while a unit strut with an aspect ratio of 10:1 or a 100:1 strut with a (1 nm)2 cross-section showed end-to-end length fluctuations of only ~0.1 nm. The authors concluded that support struts with a cross section of at least (1 nm)2 would be sufficient to keep both end-to-end distance and transverse fluctuations below ~0.1 nm in struts with 10:1 aspect ratios. Other work by the same group claims that many nanomachine designs may actually perform better than molecular dynamics results might suggest, or may require fewer atoms for the same positional stability.

The first fullerene to be discovered, the spherical C60 molecule, was originally named “buckminsterfullerene” by its discoverers, due to the molecule’s similarity to the geodesic domes designed by the famous American architect and engineer, Buckminster Fuller (1895-1983). The name has since been shortened by common usage to “fullerenes,” although spherical or tubular fullerenes are sometimes informally called “buckyballs” and “buckytubes,” again reflecting the original provenance.
larger (Fig. 2.15) variants have been observed; as cages get bigger, the corners (where the 12 pentagons needed for closure reside) get sharper. The topologically smallest possible fullerene is a dodecahedron consisting of 12 pentagons and 20 carbon atoms, but the ring isomer of C_{20} (a hoop of single carbons) is apparently energetically favored over the bowl or spherical (fullerene) isomers; in 1998, C_{28} was the smallest fullerene that had been observed experimentally, and C_{36} could be made in significant quantities by arc-discharge.

The approximate diameter (measured to atomic centers) of a spheroidal fullerene of formula C_n is D_{ball} \approx 1.1(n/60)^{1/2} nanometers. In 1998, purified C_{60} cost $27.50/gm (99.5% pure) or $60/gm (99.9%), C_{70} $250/gm (98%), and C_{84} $3,750/gm (90%) from Dynamic Enterprises Ltd. of London.

Insertion of many additional equatorial belts of carbon atoms (incorporated as hexagons) into a spheroidal fullerene results in a long cylinder with graphite-like (“graphene”) walls and spherical endcaps, making a class of fullerenes known as the capsular fullerenes or single-walled carbon nanotubes (SWNT) that come in many sizes and chiral forms (Fig. 2.16). The molecular form of carbon nanotubes has been likened to rolled chicken wire. SWNTs may be synthesized by vaporizing graphite spiked with 1% catalyst from the nickel, cobalt and iron group above 3000˚C, then allowing the vapor to slowly condense. Tubes form because the metal atoms interact with dangling bonds at the end of a tube, favoring tube extension over hemispherical capping by newly arriving carbon vapor atoms. (For any given mass number, ball-shaped fullerenes are energetically favored over tube-shaped fullerenes during high-temperature synthesis in the absence of catalyst.) SWNTs typically self-assemble in 1.1 nm wide tubes, although smaller single-walled nanotubes and larger single-walled and multi-walled nested nanotubes exist. SWNT carbon atoms are bonded in virtually flawless hexagonal arrays. Simulations show that isolated flaws migrate to the ends of the tube and are eliminated, a phenomenon known as “self-healing.” Single-molecule nanotubes of C_{1,000,000} or larger, with lengths >100,000 times longer than their widths (e.g., ~1 nm long), had been synthesized by 1998. While graphite is a very brittle material, the graphene walls of a carbon nanotube are quite resilient. Computer simulations and experiments confirm that nanotubes kink when bent (Fig. 2.17), then snap back when released. Nanotube diameter is D_{tube} = 0.078(n^2 + nm + m^2)^{1/2} nanometers, where (n, m) is the “rollup vector” defined by the number of steps required for a repeat pattern along two crystallographic cylinder wall axes. The cohesive energy per atom required to curve a flat graphene sheet into a cylinder is E_{rollup} = (13 zJ-nm^2) / D_{tube}^2. (1 zeptojoule (zJ) = 10^{-21} joules.) As an example, a (10,10) nanotube has D_{tube} = 1.36 nm and E_{rollup} = 7 zJ/atom. Young’s modulus for individual SWNTs, assuming a hollow (open-ring) cross-section, has been estimated as high as ~5.5 x 10^{12} N/m^2; elastic bending modulus (measured as 1 x 10^{12} to 0.1 x 10^{12} N/m^2) decreases sharply with increasing diameter (from 8-40 nm).

From spatial geometry, it is known that hexagonal tiling produces flat sections (like cylinder walls), and inserting pentagons into an hexagonal array produces positive curvature (like spheres), making endcaps. But 7-sided heptagons can also be inserted into hexagonal arrays to induce negative curvature, thus permitting concave surface deformations such as saddle-shaped fullerenes (Fig. 2.18) or possibly helical tubular structures. Indeed, note Colbert and Smalley, “the use of hexagonal, pentagonal, and heptagonal substructures are sufficient to produce caged carbon structures of any topology.” Nature makes use of this same geometrical principle in the radiolarians, tiny protozoans with fullerene-like siliceous skeletons, and in the conical nucleoprotein core particle of the HIV-1...
virus (Fig. 2.19). R. Smalley observes that a graphene sheet has the highest tensile strength of any known 2-dimensional network, and the packing density of atoms in the sheet (e.g., atoms/m²) is higher than any other network made of any atoms in the periodic table and higher than the packing density in any 2-dimensional slice through any 3-dimensional object—even diamond which has the highest known 3-D packing density. Hence the graphene sheet is effectively impermeable under normal chemical conditions.

Carbon nanotubes have also been observed with a variety of end-cap shapes and with tubes that reduced or increased their diameter for some distance before terminating. Toroidal fullerenes can be produced experimentally; cone-like, spindle-like, and helical fullerene objects have also been examined. By 1998, the mechanical properties of fullerenes and carbon nanotubes were being extensively investigated and there were experimental demonstrations of:

1. a limited ability to cut fullerene nanotubes to specific desired lengths;
2. dissolving derivatized SWNTs in common organic solvents;
3. the synthesis of self-oriented nanotube arrays and
4. the trapping of individual C₆₀ molecules in a perforated Langmuir-Blodgett film "workpiece holder" (though an even simpler means for producing 2-nm hole arrays with 7-8 nm hole spacings was later reported).

Nanotubes were also being investigated as nanoscale sensor components.

Given an ability to synthesize a wide variety of fullerene shapes and sizes, the next task is to find ways to join them together in specific desired molecular architectures. Two hydrogens were readily added to fullerene carbon atoms, making specific isomers of C₆₀H₂ and C₇₀H₂, and in the 1990s fullerene chemistry began to be explored. A little more than a decade after the discovery of C₆₀, gram quantities of buckyballs became widely available for chemical experimentation. Progress in covalent fullerene chemistry exploded. By the late 1990s, fullerene surfaces could be regioselectively functionalized in many interesting ways, including (Fig. 2.20):

A. a fullerene dimer,
B. a fullerene-polyester polymer,
C. a fullerene dendrimer,
D. a fullerene-rotaxane and catenane,
E. a fullerene-nucleotide DNA cleaving agent,
F. a Diels-Alder fullerene adduct that is very stable against cycloreversion, and
G. 3-dimensional extended polymeric multifullerene forms such as acetylenic macro-rings and DNA/fullerene hybrid materials.

*Interestingly, conical hexagonal lattices with narrow endcaps that are closed using P pentagons are required by Euler's theorem to have a quantized cone angle θ defined by sin(θ/2) = 1 - (P/6) for P = 0, 1, ..., 6; minimum cone angle is 19.2° at P = 5.*
Fig. 2.20B. Fullerene polyester polymer. 2621

Fig. 2.20C. Fullerene dendrimer. 2620

Fig. 2.20D. Fullerene rotaxane. 2622

Fig. 2.20E. Fullerene-nucleotide DNA cleaving agent. 2623

Fig. 2.20F. Stable Diels-Alder fullerene adduct. 2624

Fig. 2.20G. Extended fullerene polymers. 2624
Theory\textsuperscript{2665} and experiment\textsuperscript{2666} suggest that SWNTs may be joined at 30° angles to create complex structures, including helices and three-way nanotube junctions,\textsuperscript{1308,2614} as suggested by Figure 2.18.

In 1998, the idea of using fullerenes as “Tinkertoys” for molecular construction was just starting to be considered.\textsuperscript{2643} For example, given the synthesis of C\textsubscript{60} exohedral monoadducts and multiple adducts described above, it may be possible to fabricate simple gears by bonding rigid ligands onto the external surfaces of carbon nanotubes in the manner of gear teeth,\textsuperscript{2644-2646} although the chemical synthesis of nanotubes with precisely positioned teeth will not be easy.\textsuperscript{2607} Given the success in 1996 of IBM scientists\textsuperscript{2647} in positioning individual organic molecules (each having a total of 173 atoms and a 1.5-nm diameter) at room temperature by purely mechanical means, it might also be possible to align and maintain these molecular gear teeth in atomically precise meshed positions. If this could be done, would such devices actually work in the manner of macroscale gears?

J. Han and colleagues at NASA/Ames\textsuperscript{2648} performed a detailed 2000-atom molecular dynamics simulation to investigate the properties of molecular gears fashioned from carbon nanotubes with teeth added via a benzyne reaction known to occur with C\textsubscript{60}. Computationally, one gear is powered by forcing the atoms near the end of the nanotube to rotate (Fig. 2.21), and a second gear is allowed to rotate by keeping the atoms near the end of its nanotube constrained to a cylinder (i.e., the ends of the shaft were constrained to not elongate but were allowed to move within a plane transverse to the tube symmetry axis). The meshing aromatic gear teeth transfer angular momentum from the powered gear to the driven gear. Each gear is made of a 1.1-nm diameter (14, 0) nanotube with seven benzyne teeth. The spacing between two nanotubes is 1.8 nm and the smallest distance between a tooth atom and a tube atom is ~0.4 nm. The results show that the gears can operate up to 70 GHz in vacuo at room temperature without overheating or slipping. As rotational speed rises above 150 GHz, the gears overheat and begin to slip with tooth tilting up to 20°, but no bond or tooth breaking occurs up to at least 3000 K and slipping gears can always be returned to proper operation by lowering the temperature or the rotation rate. A related nanotube gear system (5-8 sprockets, 290-464 atoms) simulated by Robertson and colleagues\textsuperscript{2670} at the Naval Research Laboratory (NRL) showed similar overheating at 500 GHz but was stable when accelerated to only 20 GHz.

The NASA group also simulated several other types of nanotube-based gear systems. For instance, a rack and pinion system (Fig. 2.22) was designed using a gear made from a (14, 0) nanotube with teeth separated by two hexagon rings and a shaft made from a (9, 9) tube with teeth separated by three rings. Gear and shaft are 1.94 nm apart, with tooth face normal to the radial direction of nanotube (14, 0) for the gear, but in the axis direction of nanotube (9, 9) for the shaft. The gear could receive power, driving the shaft, or vice versa, and worked well for shaft translational velocities up to ~100 m/sec. (Most devices described in Drexler\textsuperscript{10} typically move at ≤1 m/sec.) Since shaft mass is almost twice gear mass, it takes more power for the gear to drive the shaft. In another simulation involving a large 1.4-nm gear coupled to a smaller 0.8-nm gear (Fig. 2.23), the large gear drives the smaller gear smoothly, but if power is instead applied to the smaller gear at a sufficiently large acceleration, then the smaller gear does not drive the larger one but instead “bounces back and forth several times, like elastic collisions of a small ball between two boards.”

Another research group at Oak Ridge National Laboratory (ORNL) used classical molecular dynamics to investigate the properties of molecular bearings consisting of an inner and an outer carbon nanotube.\textsuperscript{2662,2663} The graphite bearings ranged in size from inner shafts between 0.4-1.6 nm in diameter up to 12 nm long, and outer cylinders between 1.0-2.3 nm in diameter up to 4 nm long. The original simulations\textsuperscript{2662} found excessive vibrational motion, but subsequent work using a more complete quantum approach\textsuperscript{2663,2664} found that under certain conditions the nanobearing is “frictionless” and undergoes superrotation, a classical dynamical behavior reminiscent of superfluidity. The regime of superrotatory motion is somewhat restricted when the nanobearing is under load, which suggests that a very careful design is required to ensure optimum performance.

The ORNL group simulated a fullerene motor\textsuperscript{2668} consisting of two concentric graphite cylinders (shaft and sleeve) with one positive and one negative electric charge attached to the shaft. Rotational motion of the shaft was induced by applying one, or sometimes two, oscillating laser fields. The shaft cycled between periods of...
undesirable rotational pendulum-like behavior and good unidirectional motor-like behavior. The NASA/Ames group simulated a pulsed-laser-powered carbon nanotube gear-motor system which rotated consistently in one direction, although that direction could be either clockwise or counterclockwise.

G. Leach and colleagues also simulated various carbon nanotubes treated as nanostruts using molecular dynamics calculations, as described earlier for diamondoid block struts. The conclusions were similar though less severe. Nanotubes with 100:1 aspect ratio and 0.8-nm diameter at 300 K experienced periodic oscillations in length of ~0.4 nm at ~122 GHz, although multiple modes clearly were being excited; nanotubes of lesser aspect varied by at most ~0.04 nm.

2.3.3 Scanning Probe Technology

The third general pathway leading to molecular manufacturing involves a technology known as scanning probe microscopes (SPMs). The first of the SPMs was the Scanning Tunneling Microscope (STM) developed in the late 1970s and early 1980s by Gerd Karl Binnig and Heinrich Rohrer at an IBM research lab in Zurich, Switzerland, earning these scientists, along with Ernst Ruska, the 1986 Nobel in Physics. The STM was initially used as an imaging device, capable of resolving individual atoms by recording the quantum tunneling current that occurs when an extremely sharp conductive probe tip (usually tungsten, nickel, gold, or PtIr) is brought to within about one atomic diameter of an atom, and then adjusting the position of the tip to maintain a constant current as the tip is scanned over a bumpy atomic surface (Fig. 2.24).

A height change as small as 0.1 nm can cause tunneling current to double. The tip is connected to an arm that is moved in three dimensions by stiff ceramic piezoelectric transducers that provide sub-nanometer positional control. If the tip is atomically sharp, then the tunneling current is effectively confined to a region within ~0.1 nm of the point on the surface directly beneath the tip, thus the record of tip adjustments generates an atomic-scale topographic map of the surface. STM tips can scan samples at ~kHz frequencies, although slower scans are used for very rough surfaces. In some modern STMs (e.g., the DI Nanoscope), the sample is moved while the tip is held stationary.

A major limitation of the STM was that it only worked with conducting materials such as metals or semiconductors, but not with insulators or biological structures such as DNA. To remedy this situation, in 1986 Binnig, Quate and Gerber developed the Atomic Force Microscope (AFM) which is sensitive directly to the forces between the tip and the sample, rather than a tunneling current. An AFM can operate in at least three modes. In “attractive” or non-contact mode (NC-AFM, 0.01-1 N/m force constant), the tip is held some tens of nanometers above the sample surface where it experiences the attractive combination of van der Waals, electrostatic, or magnetostatic forces. In “repulsive” or contact mode (C-AFM, 0.01-1 N/m force constant), the tip is pressed close enough to the surface for tip and sample electron clouds to overlap, generating a repulsive electrostatic force of ~10 nN, much like the stylus riding a groove in a record player. There is also intermittent-contact mode (IC-AFM, 0.01-1 N/m force constant), which is sometimes called “tapping” mode. In any of these modes, a topographic map of the surface is generated by recording the up-and-down motions of the cantilever arm as the tip is scanned. These motions may be measured either by the deflection of a light spot reflected from a mirrored surface on the cantilever or by tiny changes in voltage generated by piezoelectric transducers attached to the moving cantilever arm.
Typical AFM cantilevers have lengths of 100–400 microns, widths of 20–50 microns, and thicknesses between 0.4 to several microns. AFM tips may be positioned with ~0.01 nm precision, compressive loads as small as 1-10 pN are routinely measured, and the tips may be operated even in liquids.

S. Vetter notes that STM technology has also improved, reaching resolutions of ~0.001 nm in the z direction (vertical) and ~0.01 nm in the xy plane, well beyond atomic resolution. The STM remains the instrument with the best resolution. The conducting-surface limitation has been overcome in some cases by coating the target with an extremely thin conducting layer, developing tips with multiple electrodes, and coating a conducting substrate with a sample so thin as to allow enough conduction even if the sample is characterized as an insulator in bulk.

By 1998, the growing family of SPMs included at least forty types of instruments and techniques that relied on interactions between a scanned surface and a nearby probe. Different instruments measured different forces and thus could be used to characterize different properties of the surface. For example, friction force microscopes (FFMs), magnetic force microscopes (MFMs), shear force microscopes (ShFFMs), scanning capacitance microscopes (SCMs), scanning conducting ion microscopes, chemical force microscopes, and electrostatic force microscopes (EFMs) measured frictional drag or other binding forces. Magnetic resonance force microscopes (MRFMs) used a field generated from a small magnet mounted on the tip of the cantilever arm to probe nuclear magnetic moments in a small region on the surface of the sample, imaging atom types and even detecting the spin of a single nuclear magnetic moments in a small region on the surface of the sample. In 1988, J. Foster and colleagues at IBM Almaden pinned small organic molecules to a graphite surface by applying a small electrical pulse through an STM. Additional pulses enlarged or erased this molecular feature, and often split the organic molecule into smaller pieces, though not in a controlled manner. Other workers used an STM to dump small clusters of gold atoms on a platinum surface, clusters of copper atoms on a gold surface, and molecules of tungsten carbide from a supply flowing over a surface. Others reported positioning carbon monoxide molecules and platinum atoms on platinum surfaces, moving clusters and single atoms of silicon across a silicon surface at room temperature, and fabricating via STM a “molecular corral” or “quantum corral” — a ring of atoms so small that the enclosed electrons were forced to exhibit quantum behavior. In 1999, Tomaneck and Kral proposed an atomic fountain pen in which a carbon nanotube filled with atoms is electrically induced to release these atoms, one by one every 15 microseconds, onto the work surface.

More precise control was achieved by Eigler and Schweizer at IBM Almaden in 1989, when they used an STM to position 35 individual xenon atoms on a nickel surface to spell out the corporate logo “IBM” (Fig. 2.25). To accomplish this, a bias voltage was applied to weaken the adsorption of each atom to its nickel substrate, then each atom was dragged, one by one, to the desired locations at a speed of ~0.4 nm/sec to build a meaningful pattern. It took 22 hours to make the entire logo, or ~38 min/atom. The experiment had to be carried out at 4 K (i.e., liquid helium temperatures) because the arrangement would have been unstable at room temperature. Numerous similar examples of “atomic graffiti” followed the IBM experiment, including atoms patterned in the shape of a world map, the word “atom” and “nanoworld,” spelled out with atoms, in Japanese; the word “Peace,” Einstein’s famous equation from a hydrogen monolayer atop a silicon base; the exposed silicon surface oxidized, producing narrow SiO2 tracks.

Perhaps more significantly, SPMs have been employed in an increasingly sophisticated manner to manipulate individual atoms and small clusters of atoms. The possibility of modifying surfaces with scanning probe tips was evident from the earliest years of STM research, since inadvertent contact between tips and surfaces routinely caused such modifications, and the possibility of transferring material between tip and sample had already been discussed. Suggestions for controlled surface modification soon appeared, and in 1985 Becker and Golovchenko used voltage pulses on an STM tip to pluck a single germanium atom from the surface of a sample. In 1988, J. Foster and colleagues at IBM Almaden pinned small organic molecules to a graphite surface by applying a small electrical pulse through an STM. Additional pulses enlarged or erased this molecular feature, and often split the organic molecule into smaller pieces, though not in a controlled manner. Other workers used an STM to dump small clusters of gold atoms on a platinum surface, clusters of copper atoms on a gold surface, and molecules of tungsten carbide from a supply flowing over a surface. Others reported positioning carbon monoxide molecules and platinum atoms on platinum surfaces, moving clusters and single atoms of silicon across a silicon surface at room temperature, and fabricating via STM a “molecular corral” or “quantum corral” — a ring of atoms so small that the enclosed electrons were forced to exhibit quantum behavior. In 1999, Tomaneck and Kral proposed an atomic fountain pen in which a carbon nanotube filled with atoms is electrically induced to release these atoms, one by one every 15 microseconds, onto the work surface.

More precise control was achieved by Eigler and Schweizer at IBM Almaden in 1989, when they used an STM to position 35 individual xenon atoms on a nickel surface to spell out the corporate logo “IBM” (Fig. 2.25). To accomplish this, a bias voltage was applied to weaken the adsorption of each atom to its nickel substrate, then each atom was dragged, one by one, to the desired locations at a speed of ~0.4 nm/sec to build a meaningful pattern. It took 22 hours to make the entire logo, or ~38 min/atom. The experiment had to be carried out at 4 K (i.e., liquid helium temperatures) because the arrangement would have been unstable at room temperature. Numerous similar examples of “atomic graffiti” followed the IBM experiment, including atoms patterned in the shape of a world map, the word “atom” and “nanoworld,” spelled out with atoms, in Japanese; the word “Peace,” Einstein’s famous equation from a hydrogen monolayer atop a silicon base; the exposed silicon surface oxidized, producing narrow SiO2 tracks.

Perhaps more significantly, SPMs have been employed in an increasingly sophisticated manner to manipulate individual atoms and small clusters of atoms. The possibility of modifying surfaces with scanning probe tips was evident from the earliest years of STM research, since inadvertent contact between tips and surfaces routinely caused such modifications, and the possibility of transferring material between tip and sample had already been discussed. Suggestions for controlled surface modification soon appeared, and in 1985 Becker and Golovchenko used voltage pulses on an STM tip to pluck a single germanium atom from the surface of a sample. In 1988, J. Foster and colleagues at IBM Almaden pinned small organic molecules to a graphite surface by applying a small electrical pulse through an STM. Additional pulses enlarged or erased this molecular feature, and often split the organic molecule into smaller pieces, though not in a controlled manner. Other workers used an STM to dump small clusters of gold atoms on a platinum surface, clusters of copper atoms on a gold surface, and molecules of tungsten carbide from a supply flowing over a surface. Others reported positioning carbon monoxide molecules and platinum atoms on platinum surfaces, moving clusters and single atoms of silicon across a silicon surface at room temperature, and fabricating via STM a “molecular corral” or “quantum corral” — a ring of atoms so small that the enclosed electrons were forced to exhibit quantum behavior. In 1999, Tomaneck and Kral proposed an atomic fountain pen in which a carbon nanotube filled with atoms is electrically induced to release these atoms, one by one every 15 microseconds, onto the work surface.

More precise control was achieved by Eigler and Schweizer at IBM Almaden in 1989, when they used an STM to position 35 individual xenon atoms on a nickel surface to spell out the corporate logo “IBM” (Fig. 2.25). To accomplish this, a bias voltage was applied to weaken the adsorption of each atom to its nickel substrate, then each atom was dragged, one by one, to the desired locations at a speed of ~0.4 nm/sec to build a meaningful pattern. It took 22 hours to make the entire logo, or ~38 min/atom. The experiment had to be carried out at 4 K (i.e., liquid helium temperatures) because the arrangement would have been unstable at room temperature. Numerous similar examples of “atomic graffiti” followed the IBM experiment, including atoms patterned in the shape of a world map, the word “atom” and “nanoworld,” spelled out with atoms, in Japanese; the word “Peace,” Einstein’s famous equation from a hydrogen monolayer atop a silicon base; the exposed silicon surface oxidized, producing narrow SiO2 tracks.

Perhaps more significantly, SPMs have been employed in an increasingly sophisticated manner to manipulate individual atoms and small clusters of atoms. The possibility of modifying surfaces with scanning probe tips was evident from the earliest years of STM research, since inadvertent contact between tips and surfaces routinely caused such modifications, and the possibility of transferring material between tip and sample had already been discussed. Suggestions for controlled surface modification soon appeared, and in 1985 Becker and Golovchenko used voltage pulses on an STM tip to pluck a single germanium atom from the surface of a sample. In 1988, J. Foster and colleagues at IBM Almaden pinned small organic molecules to a graphite surface by applying a small electrical pulse through an STM. Additional pulses enlarged or erased this molecular feature, and often split the organic molecule into smaller pieces, though not in a controlled manner. Other workers used an STM to dump small clusters of gold atoms on a platinum surface, clusters of copper atoms on a gold surface, and molecules of tungsten carbide from a supply flowing over a surface. Others reported positioning carbon monoxide molecules and platinum atoms on platinum surfaces, moving clusters and single atoms of silicon across a silicon surface at room temperature, and fabricating via STM a “molecular corral” or “quantum corral” — a ring of atoms so small that the enclosed electrons were forced to exhibit quantum behavior. In 1999, Tomaneck and Kral proposed an atomic fountain pen in which a carbon nanotube filled with atoms is electrically induced to release these atoms, one by one every 15 microseconds, onto the work surface.

More precise control was achieved by Eigler and Schweizer at IBM Almaden in 1989, when they used an STM to position 35 individual xenon atoms on a nickel surface to spell out the corporate logo “IBM” (Fig. 2.25). To accomplish this, a bias voltage was applied to weaken the adsorption of each atom to its nickel substrate, then each atom was dragged, one by one, to the desired locations at a speed of ~0.4 nm/sec to build a meaningful pattern. It took 22 hours to make the entire logo, or ~38 min/atom. The experiment had to be carried out at 4 K (i.e., liquid helium temperatures) because the arrangement would have been unstable at room temperature. Numerous similar examples of “atomic graffiti” followed the IBM experiment, including atoms patterned in the shape of a world map, the word “atom” and “nanoworld,” spelled out with atoms, in Japanese; the word “Peace,” Einstein’s famous equation from a hydrogen monolayer atop a silicon base; the exposed silicon surface oxidized, producing narrow SiO2 tracks.
“E=mc²,” and even a reproduction of the well-known Einstein portrait with the scientist’s tongue sticking out, a sketch of a “molecule man” using 28 individual CO molecules on a platinum surface; gold nanoparticles spelling out “USC” and “Zyvex”; and sulfur atoms removed from a surface to make letters 2 nm high. In 1996, Gimzewski’s group used an STM tip to manipulate individual buckyballs along terraces on a grooved copper plate, making a “molecular abacus.”

Some variant of an SPM would appear to be the most promising tool for direct molecular manipulation. Once a tip lies within ~1 nm of a surface, the potential barrier can be lowered sufficiently for atoms to be induced from the surface by field evaporation; an applied tip voltage ionizes the atom which can then be guided around by the tip. Depending on voltage and separation, atoms, atom clusters, or molecules can be pickled out, pushed in, or nudged around on a surface. By 1998, techniques for direct atomic manipulation proceeded at room temperature and ~1000 times faster than the original IBM logo had been assembled. For example, the Nanomanipulator, an interactive haptic control system created by a group at the University of North Carolina at Chapel Hill, allows near-real-time manipulation of individual gold atoms across a surface using a hand-held master-slave controller that drives an STM probe while the position of the atom is displayed on a monitoring screen visible to the user. Partially funded by NSF, Silicon Graphics, and TopoMetrix, the Nanomanipulator is a fully-integrated system that enables investigators to “feel” the interatomic forces as the user pushes atoms around on a surface. In one demonstration, a TV-watching user herded a gold atom, slipping and sliding, into a slot in the planar workplace in about 1 minute; when force feedback was turned off, movements tended to run wild. IBM Almaden has also experimented with removing individual atoms from metal surfaces using an STM hooked into a Virtual Reality Dataglove apparatus.

To manufacture atomically precise parts, it also will be necessary to manipulate covalent bonds at the probe tip. STMs have broken and created chemical bonds; ~1 volt pulses were used to pull atoms out of crystals, binding them to the tip, and then to re-insert them back into the crystal. In 1995, the first demonstration of catalysis on a nanometer scale was reported by scientists at the Molecular Design Institute at LBL. They used an AFM modified to function like an ultrafine-point pen for catalytic calligraphy to change the chemical composition of a material surface one molecule at a time. A surface was prepared as a self-assembled monolayer (SAM) of allylaldehyde molecules capped with a crown of three nitrogen atoms, then platinum-coated chromium was deposited onto an AFM silicon tip just a few atoms wide. The SAM was soaked with a hydrogen-containing solvent, then scanned by the AFM over a 100 micron area, with the platinum catalyzing a covalent bonding reaction in which hydrogen was added to the aldehydes, transforming them into amines as revealed by selective fluorescent tags.

Tip technology is an important and fast-growing subspecialty of SPM research. In 1990, Drexler and Foster suggested the use of custom-made, synthetic proteins to be mounted on the point of an SPM tip. Modified antibodies or specially designed proteins could serve as simple, first-generation “grippers” for binding and manipulating specific molecules, bringing them into position to react with other molecules in a precise and selective way. Functionalized SPM tips have been created to exploit antibody-antigen recognition and in the context of chemical force microscopy. A major convenience for molecular manipulation is that the same instrument that does the chemistry and the molecular orienting can also be used to inspect the results.

In 1996, Smalley’s group at Rice University, attached a single nanotube to the pyramidal silicon tip of an AFM and showed that it was quite robust and could image the bottom of deep trenches inaccessible to conventional tips. By 1998, progress had been made in functionalized carbon nanotubes for use on AFM tips. Most notably, Wong and colleagues prepared nanotube tips by oxidation in air at 700°C, burning off all but 2% of the original material and leaving the ends covered with carboxyl (-COOH) groups whose chemistry is rich and well-understood. Four different kinds of tips were created: (1) the original carboxyl tip, which is acidic; (2) an amine-terminated tip (made by forming an amide bond to one of the amine groups in ethylenediamine (H₂NCH₂CH₂NH₂)), which is basic; (3) a hydrocarbon-terminated tip (made by forming an amide bond to benzylamine (C₆H₅CH₂NH₂)), which is hydrophobic; and (4) a biotin-terminated tip (made by forming an amide bond to a biotin derivarive), which shows specific binding to streptavidin. AFM contact forces between tips and selected substrates were shown to be sensitive to pH and to the chemical details of the substrate in ways consistent with the tips’ intended chemistry.

Wong’s tips had three closely related advantages over previous techniques. First, when a functional group is attached to the apex of an Si₃N₄ or SO₂ tip, these groups usually adhere to the sides of the tip as well; upon using the tip as a tool, there is a constant hazard that contact with the sides of the tip will alter the workpiece in unwanted places. Unlike these tips built with bulk techniques, the tips of Wong’s tips are very different from their sides—the carboxyl groups are attached only on the ends of the nanotube, not on the sides. Second, Wong’s tips have lateral dimensions set by the nanoscale dimensions of nanotubes, not by top-down fabrication techniques. Note the authors: “...the small effective radius of nanotube tips significantly improves resolution beyond what can be achieved using commercial silicon tips...we have recently demonstrated that lateral resolution of <3 nm can be achieved by using COOH-terminated single-walled nanotubes tips.”

Third, single-walled nanotube tips are much closer to yielding truly single atom tips with controlled chemistry than any other alternative. J. Soreff notes that a (10, 10) nanotube with a 1.4-nm diameter has just 20 atoms at an open end. Even given a statistical distribution of tubes with varying numbers of carboxyl groups attached to their ends, one should be able to build a ligand which covers the whole end of the tube, yielding a method for ensuring that just one molecule of known structure and orientation was present at the end of the probe.

Theoretical studies have employed molecular dynamics simulations and other techniques to investigate the behavior of mechanosynthetic tool tips. For example, Drexler proposed using an acetylene radical to perform selective abstractions of hydrogen from a diamond surface (Fig. 2.26A), and this proposal was further studied in Musgrave’s ab initio quantum chemistry analysis and Sinnott’s molecular dynamics modeling of the reaction at room temperature on a diamond (111) surface. Other H-abstraction studies have been done. One possible structure for a hydrogen abstraction tool might be similar to an anthracene whose end had been modified as shown in Figure 2.26B; prior to use, the structure must be activated by removal of the terminal hydrogen. Brenner and colleagues describe further work modeling the parallel abstraction of several hydrogen atoms from the diamond (111) surface. Such reactions are similar to those involved in the well-studied chemical
Merkle's system includes positionally-controlled hydrogen abstraction capable of fabricating a large class of useful stiff hydrocarbons. However, steric hindrance of multiple tools at one site also was not explicitly addressed. Another potential problem to be overcome is loss of positional registration as the various functionalized tool heads are changed out at the AFM tip; such tips might be paired with a second positionally-invariant metrology tip on each tool head, allowing precise calibration of workspace location before proceeding back to the workpiece.

Others have suggested fabricating nanoparticles by building up structures by stacking layers of individual atoms, perhaps constrained by sacrificial joiners and scaffolding, and honed by differential etching operations.

Assuming that "nanoparts" can be fabricated to atomic precision, can SPMs also be used to assemble these nanoparticles into working nanomachines? Some preliminary work has been accomplished. For instance, arrays of nanoscale holes have been synthesized that could serve as "workpiece holders." In one experiment, C_{60} buckyballs were scattered about on a prepared surface, then an AFM pushed the individual spheres into their own little holes, securely seating them pending further processing. The C_{60}-impregnated perforated film was so durable that the buckyballs remained trapped after several months of storage. Similarly, Jung and colleagues in Gimzewski's group synthesized 4-legged porphyrin-based molecular "nanoparts" that were specifically designed for easy positioning on a copper surface. Pushing with an STM at room temperature, the authors displaced the ~1.5-nm wide nanoparticles in predefined directions and rotated the 4-legged molecules at will, arranging, for example, the parts into a precise hexagonal configuration on the copper surface. The STM has been used to induce, and to view, the rotation of an individual oxygen molecule trapped on a platinum surface, as it is intentionally and stably rotated into any one of three distinct orientations. The forces required to slide, or alternatively to roll, a carbon nanotube across a graphic surface have been directly measured by AFM.

The first known instance of a crude but purely mechanical "nanopart" assembly operation was performed in 1995 by Paul E. Sheehan and Charles M. Lieber. Two nanoscale parts with rectangular slots and a third 50-nm rectangular sliding "latch" member were milled from a MoO_{3} crystal using an AFM, then the rectangular "latch" member was slid repeatedly from one slot to the other using the same AFM, making a three-nanopart reversible mechanical latch. Noted the authors: "The lateral force needed to break the latch, 41 nN, was large considering the small latch contact area, which suggests that relatively robust assemblies can be created with such devices. Such a reversible latch could serve as the basis for mechanical logic gates....Our results...demonstrate the ability to machine complex shapes and to reversibly assemble these pieces into interlocking structures."

In 1998, manipulation of nanoscale parts in the vertical dimension (e.g., normal to the surface plane) had only just begun. The first three-dimensional structure built out of single "nanoparts" was demonstrated in 1997 by Requicha's group at the USC Laboratory for Molecular Robotics. First, an AFM was used to push 5-nm gold nanoparticles in the third dimension up and over surface protrusions or obstructions—for example, a 5-nm gold particle was shoved up onto a 2-nm-high protrusion, then slid back off again onto the surface. In another series of experiments conducted in air at room temperature with 15-30 nm gold nanoparticles deposited on silicon previously coated with a silane layer, Requicha's group used an AFM to build a simple 3-D pyramidal structure by pushing one gold "nanopart" on top of two others, then off again; the group also used the AFM to rotate and translate a dimer unit formed by two linked "nanoparts."

---

*Rather than, butadiyne, chemically distinct from the more common molecule butadiene (C_{4}H_{6}), can polymerize explosively under some conditions.*

---

*Fig. 2.26. Schematic of proposed hydrogen abstraction tool (redrawn from Merkle).*
Three-dimensional nanoassembly might perhaps be more readily achieved if nanoparts could assist in their own assembly process. R. Merkle [personal communication, 1998] suggests that an SPM could be used to position DNA-tagged molecular building blocks. A building block tagged with a specific single-stranded DNA should preferentially adhere to a surface covered with the complementary DNA. If the building block has a second single-stranded DNA tag which is complementary to single-stranded DNA on the tip of an SPM, and if it is initially attached to the SPM tip by this tag, then it could be positioned by the SPM and attached to the surface. The relative strength of attachments could be adjusted by changing the number of base pairs in the complementary region between the two strands.\(^{1066}\) A related suggestion, due to G. Fahy,\(^{322}\) is that SPM tips for assembly could be designed with more than one binding site. The existence of at least two binding sites could allow precise orientation of a nanopart, permitting the nanopart to be presented to a desired target site (or another nanopart, possibly attached to another manipulator) at exactly the location and in exactly the orientation desired, assuming that each nanopart can bind to each tip in only one possible orientation. A convenient post-assembly nanopart release mechanism is also required.

In 1997, the Avouris group\(^{2750,2771}\) at IBM Yorktown Heights demonstrated the tip of an atomic force microscope (AFM) could be used to control the shape and position of individual multiwalled carbon nanotubes dispersed on a surface. Nanotubes could be bent, straightened, translated, rotated, and (under certain conditions) cut.

In 1998, Zyvex LLC (a developmental engineering company whose goal is to create a molecular assembler capable of manufacturing atomically precise structures\(^{2774}\)) demonstrated the ability to manipulate carbon nanotubes in three dimensions inside a scanning electron microscope (SEM) having >6 nm resolution at near-video scan rates. Zyvex’s custom piezoelectric vacuum manipulator achieved positional resolutions comparable to SPMs along with the ability to manipulate objects along one rotational and three linear degrees of freedom with 0.1 nm spatial resolution.\(^{2808}\) This prototype device could probe, select and handle limited classes of nanometer-scale objects such as carbon nanotubes under real-time SEM inspection. Carbon nanotubes were attached to commercial atomic force microscope (AFM) tips either by van der Waals forces alone or by “nanosoldering” by a concentrated electron beam from the SEM. Forces applied to the nanotubes could be measured from the cantilevers’ deflections (spring constants 0.01-100 N/m; see also Neumister and Ducker\(^{2809}\)). Sometimes nanotubes were transferred from tip to tip. The manipulator could function both as a research tool for investigating properties of carbon nanotubes and other nanoscale objects without surface restrictions, and as a rudimentary building device for larger nanotube assemblies. Zyvex believed that this capability to select and manipulate nanoscale components and to examine directly their suitability as construction materials during various phases of the construction process would play an important role in enabling the technology of assembling mechanical and electronic devices from prefabricated components. Most impressively, the Zyvex system allowed the simultaneous coordinated operation of three independently controlled AFM tips within the same workspace, at different orientations. (A four-tip system was expected to be operational by mid-1999 [Mark Dyer, personal communication, 1999].) Two-handed coordinated micromanipulation under the view of an SEM had previously been demonstrated by others.\(^{2810}\)

Tuzun and colleagues\(^{2811}\) calculated the requirements for coaxial docking of two nanotubes of different diameters to form a molecular bearing. They looked at bearings formed from nanotubes 11 rings long, with 10 carbons per ring in the shaft and either 30 or 34 carbons per ring in the sleeve. For the computer simulations, Tuzun placed the sleeve of the bearing along the z axis and gave the shaft a small initial velocity towards it. A perfectly aligned shaft falls straight into a potential well from the van der Waals attraction to the sleeve; displaced or mis-oriented shafts can bounce off the edge of the sleeve. The docking envelopes for the molecular dynamics calculations and for the rigid body calculations had essentially the same shape, but with slightly different sizes. For the atomistic calculation, a shaft aligned parallel to the sleeve (with a 30 carbon ring sleeve) can be displaced by 0.26 nm before it fails to dock, while in the rigid body calculation the displacement can only be 0.16 nm. The authors note that “if an end of the shaft points closely enough to the center of the sleeve, it will fall into the non-bonded potential energy well and the two nanotubes will dock.” These calculations are important because they tell us what tolerances can be accepted during the manual assembly process. Note the authors: “A question just as important as how to design or operate nanomachines is how to assemble them.”\(^{2811}\) Interestingly, one of Zyvex’s experiments\(^{2808}\) produced a multiwalled carbon nanotube that thinned in three steps along its length, suggesting the most likely explanation that the inner shells had been pulled from the outer shells. This phenomenon has also been observed by others and is known as the “sword and sheath” failure\(^{2812}\)—the exact inverse of the purposeful nanotube insertion operation investigated by Tuzun and colleagues.

In order to produce large numbers of nanoparts and nano-assemblies, massively parallel SPM arrays and microscale SPM\(^{2772-2774}\) would be most convenient. Force-sensing devices such as piezoelectric,\(^{2813}\) piezoresistive,\(^{2814}\) and capacitive\(^{2815}\) microcantilevers made it possible to construct microscale AFMs on chips without an external deflection sensor. (We exclude fixed-tip arrays\(^{2816}\) in the following discussion.) In 1995, Itoh and colleagues\(^{2817}\) at the University of Tokyo fabricated an experimental piezoelectric ZnO$_2$-on-SiO$_2$ microcantilever array of ten tips on a single silicon chip. Each cantilever tip lay ~70 microns from its neighbor, and measured 150 microns long, 50 microns wide and 3.5 microns thick, or ~26,000 micron$^2$/device, and each of the devices could be operated independently in the z-axis (e.g., vertically) up to nearly their mechanical resonance frequencies of 145-147 KHz at an actuation sensitivity of ~20 nm/volt—for instance, 0.3-nm resolution at 125 KHz.

Parallel probe scanning and lithography has been achieved by Quate’s group at Stanford, which has progressed from simple piezoresistive microcantilever arrays with 5 tips spaced 100 microns apart and 0.04-nm resolution at 1 KHz but only one z-axis actuator for the whole array,\(^{2818}\) to arrays with integrated sensors and actuators that allow parallel imaging and lithography with feedback and independent control of each of up to 16 tips, with scanning speeds up to 3 mm/sec using a piezoresistive sensor.\(^{2819,2820}\) By 1998, Quate’s group had demonstrated\(^{2821-2827,2829}\) arrays of 50-100 independently controllable AFM probe tips mounted in 2-D patterns with 60 KHz resonances, including a 10 x 10 cantilevered tip array fabricated in closely spaced rows using throughwafer interconnects on a single chip.

MacDonald’s group at the Cornell Nanofabrication Facility has pursued similar goals. In 1991, the team fabricated their first submicron stylus, driven in the xy plane using interdigitating MEMS comb drives,\(^{2830}\) including the first opposable tip pair (Fig. 2.27). By 1993, they had produced a 25-tip array on one xyz actuator,\(^{2831,2832}\) and by 1995 a complete working micro-STM (including xyz comb drives) measuring 200 microns on an edge and a micro-AFM measuring 2 mm on an edge including a 1-mm long cantilever with
Because these components could not yet be built in 1998, such designs could not be subjected to rigorous experimental testing and validation. Designers were forced instead to rely on ab initio structural analysis and computer studies including molecular dynamics simulations. Noted Drexler:10 “Our ability to model molecular machines (systems and devices) of specific kinds, designed in part for ease of modeling, has far outrun our ability to make them. Design calculations and computational experiments enable the theoretical studies of these devices, independent of the technologies needed to implement them.”

In nanoscale design, building materials do not change continuously as they are cut and shaped, but rather must be treated as being formed from discrete atoms.2867 A nanoscale component is a supermolecule, not a finely divided solid. Any stray atoms or molecules within such a structure may act as dirt that can clog and disable the device, and the scaling of vibrations, electrical forces, thermal expansion, magnetic interaction and surface tension with size lead to dramatically different phenomena as system size shrinks from the macroscale to the nanoscale.10

Molecular bearings are perhaps the most convenient class of components to design because their structure and operation is fairly straightforward. One of the simplest examples is Drexler’s overlap-repulsion bearing design,10 shown with end views and exploded views in Figure 2.28 using both ball-and-stick and space-filling representations. This bearing has exactly 206 atoms including carbon, silicon, oxygen and hydrogen, and is comprised of a small shaft that rotates within a ring sleeve measuring 2.2 nm in diameter. The atoms of the shaft are arranged in a 6-fold symmetry, while the ring has 14-fold symmetry, a combination that provides low energy barriers to shaft rotation. * Figure 2.29 shows an exploded view of a 2808-atom strained-shell** sleeve bearing designed by Drexler and Merkle10 using molecular mechanics force fields to ensure that bond lengths, bond angles, van der Waals distances, and strain energies are reasonable. This 4.8-nm diameter bearing features an interlocking-groove interface which derives from a modified diamond {100} surface. Ridges on the shaft interlock with ridges on the sleeve, making a very stiff structure. Attempts to bob the shaft up or down, or rock it from side to side, or displace it in any direction (except axial rotation, wherein displacement is extremely smooth) encounter a very strong resistance.279 Whether these bearings would have to be assembled in unitary fashion, or instead could be assembled by inserting one part into the other without damaging either part, had not been extensively studied or modeled by 1998.

Molecular gears are another convenient component system for molecular manufacturing design-ahead. For example, Drexler and Merkle10 designed a 3557-atom planetary gear, shown in side-, end-, and exploded views in Figure 2.30. The entire assembly has twelve moving parts and is 4.3 nm in diameter and 4.4 nm in length, with a molecular weight of 51,009,844 daltons and a molecular volume of 33,458 nm³. An animation of the computer simulation shows

---

* At the atomic scale, the two opposing surfaces have periodic bumps and hollows, but the periods of these bumps are different for the two surfaces—that is, they are “incommensurate.”10,309 Two incommensurate surfaces cannot lock up in any particular position, hence the barrier to free rotation is very low, on the order of ~0.001 kT (thermal noise at room temperature).10

** Components of high rotational symmetry may consist of (a) intrinsically curved, (b) strained-shell, or (c) special-case structures.10 In the case of (b), the bearing illustrated in Figure 2.29 has bond strains of around ~10% (~38 zJ/atom), and similar hydrocarbon bearings have been designed with bond strains of ~5% (~11 zJ/atom) [R. Merkle, personal communication, 1998]. For comparison, strain energies1866,2615 are ~3 zJ/atom for diamond lattice, ~25 zJ/atom in C_{60}, ~7-27 zJ/atom in the walls of infinite carbon nanotubes of diameter 0.7-1.3 nm, up to ~59 zJ (~12% strain) for some bonds around a Lomer dislocation in diamond,2860-70 zJ/atom in C_{60}, and at least ~80 zJ/atom for C_{70}. Fullerenes are among the most highly strained natural molecules ever isolated. For symmetrical diamondoid structures with negligible hoop stress, permissible bond strains may in theory be as large as ~140 zJ/atom producing a ~23% bond strain;10 nanotube breaking strain is 20-30% for various chiral forms, and buckling strain is ~8% in axial compression. Bond strain in a simple strained-shell bearing can be lowered by making the bearing bigger, thereby reducing the curvature. Thus strained shell bearings are feasible, although in 1998 it remained unclear exactly how small they could be before becoming unstable.

---
the central shaft rotating rapidly and the peripheral output shaft rotating slowly. The small planetary gears rotate around the central shaft, and they are surrounded by a ring gear that holds the planets in place and ensures that all of the components move in the proper fashion. The ring gear is a strained silicon shell with sulfur atom termination; the sun gear is a structure related to an oxygen-terminated diamond [100] surface; the planet gears resemble multiple hexasterane structures with oxygen rather than CH2 bridges between the parallel rings; and the planet carrier is adapted from a Lomer-dislocation array created by R. Merkle and L. Balasubramaniam, and linked to the planet gears using C-C bonded bearings. View (c) retains the elastic deformations that are hidden in (a)—the gears are bowed. In the macroscale world, planetary gears are used in automobiles and other machines where it is necessary to transform the speeds of rotating shafts.

W. Goddard and colleagues at CalTech performed a rotational impulse dynamics study of this "first-generation" planetary gear. At the normal operational rotation rates for which this component was designed (e.g., <1 GHz for <10 m/sec interfacial velocities), the gear worked as intended and did not overheat. Started from room temperature, the gear took a few cycles to engage, then rotated thermally stably at ~400 K. However, when the gear was driven to ~100 GHz, significant instabilities appeared although the device still did not self-destruct. One run at ~80 GHz showed excess kinetic energy causing gear temperature to oscillate up to 450 K above baseline. One animation of the simulation shows that the ring gear wiggles violently because it is rather thin. In an actual nanorobot incorporating numerous mechanical components, the ring gear would be part of a larger wall that would hold it solidly in place and would eliminate these convulsive motions which, in any case, are seen in the simulation only at unrealistically high operating frequencies.

Drexler and Merkle later proposed a "second-generation" planetary gear design (Fig. 2.31) with 4235 atoms, a molecular
weight of 72,491.947 daltons and a molecular volume of 47.586 nm³. This new version was indeed more stable but still had too much slip at the highest frequencies. Commenting on the ongoing design effort, Goddard suggested that an optimal configuration could have the functionality of a planetary gear but might have an appearance completely different from the macroscopic system, and offered an example: “Because a gear tooth in the xy plane cannot be atomically smooth in the z-direction, we may develop a Vee design so that the Vee shape of the gear tooth in the z-direction nestles within a Vee notch in the race to retain stability in the z-direction as the teeth contact in the xy plane. This design would make no sense for a macroscopic gear system since the gear could never be placed inside the race. However, for a molecular system one could imagine that the gear is constructed and that the race is constructed all except for a last joining unit. The parts could be assembled and then the final connections on the face made to complete the design analogous to the ZARBI system of Rebek.

Another class of nanodevice that has been designed is a gas-powered molecular motor or pump. The pump and chamber wall segment shown in Figure 2.32 contain 6165 atoms with a molecular weight of 88,190.813 daltons and a molecular volume of 63.984 nm³. The device can serve either as a pump for neon gas atoms or (if run backwards) as a motor that can convert neon gas pressure into rotary power. The helical rotor has a grooved cylindrical bearing surface at each end, supporting a screw-threaded cylindrical segment in the middle. In operation, rotation of the shaft moves a helical groove past longitudinal grooves inside the pump housing. There is room enough for small gas molecules only where facing grooves cross, and these crossing points move from one side to the other as the shaft turns, moving the neon atoms along. Goddard reported that preliminary molecular dynamics simulations of the device showed that it could indeed function as a pump, although “structural deformations of the rotor can cause instabilities at low and high rotational frequencies. The forced translations show that at very low perpendicular forces due to pump action, the total energy rises significantly and again the structure deforms.” Merkle acknowledged that the pump moves neon atoms at an energy cost of 185 Kcal/mole-Angstrom (12,900 zJ/atom-nm), which is not very energy-efficient. Further refinement of this crude design is clearly warranted.

Conveyor systems would also be useful. Employing a primitive molecular CAD software package called Crystal Sketchpad, G. Leach created a design for a 5-nanopart, ~2500-atom conveyor belt system comprised of two rollers, two axles, and a belt consisting of a strained thin-walled diamond sheet. The design has not been subjected to further computational analysis, either to minimize rotational energy barriers or to optimize rotational dynamics or operational stability.

Drexler and Merkle have also produced a preliminary design for a 2596-atom fine-motion controller (Fig. 2.33). A general-purpose molecular assembler arm must be able to move its “hand” by many atomic diameters, position it with fractional-atomic-diameter accuracy, and then execute finely-controlled motions, perhaps to transfer one or a few atoms in a guided chemical reaction. Human arms use large muscles and joints for large motions and more finely-controlled finger motions for precision; the device presented here can execute precise finger-like motions over several atomic diameters with associated 90-degree rotations. The core of the device consists of a shaft linking two hexagonal endplates, sandwiching a stack of eight rings, making a modified Stewart platform (Section 9.3.1.5). In a complete system, each ring would be rotated by a lever driven by a cam mechanism. Each ring supports a strut linked to a central platform (here shown raised.

---

Fig. 2.31. Side and top views of a 4235-atom “second generation” planetary gear (courtesy of K.E. Drexler and R.C. Merkle).

Fig. 2.32. Side views of a 6165-atom neon gas pump/motor (courtesy of K.E. Drexler and R.C. Merkle).

Fig. 2.33. Side view of a 2695-atom fine motion controller (courtesy of K.E. Drexler and R.C. Merkle).
In 1998, the capacity of such systems to design and manipulate large nanoscale mechanical components was extremely limited. A few very primitive molecular nanotechnology design packages had been attempted, including Crystal Sketchpad, DiamondCAD (www.zyvex.com/diamond.html), Molecular Assembly Sequence Software (www.carol.com/mass.shtml), Molecular Modelling Toolkit (starship.python.net/crew/hinsen/mmtk.html), and NanoCAD (world.std.com/~wware/ncad.html). The possible design of molecules for specific purposes using genetic software techniques had also been investigated.

2.4.2 Molecular Assemblers

Drexler has proposed the molecular assembler, a device resembling an industrial robot which would be capable of holding and positioning reactive moieties in order to control the precise location at which chemical reactions take place. This general approach would allow the construction of large atomically precise objects by a sequence of precisely controlled chemical reactions. Much like the ribosome in biology (Section 2.3.1), an assembler would build various classes of useful molecular structures following a sequence of instructions. During this process, the assembler would provide three-dimensional positional and full orientational control over the molecular component (analogous to the individual amino acid in the ribosome model) that is being added to a growing complex molecular structure (analogous to the growing polypeptide in the ribosomal model). In one approach, a molecular assembler may be capable of forming any one of several different kinds of chemical bonds (e.g., by changing tool tips), not just a single kind such as the peptide bond that the ribosome makes. In bonding atoms or molecules to one another, the assembler would provide any needed energy (especially if the reaction happens not to be energetically favored) through physical force, thus performing mecanosynthesis (as opposed to the traditional means of chemical synthesis in solution).

In another approach, a molecular assembler might be capable only of noncovalent assembly operations, wherein nanoparticles are fabricated by other means and then presented to the assembler, which assembles the nanoparticles into working nanomachines.

The first simple molecular assembler will almost certainly be a macroscale device, perhaps a modified SPM system as was being pursued by Zyvex in 1998. Multiple SPM heads could be equipped with a small number of nanoscale tool tips. In one scenario, nanoparticles fabricated using bulk chemistry techniques would be inspected and selected by the SPM, then assembled one by one into working nanomachines (e.g., the desired useful nanoscale products). Such assembly operations will be very slow, because the placement of each new component may require simultaneous rotations and translations of large macroscale SPM components. Assembly time scales roughly linearly with assembler size because smaller assembler components moving at a given velocity need to travel less distance to accomplish a given physical operation, hence consuming less time and energy per physical operation. An important early development goal thus will be to design and fabricate macroscale molecular assemblers.

*Existing CASD software/database packages include CAMEO (reaction chemistry assistant; William L. Jorgensen, Yale Univ.; zarbi.chem.yale.edu/programs/cameo.html), CAS (reaction data base; www.cas.org/CASFILES/casreact.html), Chiron (retrosynthetic analysis of chiral precursors; Steve Hanessian, Univ. of Montreal; www.netsci.org/Resources/Software/Cheminfo/chiron.html), CIARA (Vogel Scientific Software Inc.; www.vogelscientific.com), Crossfire (Computer Assisted Synthesis, Beilstein Information Systems; www.beilstein.com), ERDOS (Johann Gasteiger, Erlangen University; schiele.org/organik.uni-erlangen.de), LAHSA (retrosynthetic analysis; Alan Long, Harvard University; long@midas.harvard.edu), REACCS (MDL Information Systems Inc.; www.mdl.com), SYNGEN (Jim Hendrickson, Brandeis University; synegn2.chem.brandeis.edu/syngen.html), SYNLIB (W. Clark Still, Columbia University), and SYNTREE (Trinity Software Inc.; www.trinitysoftware.com/Trinity/orgchem/SYNTREE.HTM).
At its most basic level, the simplest possible nanoscale molecular assembler may be comprised of one or more nanoscale manipulators (Section 9.3). For example, the diamondoid nanoscale manipulator arm described in Section 9.3.1.4 (Figs. 9.8 and 9.9) has about 4 million atoms excluding the base and control and power structures; doubling the size to account for support structures gives a molecular weight per arm of ~100 megadaltons and a total molecular volume of ~140,000 nm³. Designed for high strength and stiffness, this robot arm should be able to hold a molecular fragment stiffly enough to make it react with a chosen end of a carbon-carbon double bond with an error rate of only 10⁻¹⁵.¹⁰ The robot arm could guide chemical reactions with high reliability at room temperature in vacuo, with little need for sensing the positions of the molecules with which it is working.²⁷⁹ Alternatively, with appropriate tool tips, such an arm could grasp and manipulate individual prefabricated nanoparticles in solution phase. Assembly of nanoparticles without fabrication may require only a very small tool set.

The manipulator arm must be driven by a detailed sequence of control signals, just as the ribosome needs mRNA to guide its actions. However, such detailed control signals can be provided by external acoustic, electrical, or chemical signals that are received by the robot arm via an onboard sensor or power transducer, using a simple “broadcast architecture”¹⁰,²⁸⁰,²⁸⁷ (Chapter 12), a technique which can also be used to import power (e.g., Section 6.3.3). Such transducers may be extremely small, on the order of (~10 nm)³ each.¹⁰ Interestingly, the biological cell may be regarded as an example of a broadcast architecture:²⁸⁷ the nucleus, located external to the cytoplasm, broadcasts mRNA chemical signals to millions of spatially diverse cytoplasmic ribosomes, thereby remotely controlling the construction of cellular proteins.

Thus a two-armed mechanical molecular assembler that receives power and instructions from some external agency may have a total molecular volume of ~300,000 nm³ (a cube ~67 nm on an edge), containing ~16 million atoms with a molecular weight of ~200 megadaltons—very roughly the mass and scale of a medium-size virus particle, such as an adenovirus. For comparison, the average enzyme (a biochemical chipping tool) weighs about ~0.1 megadalton, while a typical ribosome (a primitive “protein assembler”) weighs ~4.2 megadaltons. A simple mechanical assembler in the 10-100 megadalton range cannot be ruled out. However, such a device might have a very small set of tool tips and an extremely limited manufacturing repertoire.

Multiple nanoscale assemblers each capable of independent simultaneous actuation may require control signals more conveniently provided by an onboard nanocomputer (Section 10.2). This programmable nanocomputer must be able to accept stored instructions which are sequentially executed to direct the manipulator arm to place the correct moiety or nanopart in the desired position and orientation, thus giving precise control over the timing and locations of chemical reactions or assembly operations. The mechanical nanocomputer analyzed by Drexler¹⁰ requires ~16 nm³ per logic gate and ~40 nm³ per data register. If such components could be used to construct the equivalent of the most primitive 4-bit Intel 4004 microprocessor, then the nanocomputer could process ~10⁵ bits/sec (~25,000 ops/sec) at a ~1 KHz clock speed within a mechanism volume of ~36,000 nm³ (Section 10.2.1), again neglecting power supply, I/O linkages, and the like. An additional ~160,000 nm³ of rod logic registers (Section 10.2.1) adds ~1 kilobyte of onboard RAM memory or ~40 kilobytes of internal tape memory (conservatively assuming a tape storage density comparable to linear DNA). (A memory tape containing all bits needed for complete self-description (Table 2.1) may be considerably longer, even allowing for substantial data compression.) Doubling the total volume to account for support structure and other overhead gives a minimum mechanical nanocomputer molecular volume of ~400,000 nm³, roughly 70 million atoms with a molecular weight of ~800 megadaltons. Thus the smallest nanocomputer-driven nanoscale molecular assembler with two manipulator arms may have a total molecular volume of ~700,000 nm³ (a cube ~88 nm on an edge), containing ~86 million atoms with a molecular weight of ~1 gigadalton.

By 1998, only a small amount of research targeted at actual assembler design had begun. Following Drexler’s original discussions,¹⁰ during 1991-1998 R. Merkle authored or co-authored a continuing series of papers discussing various operational aspects and specific components of assembler design, including mecanosynthetic positional control,²⁷⁶,²⁷⁶ general design considerations for assemblers,²⁸⁰,²⁸⁶ the broadcast architecture,²⁸⁰,²⁸⁷ convergent assembly,²⁸⁶ binding sites,¹¹⁹ positioning devices,¹²⁵³ mecanosynthetic path sets,²⁶⁸⁵ designs for a neon pump,²⁶⁸⁵ and a fine motion controller,²⁵⁸⁹ and possible assembler casings.²⁸¹ J.S. Hall has considered high-level designs for a nanoscale parts-fabrication and parts-assembly nanorobot,²⁸⁰ Zytex,²⁷⁹³ founded in 1996 by James von Ehr, has set itself the task of building the first programmable nanoscale assembly in a 5-10 year time frame (e.g., by ~2006). W. Goddard and colleagues²⁸³ have proposed a series of molecular dynamics simulations of simple assemblers, although by 1998 these studies evidently had not yet begun. According to Goddard’s original proposal:

“Ultimately we need a programmable synthetic system to make a real device. Even though we may not have tools for all the chemical steps and may not have designs for all the pumps, engines, and transmissions needed, we propose to study the dynamics of simplified prototype assemblers. In these studies we anticipate having (1) a reservoir or supplies area for providing the various building units (atoms and fragments) required, (2) a work area in which we construct the nanomachine device (initially we will consider assembling the structure on top of a diamond surface), and (3) a molecular scale nano-hand which will extract the atoms from (1) and carry them to (2). We will then use extensions of our massive

---

**Table 2.1 Information Required to Describe Some Self-Replicating Systems**

<table>
<thead>
<tr>
<th>Self-Replicating System</th>
<th># of Bits Needed to Specify System</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penrose ratcheting 2-blocks</td>
<td>0.238 x 10³</td>
<td>2879</td>
</tr>
<tr>
<td>Rebek’s self-replicating molecules</td>
<td>~1 x 10⁴</td>
<td>131,2452</td>
</tr>
<tr>
<td>Human PrP protein (prion)</td>
<td>1.5 x 10⁵</td>
<td>2877,2878</td>
</tr>
<tr>
<td>Ebola virus</td>
<td>25.4 x 10⁴</td>
<td>2875</td>
</tr>
<tr>
<td>Penrose ratcheting 12-blocks</td>
<td>~49 x 10⁵</td>
<td>2880,2881</td>
</tr>
<tr>
<td>Typical human ribosome</td>
<td>280 x 10⁷</td>
<td>997</td>
</tr>
<tr>
<td>von Neumann’s Universal Constructor</td>
<td>~500 x 10³</td>
<td>280,1985</td>
</tr>
<tr>
<td>1988 Internet Worm</td>
<td>~500 x 10³</td>
<td>280,2874</td>
</tr>
<tr>
<td>Pyrenomus salina (algae)</td>
<td>1.32 x 10⁶</td>
<td>997</td>
</tr>
<tr>
<td>M. capricolum (mycoplasma)</td>
<td>1.448-2.2 x 10⁶</td>
<td>2876</td>
</tr>
<tr>
<td>E. coli (bacterium)</td>
<td>8.4 x 10⁶</td>
<td>997</td>
</tr>
<tr>
<td>S. cerevisiae (yeast)</td>
<td>26 x 10⁶</td>
<td>997</td>
</tr>
<tr>
<td>D. melanogaster (insect)</td>
<td>280 x 10⁶</td>
<td>997</td>
</tr>
<tr>
<td>200-megadalton 2-arm nanoassembler</td>
<td>384 x 10⁴</td>
<td>Section 2.4.2</td>
</tr>
<tr>
<td>1-gigadalton 2-arm nanoassembler</td>
<td>2.32 x 10⁹</td>
<td>Section 2.4.2</td>
</tr>
<tr>
<td>G. domesticus (bird)</td>
<td>2.4 x 10⁹</td>
<td>997</td>
</tr>
<tr>
<td>X. laevis (amphibian)</td>
<td>6.2 x 10⁹</td>
<td>997</td>
</tr>
<tr>
<td>H. sapiens (human genome)</td>
<td>6.6 x 10⁹</td>
<td>997</td>
</tr>
<tr>
<td>Various flowering plant species</td>
<td>0.12-220 x 10¹⁰</td>
<td>997</td>
</tr>
<tr>
<td>NASA Lunar Manufacturing Facility</td>
<td>272 x 10⁷</td>
<td>115</td>
</tr>
</tbody>
</table>
molecular dynamics program to operate the system: moving the tip from reservoir to work area, moving it to the contribute appropriate surface site, moving it to regenerate the active tip, and then moving it back to add new atoms and molecules. This will include proper temperature effects, molecular vibrations, energy release upon the various chemical steps, etc. The ground rules here are that a realistic force field be used and that all pieces be treated at the atomic level (but some might be semi-rigid). This will use the force field developed for nanosynthesis. The purpose of these simulations is to examine issues of vibration caused by chemical forces as the tool picks up and delivers atoms to the growing surface. Also we want to consider the effect of energy release in the chemical steps on the thermal fluctuations in these systems (which may cause displacements and vibrations)."

Building mega-atom or giga-atom nanoproducts one at a time would be incredibly expensive and time consuming. For example, imagine that we wish to construct a simple medical nanorobot such as the 1-micron spherical respirocyte\(^{1400}\) described in Chapter 22, which consists of ~18 billion atoms (dry structure). A factory employing a coordinated team of 100 macroscale SPM assemblers, each able to place ~1 atom/sec-SPM on a convergently-assembled workpiece achieves a pitiful manufacturing throughput rate of ~2 respirocytes per decade. Nanoscale assemblers with appendages ~10\(^6\) times smaller than macroscale SPM assemblers might plausibly achieve net assembly rates of ~10\(^7\) atoms/sec-SPM assemblers, but even a production line employing ~300 such nanoassemblers (which the abovementioned 100-SPM factory team could build in ~1 year, assuming ~10\(^7\) atoms/nanoassembler) can only manufacture ~1 respirocyte nanorobot per minute. At that rate, it would take ~2 million years to build the first ~1 cm\(^3\) therapeutic dosage containing ~10\(^12\) respirocyte nanorobots.

The necessary solution to this mass-production bottleneck is to employ any of several massively parallel approaches to manufacturing, including such techniques as self-assembly (Sections 2.3.1 and 2.3.2), convergent assembly,\(^{10,2869}\) or, most usefully, self-replication (Chapter 14). The basic advantage of self-replication is readily illustrated. Consider a single ~10\(^5\)-atom "seed" nanoassembler (having an onboard nanocomputer) that has been painstakingly built in ~10\(^6\) sec using the abovementioned 100-SPM macroscale assembler team, working at ~1 atom/sec-SPM. The seed nanoassembler is first programmed to build a copy of itself, which it accomplishes (working at ~10\(^6\) atoms/sec-nanoassembler) in ~100 sec. These two nanoassemblers then each build a copy of themselves in another ~100 sec; now there are four nanoassemblers. After ~48 generations, requiring a total of ~80 minutes to complete, there are ~3 x 10\(^14\) nanoassemblers.* These nanoassemblers are then reprogrammed for the manufacture of 18-billion-atom respirocytes and are fed the appropriate, presumably different, feedstock. This vastly expanded nanoassembler manufacturing system can now produce ~10\(^12\) respirocytes (~1 cm\(^3\) therapeutic dose) per minute.

The design of machines able to make copies of themselves was first described by von Neumann.\(^{1985}\) Many variations on this theme have been reviewed by Freitas and Gilbreath\(^{115}\) and Sipper,\(^{2871}\) and design for self-replication in the context of nanoscale assemblers has been considered by Drexler,\(^{8,10}\) Merkle,\(^{116,2868,2872,2873}\) and Hall.\(^{2870}\) As Drexler\(^{10}\) notes: "It may seem somehow paradoxical that a machine can contain all the instructions needed to make a copy of itself, including those selfsame complex instructions, but this is easily resolved. In the simplest approach, the machine reads the instructions twice: first as commands to be obeyed, and then as data to be copied. Adding more data does not increase the complexity of the data-copying process, hence the set of instructions can be made as complex as is necessary to specify the rest of the system. By the same token, the instructions transmitted in a replication cycle can specify the construction of an indefinitely large number of other artifacts." The estimated information content of self-replicating systems—the length of the instruction tape—can be surprisingly small (Table 2.1). Von Neumann's original analysis\(^{1985}\) concluded that perhaps twelve different kinds of units of unknown complexity could be required as building materials, and Haldane\(^{2884}\) inferred that as many as ~10\(^7\) individual parts might be needed to make a replicator. This inference was refuted just three years later with the arrival of the first in a stream of very simple but ingenious designs for mechanical self-replicating machines that were assembled and operated in the late 1950s. In the first example,\(^{2879}\) a pair of joined raftletike blocks,\(^{2886}\) when placed in a "sea" of left and right blocks and then physically agitated, replicated itself from this well-ordered input substrate, making more block-pairs until all the single blocks were used up, a process similar to chemical autocatalysis. More complex congeeries of blocks, including clever four-block, eight-block, and 12-block replicators that could replicate themselves in a sea of blocks were presented by Penrose,\(^{2880,2881}\) Jacobson,\(^{2882}\) and demonstrated a 3-unit replicator consisting of toy train engines circulating on HO model railroad tracks, and Morowitz\(^{2883}\) designed a simple two-unit device with one unit comprised of about a dozen components including switches, batteries and electromagnets, that could assemble copies of itself from parts floating on the water surface of a bathtub. In 1998, Lohn and colleagues\(^{2885}\) gave two designs for simple self-replicating systems that could be constructed out of wood, batteries and electromagnets, with explicit analogies to nanoscale fabrication. Several important conclusions may be drawn from these examples:

First, replication is fundamentally so simple a task that machines capable of displaying this behavior pre-date most of the modern electronic computer era.

Second, assembly is an inherently simpler operation than fabrication. A complex part may embody hundreds or thousands of prior fabrication and assembly operations, yet may be assembled within an assemblage in a single step. Hence nanopart assembly may be an easier candidate for early implementation in first-generation self-replicating molecular manufacturing systems than is molecular fabrication.

Third, and most important, the simplest nanoreplicator may require the ability to assemble, but not to (atomically) fabricate, in order to replicate itself. A nanomachine capable of assembly alone can replicate itself only from a very limited set of well-ordered input materials of relatively high complexity, but such a machine can be extremely simple both in structure and function. Certainly a nanomachine capable of both assembly and fabrication can replicate itself from a more diverse and disordered set of more elementary input materials—but only at the expense of far greater internal structural and functional complexity. At least two distinct models of replication have been identified in replicating systems design.\(^{112}\) The first model may be called the "unit replication" or organismic model (Fig. 2.34), in which the replicator is an independent unit which employs the surrounding substrate to directly produce an identical copy of itself; both the original and the copy remain fertile and may replicate again, thus

*Numerous complexities and limitations have been ignored here. For example, exponential growth cannot continue indefinitely in a finite environment with limited materials transport speeds—e.g., a 2-nm wide nanorobot suspended in 310 K water diffuses only~26 nm in 10\(^4\) sec (Eqn. 3.1). Mechanical replicating systems designed purely for molecular manufacturing may be inflexible and brittle, employing limited energy resources and materials feedstocks not found outside of the immediate manufacturing environment.
exponentiating their numbers. The second model may be called the “unit growth” or factory model (Fig. 2.35), in which a population of specialist devices, each one individually incapable of self-replication, can collectively fabricate and assemble all necessary components comprising all specialist devices within the system, hence the factory is capable of expanding its size (or of manufacturing duplicate factory systems) indefinitely in an appropriate environment. The factory model is sometimes called “bootstrapping,” which is any production of more productive capacity, from less. T. McKendree [personal communication, 1999] likens the situation to an ant colony: “There necessarily is the line of reproducing queens, making a sufficient number of ants feasible; [however,] most of the queen’s products are ‘workers’ that perform useful functions, but themselves are incapable of reproduction.”

It cannot be emphasized too strongly that mechanical medical nanodevices will not self-replicate inside the human body, nor will they have any need for self-replication themselves (Section 1.3.3). Machines that perform medical tasks are fundamentally different from machines that manufacture other machines. As R. Merkle explains: 373 “While self-replicating systems are the key to low cost [manufacture], there is no need (and little desire) to have such systems function in the outside world. Instead, in an artificial and controlled environment they can manufacture simpler and more rugged systems that can then be transferred to their final destination. Medical devices designed to operate in the human body don’t have to self-replicate: we can manufacture them in a controlled environment and then inject them into the patient as needed. The resulting medical device will be simpler, smaller, more efficient and more precisely designed for the task at hand than a device designed to perform the same function and self-replicate. This conclusion should hold generally: optimize [product] device design for the desired function, manufacture the [product] device in an environment optimized for manufacturing, then transport the [product] device from the manufacturing environment to the environment for which it was designed. A single device able to do everything would be harder to design and less efficient.”

In an effort to stimulate scientific and engineering interest in constructing the first nanoassembler, in November 1995 the Foresight Institute created the $250,000 Feynman Grand Prize, with funding contributed by Zyvex founder James von Ehr and St. Louis venture capitalist Marc Arnold. The prize will be awarded to the individual or group that first achieves both of two significant nanotechnology breakthroughs—first, the design and construction of a functional nanometer-scale robotic arm, and second, the design and construction of a functional 8-bit adder computing nanodevice. According to Grand Prize rules, the robotic nanomanipulator must fit entirely inside a 100-nm cube, carry out actions directed by input signals of specified types, be able to move to a directed sequence of positions anywhere within a 50-nm cube, complete all directed actions with a positioning accuracy of 0.1 nanometer or better, and perform at least 1,000 accurate, nanometer-scale positioning motions per second for at least 60 consecutive seconds. The adder must fit entirely within a 50-nm cube, and must be capable of adding accurately any pair of 8-bit binary numbers, discarding overflow, accepting input signals of specified types, and producing its output as a pattern of raised nanometer-scale bumps on an atomically precise and level surface. (J.S. Hall notes that a conventional 8-bit adder may be constructed using a total of 94 AND, OR, and NOT gates; using XOR gates, the total may be reduced to 37 gates.) Both devices may accept inputs from acoustic, electrical, optical, diffusive chemical, or mechanical means, although
any mechanical driving mechanism used for input must be limited to a single linkage that either slides or rotates on a single axis. To demonstrate the capacity for mass-production, at least 32 copies of each device must be provided for analysis and destructive testing by judges.

Once nanoassemblers are available, the design of far more complex nanomachines each containing tens or hundreds of billions of precisely arranged atoms will require new molecular CAD (Section 2.4.1) tools and techniques, including automated hierarchical design decomposition of objects (that must be built up from an array of nanoparticles) and a shape description language.⁹ Nanosystem design compilers, conceptually similar to silicon compilers that generate a complex pattern of transistors and conductors from an abstract specification of the properties of a digital circuit, will also be required for efficient nanorobot design.⁹ Given a fully specified design that meets all constraints necessary to physically permit assembly, all operations of the assembly process must be specified using an assembly-process compiler. Explains Drexler:¹⁰ “The design of structures and assembly procedures by hierarchical decomposition directly generates a tree of assembly steps. If this tree has been chosen to generate parts of the appropriate sizes and numbers, then it can be mapped onto a manufacturing layout. The vast number of manipulator motions to be specified at the finest, earliest, and most-dispersed levels of the assembly process can be planned in parallel by identical software systems running almost independently on separate processors. The result of this parallel assembly-process compilation is a set of instructions that, when executed by the transportation system and manipulator controllers of a manufacturing mechanism, will result in the assembly of an object corresponding to the initial design.” Scale-up to larger systems, including the control, coordination, and programming of large aggregates of cooperating nanorobots, is discussed in Volume II.

Since molecular manufacturing systems can be used to make more molecular manufacturing systems, it is believed that the capital costs of production can be quite low. Drexler²⁴³ claims that an analysis of inputs, outputs, and productivity suggests the total cost of production can be in the range familiar in agriculture and in the production of industrial chemicals—on the order of tens of cents per pound. Merkle²⁶⁷ has made similar assertions, for example: “Eventually (after amortization of possibly quite high development costs), the price of assemblers (and of the objects they build) should be no higher than the price of other complex structures made by self-replicating systems. Potatoes—which have a staggering design complexity involving tens of thousands of different genes and different proteins directed by many megabits of genetic information—cost well under a dollar per pound.” Such analogies from conventional agriculture or industrial chemistry are not strictly applicable to the medical products and pharmaceutical areas, where structural precision, product repeatability, and intensive quality control are of supreme importance, and where most products are subject to rigorous government regulation—unlike potatoes, which vary widely in size, shape, and molecular constitution, and generally do not require stringent product liability insurance, multi-phase clinical trials, or FDA approval.

Still, it is likely that molecular manufacturing will eventually allow reasonably inexpensive production of large batches of designed medical nanorobots. As a very crude analogy, even in 1995 a typical mail-order bioscience vendor²⁸⁸⁷ could manufacture 15 micromoles (~9 x 10¹⁸ product items) of customized oligonucleotides for $18/base up to at least 110 bases with a convenient 48-hour turnaround on orders; other online vendors offer similar or slightly higher prices up to 200-mer sequences with additional charges for necessary purifications.²⁸⁸⁸ If a billion-base structure could be induced to self-assemble (Section 2.3.1) from 10 million different sets of 100-mer custom-built oligonucleotides (~20 billion atoms), forming a desired -20-billion-atom nanomachine, then, optimistically assuming a -100% yield, the cost would be -$2 x 10⁻⁹ per nanomachine or -$2000 per -1 cm⁻³ of product volume which would include 10¹² finished nanomachines. In 1998, many specialty biotechnology-related drug treatments were comparably expensive. For example, treatments for multiple sclerosis using two closely related -22.5 kilodalton interferons were administered in dosages of -1 milligram/week at a treatment cost of ~$10,000/yr.²⁹⁵⁰ which is equivalent to a treatment cost of -$2 x 10⁻⁹ (2 “nanodollars”) per giga-atom. The famous HeLa (Henrietta Lacks) cell line²⁸⁹¹ is an example of a once tiny population of unique cancerous cells that were purposely replicated in vitro (~24-hour replication time) over many decades and are now widely available worldwide at very low cost (e.g., $216 per 1-cm³ ampule of the ATCC CCL-2 HeLa cell line²⁸⁹¹). Swallowable therapeutic pills containing bacteria...
(i.e., natural biological nanomachines) are already widely available over the counter for gastrointestinal reflation, as for example Salivarex™ which "contains a minimum of ~2.9 billion beneficial bacteria per capsule", and Alkadophilus™ which "contains 1.5 billion organisms per capsule", both at a 1998 price of ~$0.2 x 10^9 per bacterium.

Another useful analogic approach to cost estimation is suggested by the plunging price of computer power during the 20th century. Figure 2.36 shows the cost, in constant 1998 U.S. dollars, of a computer processing capability equivalent to -1 TeraFLOP (10^{12} floating-point operations per second, -10^{14} bits/sec assuming 64-bit words). The plotted data is from a compilation of 67 historical computers by Moravec through 1989, with an additional 27 data points for other computers spanning 1973-1999 added by the author. The chart shows that the cost of 1 TFLOP has fallen from $2 x 10^{20} in 1908 (the mechanical Hollerith Tabulator) to just $6 x 10^6 in 1998 (the SGI/Cray T3E-1200E), a 90-year price decline of almost 14 orders of magnitude—on average, a halving every ~2 years, a century-long example of Moore's Law. This trend may give us some confidence that even if the first individual 1 micron³ nanorobot costs ~$100,000 to assemble (much like the first Nippondenso microcar: Section 2.2), eventually the unit price may fall at least 14 orders of magnitude to below ~$10^{-9} per nanorobot or <$1000 per -1 cm³ dosage (10^{12} nanorobots). Recycling or remanufacturing each nanomachine ~1000 times before final discard would then imply a net ~$1/cm³ treatment cost. Of course, in 1998 all such estimates were mere guesswork, but the ultimate expectation of relatively low-cost nanomedical treatments was certainly rational, if not provable.

This Chapter has demonstrated that there are many different pathways leading toward molecular nanotechnology, each one of which provides incremental benefits to motivate further travel down the paths (Fig. 2.37). The likelihood of all the paths being blocked is low, even if we have little confidence that we can predict which approach will succeed first. Progress along the many pathways will provide precursor products, including some early medical applications, but the ultimate achievement of molecular manufacturing will finally make nanomedicine feasible.

As Drexler concluded Nanosystems: "Each step along [these pathways] will present great practical challenges, but each step will also bring valuable new capabilities. The long-term rewards, measured in terms of scientific and technological capabilities, appear large." This author agrees. For the remainder of the present work, we shall assume that a molecular manufacturing technology will someday exist that can economically manufacture macroscopic batches of microscopic machines comprised of atomically-precise nanoscale components. The medical applications and implications of such a technological capability are the principle subject of this three-volume book.
**CHAPTER 3**

**Molecular Transport and Sortation**

### 3.1 Human Body Chemical Composition

The human body consists of ~7 x 10²⁷ atoms arranged in a highly aperiodic physical structure. Although 41 chemical elements are commonly found in the body’s construction (Table 3.1), CHON comprises 99% of its atoms. Fully 87% of human body atoms are either hydrogen or oxygen.

Somatic atoms are generally present in combined form as molecules or ions, not individual atoms. The molecules of greatest nanomedical interest are incorporated into cells or circulate freely in blood plasma or the interstitial fluid. Table 3.2 summarizes the gross molecular contents of the typical human cell, which is 99.5% water and salts, by molecule count, and contains ~5000 different types of molecules. Appendix B lists 261 of the most common molecular and cellular constituents of human blood, and their normal concentrations in whole blood and plasma. This listing is far from complete. The human body is comprised of ~10⁵ different molecular species, mostly proteins—a large but nonetheless finite molecular parts list. By 1998, at least ~10⁴ of these proteins had been sequenced, ~10³ had been spatially mapped, and ~7,000 structures (including proteins, peptides, viruses, protein/nucleic acid complexes, nucleic acids, and carbohydrates) had been registered in the Protein Data Bank which at that time was maintained at Brookhaven National Laboratory.

Given the current accelerating pace of improving technology, it is likely that the sequences and 3-D or tertiary structures of all human proteins will have been determined by the second decade of the 21st century.

Transporting and sorting such a broad range of essential molecular species will be an important basic capability of many nanomedical systems. The three principal methods for distinguishing and conveying molecules that are most useful in nanomedicine are diffusion transport (Section 3.2), membrane filtration (Section 3.3), and receptor-based transport (Section 3.4). The Chapter ends with a brief discussion of binding site engineering (Section 3.5).

### Table 3.1 Estimated Atomic Composition of the Lean 70-kg Male Human Body

<table>
<thead>
<tr>
<th>Element</th>
<th>Sym</th>
<th># of Atoms</th>
<th>Element</th>
<th>Sym</th>
<th># of Atoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
<td>H</td>
<td>4.22 x 10²⁷</td>
<td>Cadmium</td>
<td>Cd</td>
<td>3 x 10¹⁰</td>
</tr>
<tr>
<td>Oxygen</td>
<td>O</td>
<td>1.61 x 10²⁷</td>
<td>Boron</td>
<td>B</td>
<td>2 x 10⁹</td>
</tr>
<tr>
<td>Carbon</td>
<td>C</td>
<td>8.03 x 10²⁶</td>
<td>Manganese</td>
<td>Mn</td>
<td>1 x 10⁷</td>
</tr>
<tr>
<td>Nitrogen</td>
<td>N</td>
<td>3.9 x 10²³</td>
<td>Nickel</td>
<td>Ni</td>
<td>1 x 10⁷</td>
</tr>
<tr>
<td>Calcium</td>
<td>Ca</td>
<td>1.6 x 10²³</td>
<td>Lithium</td>
<td>Li</td>
<td>1 x 10⁷</td>
</tr>
<tr>
<td>Phosphorus</td>
<td>P</td>
<td>9.6 x 10²³</td>
<td>Barium</td>
<td>Ba</td>
<td>8 x 10⁶</td>
</tr>
<tr>
<td>Sulfur</td>
<td>S</td>
<td>2.6 x 10²⁴</td>
<td>Iodine</td>
<td>I</td>
<td>5 x 10⁵</td>
</tr>
<tr>
<td>Sodium</td>
<td>Na</td>
<td>2.5 x 10²⁴</td>
<td>Tin</td>
<td>Sn</td>
<td>4 x 10⁵</td>
</tr>
<tr>
<td>Potassium</td>
<td>K</td>
<td>2.2 x 10²⁴</td>
<td>Gold</td>
<td>Au</td>
<td>2 x 10⁵</td>
</tr>
<tr>
<td>Chlorine</td>
<td>Cl</td>
<td>1.6 x 10²⁴</td>
<td>Zinc</td>
<td>Zn</td>
<td>2 x 10⁵</td>
</tr>
<tr>
<td>Magnesium</td>
<td>Mg</td>
<td>4.7 x 10²⁴</td>
<td>Cobalt</td>
<td>Co</td>
<td>2 x 10⁴</td>
</tr>
<tr>
<td>Silicon</td>
<td>Si</td>
<td>3.9 x 10²⁴</td>
<td>Cesium</td>
<td>Cs</td>
<td>7 x 10³</td>
</tr>
<tr>
<td>Fluorine</td>
<td>F</td>
<td>8.3 x 10²³</td>
<td>Mercury</td>
<td>Hg</td>
<td>6 x 10³</td>
</tr>
<tr>
<td>Iron</td>
<td>Fe</td>
<td>4.5 x 10²³</td>
<td>Arsenic</td>
<td>As</td>
<td>6 x 10³</td>
</tr>
<tr>
<td>Zinc</td>
<td>Zn</td>
<td>2.1 x 10²³</td>
<td>Chromium</td>
<td>Cr</td>
<td>6 x 10³</td>
</tr>
<tr>
<td>Rubidium</td>
<td>Rb</td>
<td>2.2 x 10²²</td>
<td>Molybdenum</td>
<td>Mo</td>
<td>3 x 10³</td>
</tr>
<tr>
<td>Strontium</td>
<td>Sr</td>
<td>2.2 x 10²²</td>
<td>Selenium</td>
<td>Se</td>
<td>3 x 10³</td>
</tr>
<tr>
<td>Bromine</td>
<td>Br</td>
<td>2 x 10²¹</td>
<td>Beryllium</td>
<td>Be</td>
<td>3 x 10³</td>
</tr>
<tr>
<td>Aluminum</td>
<td>Al</td>
<td>1 x 10²¹</td>
<td>Vanadium</td>
<td>V</td>
<td>8 x 10²</td>
</tr>
<tr>
<td>Copper</td>
<td>Cu</td>
<td>7 x 10²⁰</td>
<td>Uranium</td>
<td>U</td>
<td>2 x 10⁷</td>
</tr>
<tr>
<td>Lead</td>
<td>Pb</td>
<td>3 x 10²⁰</td>
<td>Radium</td>
<td>Ra</td>
<td>8 x 10³</td>
</tr>
</tbody>
</table>

**Total: 6.71 x 10²⁷**

### Table 3.2 Estimated Gross Molecular Contents of a Typical 20-µm Human Cell

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Mass % (daltons)</th>
<th># Molecules</th>
<th>Molecule %</th>
<th>Number of Molecular Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>65</td>
<td>1.74 x 10¹⁵</td>
<td>98.73</td>
<td>1</td>
</tr>
<tr>
<td>Other Inorganic</td>
<td>1.5</td>
<td>1.31 x 10¹⁲</td>
<td>0.74</td>
<td>20</td>
</tr>
<tr>
<td>Lipid</td>
<td>12</td>
<td>8.4 x 10¹⁰</td>
<td>0.475</td>
<td>50</td>
</tr>
<tr>
<td>Other Organic</td>
<td>0.4</td>
<td>7.7 x 10⁹</td>
<td>0.044</td>
<td>~200</td>
</tr>
<tr>
<td>Protein</td>
<td>20</td>
<td>50,000</td>
<td>0.011</td>
<td>~5,000</td>
</tr>
<tr>
<td>RNA</td>
<td>1.0</td>
<td>1 x 10⁴</td>
<td>3 x 10⁻³</td>
<td>---</td>
</tr>
<tr>
<td>DNA</td>
<td>0.1</td>
<td>1 x 10⁵</td>
<td>46</td>
<td>3 x 10⁻¹¹</td>
</tr>
</tbody>
</table>

**TOTALS 100% --- 1.76 x 10¹⁴ 100% ---**
9.4.2.1). Molecular transport to and from such nanodevices is governed by diffusion, not by bulk flow.

### 3.2.1 Brownian Motion

A particle suspended in a fluid is subjected to continuous collisions, from all directions, with the surrounding molecules. If the velocities of all molecules were the same all the time, the particle would experience no net movement. However, molecules do not have a single velocity at a given temperature, but rather have a distribution of velocities of varying degrees of probability. Thus from time to time, a suspended particle receives a finite momentum of unpredictable direction and magnitude. The velocity vector of the particle changes continuously, resulting in an observable random zigzag movement called Brownian motion.

Einstein\(^{385}\) approximated the RMS (root mean square) displacement of a particle of radius \(R\) suspended in a fluid of absolute viscosity \(\eta\) and temperature \(T\), after an observation period \(\tau\), as:

\[
\Delta X = \left(\frac{k T}{3 \pi \eta R}\right)^{1/2} \text{ (meters)}
\]

(Eqn. 3.1)

where \(k = 1.381 \times 10^{-23} \text{ joule/kelvin} (K)\) or 0.01381 \(\text{zJ/K} (\text{Boltzmann’s constant})\). Particles under bombardment also experience a rotational Brownian motion around randomly oriented axes, with the RMS angle of rotation:

\[
\Delta \alpha = \left(\frac{k T}{4 \pi \eta R^3}\right)^{1/2} \text{ (radians)}
\]

(Eqn. 3.2)

although for \(\tau < \tau_{\text{min}} = M / 15 \pi \eta R\), where \(M\) is particle mass (see below), rotation is ballistic.

In human blood plasma, with \(\eta = 1.1 \text{ centipoise} (1.1 \times 10^{-3} \text{ kg/m/sec})\) and \(T = 310\) K, a spherical 1-micron diameter nanodevice (\(R = 0.5\) micron) translates \(-1\) micron in 1 sec (\(v_{\text{brownian}} = 10^{-6} \text{ m/sec}\)) or \(-8\) microns (= the width of a capillary) after 77 sec (\(v_{\text{brownian}} = 10^{-7} \text{ m/sec}\)), and rotates once in \(-16\) sec (\(\tau_{\text{min}} = 2 \times 10^{-8} \text{ sec}\)). In the same environment, a rigid 10-nm particle (roughly the diameter of a globular protein) would translate \(-8\) microns in one second (\(v_{\text{brownian}} = 10^{-7} \text{ m/sec}\)) while rotating \(-250\) times, due to Brownian motion (\(\tau_{\text{min}} = 2 \times 10^{-15} \text{ sec}\)).

The instantaneous thermal velocity over one mean free path, the average distance between collisions, is much higher than the net Brownian translational velocity would suggest. For a particle of mass \(M = 4/3 \pi \rho R^3\) with mean (working) density \(\rho\), the mean thermal velocity is:

\[
v_{\text{thermal}} = \left(3k T / M\right)^{1/2}
\]

(Eqn. 3.3)

At \(T = 310\) K, a spherical 1-micron diameter nanodevice of normal density (e.g., taking \(\rho = \rho_{\text{H2O}} = 993.4 \text{ kg/m}^3\) to minimize ballasting requirements; Section 10.3.6) has \(v_{\text{thermal}} = 5 \times 10^{-4} \text{ m/sec}\); for a spherical 10-nm diameter protein with \(\rho = 1500 \text{ kg/m}^3\), \(v_{\text{thermal}} = 4 \times 10^{-6} \text{ m/sec}\).

#### 3.2.2 Passive Diffusive Intake

Medical nanodevices will frequently be called upon to absorb some particular material from the external aqueous operating environment. Molecular diffusion presents a fundamental limit to the speed at which this absorption can occur. (Once a block of solution has passed into the interior of a nanodevice, it may be divergently subdivided and transported at \(-0.01\) m/sec along internal pathways of characteristic dimension \(-1\) micron far faster than the \(-1\) mm/sec bulk diffusion velocity across 1 micron distances; Section 9.2.7.5.)

For a spherical nanodevice of radius \(R\), the maximum diffusive intake current is:

\[
J = 4 \pi R D C
\]

(Eqn. 3.4)

where \(J\) is the number of molecules/sec presented to the entire surface of the device, assumed to be 100% absorbed (but see Section 4.2.5), \(D\) (m\(^2\)/sec) is the translational Brownian diffusion coefficient for the molecule to be absorbed, and \(C\) (molecules/m\(^3\)) is the steady-state concentration of the molecule far from the device.\(^{337}\) (Blood concentrations in gm/cm\(^3\) from Appendix B are converted to molecules/m\(^3\) by multiplying Appendix B figures by \((10^6 \times N_A/MW)\), where \(N_A = 6.023 \times 10^{23}\) molecules/mole (Avogadro’s number) and MW = molecular weight in gm/mole or daltons.) For rigid spherical particles of radius \(R\), where \(R >> R_{\text{H2O}}\), the Einstein-Stokes equation\(^{363}\) gives:

\[
D = k T / (6 \pi \eta R)
\]

(Eqn. 3.5)

though this is only an approximation because \(D\) varies slightly with concentration, with departure from molecular sphericalness, and other factors.

Measured diffusion coefficients in water for various molecules of physiological interest, converted to 310 K, are in Table 3.3. Diffusion coefficient data for ionic salts such as NaCl and KCl, which dissociate in water and diffuse as independent ions, are for solvated electrolytes. A 1-micron (diameter) spherical nanodevice suspended in arterial blood plasma at 310 K, with \(C = 7.3 \times 10^{22}\) molecules/m\(^3\) of oxygen and \(D = 2.0 \times 10^{-9} \text{ m}^2/\text{sec}\), encounters a flow rate of \(J = 9.2 \times 10^8\) molecules/sec of \(O_2\) impinging upon its surface. (The same calculation applied to serum glucose yields \(J = 1.3 \times 10^{10}\) molecules/sec.) The characteristic time for change mediated by diffusion in a region of size \(L\) scales as \(L^2/D\) (Eqn. 3.9, other factors.

In blood, the diffusivity of larger particles is significantly elevated because local fluid motions generated by individual red cell rotation lead to greater random excursions of the particles.\(^{388}\) The effective diffusivity \(D_e = D + D_r\), where the rotation-induced increase in diffusivity \(D_r = 0.25 R_{\text{rbc}}^2 g\), with red cell radius \(R_{\text{rbc}} = 2.8\) microns (taken for convenience as a spherical volume equivalent) and a typical blood shear rate (Section 9.4.1.1) \(\gamma = 500 \text{ sec}^{-1}\), giving \(D_r = 10^{-9} \text{ m}^2/\text{sec}\) in normal whole blood. The elevation of diffusivity caused by red cell stirring is just 50% for \(O_2\) molecules. However, for large proteins and viruses the effective diffusivity increases 10-100 times, and the effective diffusivity of particles the size of platelets is a factor of 10,000 higher than for Brownian molecular diffusion.

The diffusion current to the surface of a nanodevice can also be estimated for various nonspherical configurations.\(^{387}\) For instance, the diffusion current to both sides of an isolated thin disk of radius \(R\) is given by \(J = 8 R D C\). The two-sided current to a square thin plate of area \(L^2\) is \(J = (8/\pi) D L C\). The steady-state diffusion current to an isolated cylinder of length \(L_c\) and radius \(R\) is approximated by \(J = 2 \pi L_c D C / (\ln(2L_c/R) - 1)\), for \(L_c >> R\). The diffusion current through a circular hole of radius \(R\) in an impermeable wall separating regions of concentration \(c_1\) and \(c_2\) is \(J = 4 R D (c_1-c_2)\).

#### 3.2.3 Active Diffusive Intake

Foraging nanodevices operating in aqueous environments may only modestly exceed the maximum rates of passive diffusive intake described in Section 3.2.2 by engaging in active physical movements designed to increase access to the desired molecules, as described quantitatively below.
Section 3.2.2, which defines a characteristic diffusion frequency.

\[ \nu_{\text{diff}} = \frac{D}{L_a^2} \text{sec}^{-1} \]

Stirring will be more effective than diffusion only if \( \nu_{\text{stir}} > \nu_{\text{diff}} \), that is, if \( \nu_a > D / L_a \). For local stirring, \( L_a \) cannot be much larger than the size of the nanodevice itself. Assuming \( L_a = 1 \) micron and \( D = 10^{-9} \) m²/sec for small molecules, then \( \nu_a > 1000 \) microns/sec, a faster motion than is exhibited by bacterial cells but quite modest for nanomechanical devices (Section 9.3.1). With \( D = 10^{-11} \) m²/sec for large proteins and virus particles, \( \nu_a > 10 \) microns/sec, well within the normal microbiological range.

The ratio of stirring time to diffusion time, or Sherwood number, is:

\[ N_{Sh} = L_a \nu_a / D \]  

(Eqn. 3.6)

provides a dimensionless measure of the effectiveness of stirring vs. diffusion. For bacteria absorbing small molecules, \( N_{Sh} \sim 10^{-2} \). Micron-scale nanodevices with 1-micron appendages capable of 0.01-1 m/sec movement can achieve \( N_{Sh} \sim 10-1000 \) for small to large molecules, hence could be considerably more effective stirrers.

In a classic paper, Berg and Purcell\(^{337} \) analyzed the viscous frictional energy cost of moving the stirring appendages so that the fluid surrounding a spherical object (e.g., a nanodevice) of radius \( R \), out to some maximum stirring radius \( R_s \), is maintained approximately uniform in concentration. The objective is to transfer fluid from a distant region of relatively high concentration to a place much closer to the nanodevice, thereby increasing the concentration gradient near the absorbing surface. To double the passive diffusion current by stirring, the minimum required power density is:

\[ P_d \approx \frac{12 \eta D^2}{R^4} \left( \frac{R_s + 2 R}{R_s - 2 R} \right)^3 \]  

(watts/m³)  

(Eqn. 3.7)

If \( \eta = 1.1 \times 10^{-3} \) kg/m-sec, \( R = 0.5 \) micron, \( D = 10^{-9} \) m²/sec for small molecules, and using a modest \( L_a = 1 \) micron stirring apparatus giving \( R_s = 3R \), then \( P_d \sim 3 \times 10^2 \) watts/m³. This greatly exceeds the \( 10^{-2}-10^{-8} \) watts/m³ power density commonly available to biological cells (Table 6.8) but lies well within the normal range for nanomechanical systems which typically operate at up to \( 10^2 \) watts/m³. (Nanomedically safe in vivo power densities are discussed at length in Sections 6.5.2 and 6.5.3.) For \( D = 10^{-11} \) m²/sec for large molecules, \( P_d \sim 3 \times 10^3 \) watts/m³, which is reasonable even by biological standards. The maximum possible gain from stirring is \(-R/R_s\), because the current is ultimately limited to what can diffuse into the stirred region.

Local heating due to stirring is minor. Given device volume \( V \sim 1 \) micron\(^3\), \( P_d \sim 3 \times 10^3 \) watts/m³, mixing distance \( L_{\text{mix}} \sim 5 \) microns, and thermal conductivity \( K_t = 0.623 \) watts/m-K for water, then \( \Delta T = (P_d V / L_{\text{mix}} K_t) \) = 0.01 microkelvins; taking heat capacity \( C_V = 4.19 \times 10^3 \) J/m³-K for water, thermal equilibration time \( t_{\text{eq}} = L_{\text{mix}}^2 C_V / K_t \) = 0.2 milliseconds.

3.2.3.2 Diffusive Swimming

The second strategy for active diffusive intake by swimming, Again, the nanodevice is equipped with suitable active propulsion equipment (Section 9.4) which enables it to move so as to continuously encounter the highest possible concentration gradient near its surface. Consider a spherical motile nanorobot of radius \( R \) propelled at constant velocity \( v_{\text{swim}} \) through a fluid containing a desired molecule for which the surface of the device is essentially a perfect sink (Section 4.2.5). Applying the Stokes velocity field flow around the sphere to the standard diffusion equation, a numerical solution by Berg and Purcell\(^ {337} \) found that the fractional increase in the diffusion current due to swimming is proportional to \( v_{\text{swim}} \) for \( v_{\text{swim}} \ll D / R \), and to \( v_{\text{swim}}^{2/3} \) for \( v_{\text{swim}} \gg D / R \).

### Table 3.3 Translational Brownian Diffusion Coefficients for Physiologically Important Molecules Suspended in Water at 310 K (most values converted from measured data at 20 °C).\(^{390,754,761-763} \)

<table>
<thead>
<tr>
<th>Diffusing Particle in Water</th>
<th>Mol. Wt. (gm/mole)</th>
<th>Diff. Coeff. D (m²/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂</td>
<td>2</td>
<td>5.4 x 10⁻⁵</td>
</tr>
<tr>
<td>H₂O</td>
<td>18</td>
<td>2.31 x 10⁻⁵</td>
</tr>
<tr>
<td>O₂</td>
<td>32</td>
<td>2.0 x 10⁻⁵</td>
</tr>
<tr>
<td>Methanol</td>
<td>32</td>
<td>1.5 x 10⁻⁵</td>
</tr>
<tr>
<td>HCl</td>
<td>36.5</td>
<td>3.6 x 10⁻⁵</td>
</tr>
<tr>
<td>CO₂</td>
<td>44</td>
<td>1.9 x 10⁻⁵</td>
</tr>
<tr>
<td>NaCl</td>
<td>58.5</td>
<td>1.5 x 10⁻⁵</td>
</tr>
<tr>
<td>Urea</td>
<td>60</td>
<td>1.3 x 10⁻⁵</td>
</tr>
<tr>
<td>Glycine</td>
<td>75</td>
<td>1.0 x 10⁻⁵</td>
</tr>
<tr>
<td>KCl</td>
<td>75</td>
<td>2.0 x 10⁻⁵</td>
</tr>
<tr>
<td>α-Alanine isomer</td>
<td>89</td>
<td>9.5 x 10⁻⁴</td>
</tr>
<tr>
<td>β-Alanine isomer</td>
<td>89</td>
<td>9.7 x 10⁻⁴</td>
</tr>
<tr>
<td>Glycerol</td>
<td>92</td>
<td>8.8 x 10⁻⁴</td>
</tr>
<tr>
<td>CaCl₂</td>
<td>111</td>
<td>1.2 x 10⁻³</td>
</tr>
<tr>
<td>Glucose</td>
<td>180</td>
<td>7.1 x 10⁻³</td>
</tr>
<tr>
<td>Mannitol</td>
<td>182</td>
<td>7.1 x 10⁻³</td>
</tr>
<tr>
<td>Citric acid</td>
<td>192</td>
<td>6.9 x 10⁻³</td>
</tr>
<tr>
<td>Sucrose</td>
<td>342</td>
<td>5.6 x 10⁻³</td>
</tr>
<tr>
<td>Milk lipase</td>
<td>6,669</td>
<td>1.5 x 10⁻⁴</td>
</tr>
<tr>
<td>Ribonuclease</td>
<td>13,683</td>
<td>1.3 x 10⁻³</td>
</tr>
<tr>
<td>Insulin</td>
<td>24,430</td>
<td>7.7 x 10⁻³</td>
</tr>
<tr>
<td>Scarlet fever toxin</td>
<td>27,000</td>
<td>1.0 x 10⁻²</td>
</tr>
<tr>
<td>Somatostatin</td>
<td>27,100</td>
<td>9.6 x 10⁻³</td>
</tr>
<tr>
<td>Carbonic anhydrate Y</td>
<td>30,640</td>
<td>1.1 x 10⁻³</td>
</tr>
<tr>
<td>Plasma mucoprotein</td>
<td>44,070</td>
<td>5.6 x 10⁻³</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43,500</td>
<td>8.1 x 10⁻³</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>68,000</td>
<td>7.3 x 10⁻³</td>
</tr>
<tr>
<td>Serum albumin</td>
<td>68,460</td>
<td>6.5 x 10⁻³</td>
</tr>
<tr>
<td>Transferrin</td>
<td>74,000</td>
<td>6.2 x 10⁻³</td>
</tr>
<tr>
<td>Gonadotrophin</td>
<td>98,630</td>
<td>4.7 x 10⁻³</td>
</tr>
<tr>
<td>Collagenase</td>
<td>109,000</td>
<td>4.5 x 10⁻³</td>
</tr>
<tr>
<td>Lipase</td>
<td>120,000</td>
<td>5.3 x 10⁻³</td>
</tr>
<tr>
<td>Plasminogen (proﬁbrolysin)</td>
<td>134,000</td>
<td>3.1 x 10⁻³</td>
</tr>
<tr>
<td>Ceruloplasmin</td>
<td>143,300</td>
<td>5.0 x 10⁻³</td>
</tr>
<tr>
<td>γ-Globulin</td>
<td>153,100</td>
<td>4.2 x 10⁻³</td>
</tr>
<tr>
<td>Immunoglobulin G (IgG)</td>
<td>158,500</td>
<td>4.2 x 10⁻³</td>
</tr>
<tr>
<td>Hyaluronic acid</td>
<td>177,600</td>
<td>1.3 x 10⁻³</td>
</tr>
<tr>
<td>Glucose dehydrogenase</td>
<td>190,000</td>
<td>3.6 x 10⁻³</td>
</tr>
<tr>
<td>Fibrinogen</td>
<td>339,700</td>
<td>2.1 x 10⁻³</td>
</tr>
<tr>
<td>Collagen</td>
<td>345,000</td>
<td>7.1 x 10⁻³</td>
</tr>
<tr>
<td>Urease</td>
<td>482,700</td>
<td>3.7 x 10⁻³</td>
</tr>
<tr>
<td>Cytochrome a</td>
<td>529,800</td>
<td>3.8 x 10⁻³</td>
</tr>
<tr>
<td>α-Macroglobulin</td>
<td>820,000</td>
<td>2.5 x 10⁻²</td>
</tr>
<tr>
<td>β-Lipoprotein</td>
<td>2,663,000</td>
<td>1.8 x 10⁻²</td>
</tr>
<tr>
<td>Ribosome</td>
<td>4,200,000</td>
<td>1.3 x 10⁻²</td>
</tr>
<tr>
<td>Vical DNA</td>
<td>6,000,000</td>
<td>1.4 x 10⁻¹</td>
</tr>
<tr>
<td>Uronic mucoprotein</td>
<td>7,000,000</td>
<td>3.6 x 10⁻⁵</td>
</tr>
<tr>
<td>Tobacco mosaic virus</td>
<td>31,340,000</td>
<td>5.6 x 10⁻²</td>
</tr>
<tr>
<td>T7 Bacteriophage</td>
<td>37,500,000</td>
<td>9.5 x 10⁻¹</td>
</tr>
<tr>
<td>Polyhedral silkworm virus</td>
<td>916,200,000</td>
<td>2.3 x 10⁻²</td>
</tr>
<tr>
<td>1-µm spherical nanodevice</td>
<td>-8 x 10⁻¹</td>
<td>4.1 x 10⁻³</td>
</tr>
<tr>
<td>Platelet (~2.4 µm)</td>
<td>-9 x 10⁻¹</td>
<td>1.6 x 10⁻³</td>
</tr>
<tr>
<td>Red Blood Cell (~5.6 microns)</td>
<td>-6 x 10⁻¹</td>
<td>6.8 x 10⁻⁴</td>
</tr>
</tbody>
</table>
Diffusive intake is doubled at a swimming speed \( v_{\text{swim}} = 2.5 \, \text{D/R} \), which for 1-micron devices is \(-5000\) microns/sec when absorbing small molecules, \(-50\) microns/sec for large molecules. The viscous frictional energy cost to drive the nanodevice through the fluid, derived from Stokes’ law (Eqn. 9.73), requires an onboard power density of:

\[
P_d = 9 \, \eta \, v_{\text{swim}}^2 / 2 \, R^2 \tag{Eqn. 3.8}
\]

If \( \eta = 1.1 \times 10^{-3} \, \text{kg/m-sec} \), \( v_{\text{swim}} = 2.5 \, \text{D/R} \), \( R = 0.5 \, \text{micron} \), \( D = 10^{-9} \, \text{m}^2/\text{sec} \) for small molecules, then \( P_d \approx 5 \times 10^5 \, \text{watts/m}^3 \).

For large molecules with \( D = 10^{-11} \, \text{m}^2/\text{sec} \), \( P_d \approx 50 \, \text{watts/m}^3 \). Thus the energy cost of diffusive swimming appears modest for nanomechanical systems; gains in diffusion by swimming for nanodevices will be restricted primarily by the maximum safe velocity that may be employed in vivo (Section 9.4.2.6).

In general, outswimming diffusion requires movement over a characteristic distance \( L_s \approx D/v_{\text{swim}} \). For bacteria moving at \(-30\) micron/sec and absorbing small molecules, then \( L_s \approx 30 \, \text{microns} \), roughly the sprint distance exhibited by flagellar microbes such as *E. coli*. For micron-scale nanodevices moving at \(-1\, \text{cm/sec} \) (Section 9.4), \( L_s \approx 1-100 \, \text{nm} \) for large to small molecules.

### 3.2.4 Diffusion Cascade Sortation

Nanodevices may also use diffusion to sort molecules. One of the remarkable features of diffusive sortation is that an input sample consisting of a complex mixture of many different molecular species can sometimes be completely resolved into pure fractions without having any direct knowledge of the precise shapes or electrochemical characteristics of the molecules being sorted. This can be a tremendous advantage for nanodevices operating in environments containing a large number of unknown substances. Another major advantage is the ability to readily distinguish isomeric (though not chiral) molecules. As one example of many possible, molecules suspended in water will diffuse into an adjacent region of pure water at different speeds, giving rise to dissimilar time-dependent concentration gradients which may be exploited for sortation by interrupting the process before complete diffusive equilibrium is reached.

For simplicity, assume we wish to separate two molecular species initially present in solution in equal concentrations (\( c_1 = c_2 \)), but having unequal diffusion coefficients (\( D_1 < D_2 \)). Consider a separation apparatus with two chambers. Chamber A contains input sample concentrate. Chamber B contains pure water. A dilating gate (Section 3.3.2) separates the two chambers. The gate is opened for a time \( \Delta t \) approximated by:

\[
\Delta t = \frac{(\Delta X)^2}{2 \, D_2} - \frac{L^2}{2 \, D_2} \tag{Eqn. 3.9}
\]

which relates the diffusion coefficient to the mean displacement \( \Delta X \), taken here as \( L \), the length of Chamber B. Table 3.4 gives an estimate of the time required for diffusion to reach 90% completion for glycine, a typical small molecule, in aqueous solution.

After \( \Delta t \) has elapsed, the gate is closed. (A gate with 10-nm sliding segments moving at 10 cm/sec closes in 0.1 microsec.) The faster-diffusing component \( D_2 \) approaches diffusive equilibrium in Chamber B, but the slower-diffusing component does not; it is present only in smaller amounts. This gives a separation factor \( c_2/c_1 \) - \( D_2/D_1 \) for each diffusion sortation unit. If \( n \) units are connected in series, with each unit receiving as input the output of the previous unit, the net concentration achieved by the entire cascade is

\[
-(D_2/D_1)^n \text{. Such cascades are commonplace in gaseous diffusion isotope separation and other applications.}
\]

Figure 3.1 shows a 2-dimensional representation of an efficient design for a simple diffusion unit that might be used in a sortation cascade. Each unit consists of 5 chambers of equal volume, 7 dilating gates, 3 flap valves, 3 pistons, and 2 sieves which pass only water (or smaller) molecules. Each chamber is roughly cubical with \( L = 35 \, \text{nm} \) along the inside edge; including full piston throws and drives, controls, interunit piping and other support structures, each unit measures \(-125 \, \text{nm} \times 100 \, \text{nm} \times 80 \, \text{nm} \) or \(-0.001 \, \text{micron}^3 \) with a mass of \(-10^{-18} \, \text{kg} \).

The following is a precise description of one complete cycle of operation for each unit:

1. The cycle begins with fluid to be sorted in Chamber A, Chambers B and W full of pure water with piston W all the way out, Chambers R and D empty with pistons R and D all the way in, and all valves and gates closed.

2. Gate AB is opened for a time \( \Delta t \), then closed. For a small molecule such as urea (\( \text{MW} = 60 \, \text{daltons} \)), \( \Delta t = 1 \, \text{microsec} \) for a large molecule such as the enzyme urease (\( \text{MW} = 482,700 \, \text{daltons} \)), \( \Delta t = 35 \, \text{microsec} \).

3. Valves AI- and AI+, and gate AR, are opened. Piston R is drawn fully out, slowly to preserve laminar flow and to prevent mixing. Fluid in Chamber A is drawn into Chamber R. Fluid passing through the DO gate of the previous unit in the cascade enters Chamber A through valve AI-. Fluid passing through the RO valve of the subsequent unit in the cascade enters Chamber A through valve AI+. All valves and gates are closed. Chamber A is now ready for the next cycle.

4. Gates WB and BD are opened. Piston W is slowly pushed all the way in while piston D is slowly pulled all the way out. Concentrated solution in Chamber B is transferred into Chamber D as pure water in Chamber W is transferred into Chamber B, again preserving laminar flow. Both gates are closed; Chamber B is now ready for the next cycle.

5. Gates RW and DW are opened. Pistons R and D are slowly and simultaneously pushed halfway in while piston W is pulled all the way out. Forced at high pressure (~160 atm) through ~0.3 nm diameter sieve pores (Section 3.3.1), half of the solvent water present in Chambers R and D is pushed into Chamber W, filling Chamber W with water. (This design allows for easy backflushing if sieve pores become clogged.) Both gates are closed; Chamber W is now ready for the next cycle.

<table>
<thead>
<tr>
<th>Diffusion Distance</th>
<th>Diffusion Time (sec)</th>
<th>Mean Velocity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 nm</td>
<td>(-10^{-9})</td>
<td>1 m/sec</td>
</tr>
<tr>
<td>10 nm</td>
<td>(-10^{-7})</td>
<td>100 mm/sec</td>
</tr>
<tr>
<td>100 nm</td>
<td>(-10^{-5})</td>
<td>10 mm/sec</td>
</tr>
<tr>
<td>1 micron</td>
<td>(-10^{-4})</td>
<td>1 mm/sec</td>
</tr>
<tr>
<td>10 microns</td>
<td>(-10^{-4})</td>
<td>100 microns/sec</td>
</tr>
<tr>
<td>100 microns</td>
<td>10</td>
<td>100 microns/sec</td>
</tr>
<tr>
<td>1 mm</td>
<td>1000 (17 min)</td>
<td>1 micron/sec</td>
</tr>
<tr>
<td>1 cm</td>
<td>(-10^{3}) (28 hr)</td>
<td>0.1 micron/sec</td>
</tr>
</tbody>
</table>

---

Table 3.4 Estimated Time for Diffusion to Reach 90% Completion for Glycine in Aqueous Solution at 310 K.\(^{397}\)
6. Valve RO and gate DO are opened. Pistons R and D are slowly and simultaneously pushed the rest of the way in. Concentrated return fluid passes through valve RO and back to the AI+ input port of the previous unit in the cascade for further extraction. Concentrated diffusant fluid passes through gate DO and on to the AI- input port of the subsequent unit in the cascade for further purification. The valve and gate are closed; Chambers R and D are now empty and ready for the next cycle.

7. Return to Step (1). Adjacent units operate in counterphase while previous and subsequent units operate in synchrony, in a two-phase system.

Increasingly purified sample passes through a multi-unit sortation cascade as described above. For small molecules, a cascade of n ~1000 units (total device volume ~1 micron$^3$) completely resolves two mixed molecular species with $D_2/D_1 = 1.01$. As a crude approximation, $D \approx 1 / MW^{1/3}$ for small spherical particles, so this cascade separates small molecules differing by the mass of one hydrogen atom, which should be sufficient for most purposes. Structural isomeric forms of the same molecule, such as $\alpha$-alanine and $\beta$-alanine, often have slightly different diffusion coefficients, thus are also easily separable using a diffusion cascade. However, stereoisomeric (chiral) forms cannot be sorted by diffusion through an optically inactive solvent like water.

For large molecules, a 1 million-unit cascade (total device volume ~1000 micron$^3$) provides $D_2/D_1 \approx 1.00001$, sufficient to completely separate large molecules differing by the mass of a single carbon atom. The fidelity of such fine resolutions depends strongly upon the ability to hold constant the temperature of the chamber, since $D$ varies directly with temperature (Eqn. 3.5). Device temperature stability will be determined by at least three factors: (1) the accuracy of onboard thermal sensors in measuring $T$ ($\Delta T/T < 10^{-6}$; Section 4.6.1), (2) the rapidity with which the temperature measurement can be taken ($10^{-9}$ to $10^{-6}$ sec; Section 4.6.1), and (3) the time that elapses between the temperature measurement and the end of the diffusive sortation process (which may be of the same order as the gate closing time, $10^{-6}$ sec).

Most of the waste heat is generated in this device by forced water sieving (Section 3.3.1). To remain within biocompatible thermogenic limits (~$10^5$ watts/m$^3$), each unit may be cycled once every ~3 milliseconds, a 0.8% duty cycle of a ~23 microsec sieving stroke. Subject to this restriction, each unit would consume ~1 picowatt in continuous operation. A unit presented with a ~0.1 M concentration of small molecules processes ~10$^6$ molecules/sec (e.g., ~1 gm/hour of glucose using 1 cm$^3$ of n = 1000-unit cascades), or ~10$^4$ molecules/sec for a unit presented with large molecules at ~0.001 M, circulating ~10$^9$ molecules/sec of water as working fluid while running at 340 cycles/sec. Additional chamber segments on each unit, combined with more complex diffusion circuits among the many units in a cascade, should permit the simultaneous complete fractionation of the input feedstock even if hundreds of distinct molecular species are present.

By 1998, diffusion-based separation had been demonstrated in microfluidic devices.2689
3.2.5 Nanocentrifugal Sortation

Nanoscale centrifuges offer yet another method for rapid molecular sortation, by biasing diffusive forces with a strong external field. The well-known effect of gravitational acceleration on spherical particles suspended in a fluid is described by Stokes’ Law for sedimentation:

\[ v_t = 2 g R^2 \left( \frac{\rho_{\text{particle}} - \rho_{\text{fluid}}}{\eta} \right) \]

where \( v_t \) is terminal velocity, \( g \) is the acceleration of gravity (9.81 m/sec^2), \( R \) is particle radius, \( \rho_{\text{particle}} \) and \( \rho_{\text{fluid}} \) are the particle and fluid densities (kg/m^3), and \( \eta \) is coefficient of viscosity of the fluid. Particles which are more dense than the suspending liquid tend to fall. Those which are less dense tend to rise (\( \rho_{\text{particle}} / \rho_{\text{fluid}} \) -0.8 for lipids, up to -1.5 for proteins, and -1.6 for carbohydrates in water).

This separation process may be greatly enhanced by rapidly spinning the mixed-molecule sample in a nanocentrifuge device. For ideal solutions (e.g., obeying Raoult’s law) at equilibrium:

\[ \frac{c_2}{c_1} = \exp \left\{ \frac{MW_{\text{kg}}}{2 R_g T} \left( 1 - \frac{\rho_{\text{fluid}}}{\rho_{\text{particle}}} \right) \frac{t^2}{2} \right\} \]

where \( c_2 \) is the concentration at distance \( r_2 \) from the axis of a spinning centrifuge (molecules/m^3), \( c_1 \) is the concentration at distance \( r_1 \) (nearer the axis), \( MW_{\text{kg}} \) is the molecular weight of the desired molecule in kg/mole, \( \omega \) is the angular velocity of the vessel (rad/sec), \( T \) is temperature (K) and the universal gas constant \( R_g = 8.31 \) joule/mole-K. The approximate spinning time \( t \) required to reach equilibrium is

\[ t_s = \ln\left(\frac{r_2}{r_1}\right) / \left(\omega^2 S_d\right) \]

where \( S_d \) is the sedimentation coefficient, usually given in units of \( 10^{-13} \) sec or svedbergs (Table 3.5). Research ultracentrifuges have reached accelerations of ~10^9 g’s.

Consider a cylindrical diamondoid vessel of density \( \rho_{\text{vessel}} = 3510 \) kg/m^3, radius \( r_v = 200 \) nm, height \( h = 100 \) nm, and wall thickness \( x_{\text{wall}} = 10 \) nm, securely attached to an axial drive shaft of radius \( r_a = 50 \) nm (schematic in Fig. 3.2). A fluid sample containing desired molecules enters the vessel through a hollow conduit in the drive shaft, and the device is rapidly spun. If rim speed \( v_r = 1000 \) m/sec (max), then \( \omega = v_r / r_v = 5 \times 10^9 \) rad/sec (\( \omega / 2 \pi = 8 \times 10^8 \) rev/sec). The maximum bursting force \( F_b = 0.5 \rho_{\text{vessel}} v_r^2 = 2 \times 10^9 \) N/m^2, well below the 50 x 10^9 N/m^2 diamondoid tensile strength conservatively assumed by Drexler.\(^{10}\) Since \( S_d \) ranges from 0.1-200 x 10^{-13} sec for most particles of nanomedical interest (Table 3.5), minimum separation time using acceleration \( a_{\text{g}} = g \times v_r^2 / g \times r_v = 5 \times 10^{11} \) g/s, when \( r_2 = r_v \) and \( r_1 = r_a \), is \( t_s = 0.003-6.0 \times 10^{-9} \) sec. Fluid sample components migrate at ~0.1 m/sec.

Maximum centrifugation energy per particle \( E_c = (MW_{\text{kg}} / N_A) a_{\text{g}} (r_v - r_a) = 10,000 \) zJ/molecule, or ~10 zJ/bond for proteins, well below the 180-1800 zJ/bond range for covalent chemical bonds (Section 3.5.1). However, operating the nanocentrifuge at peak speed may disrupt the weakest noncovalent bonds (including hydrophobic, hydrogen, and van der Waals) which range from 4-50 zJ/bond. The nanocentrifuge has mass ~10^{-17} kg, requires ~3 picojoules to spin up to speed (bearing drag consumes ~10 nanowatts of power, and fluid drag through the internal plumbing contributes another ~5 nanowatts), completes each separation cycle in ~10^4 revs (~10^{-5} sec), and processes ~300 micron^3/sec which is ~10^{13} small molecules/sec (at 1% input concentration) or ~10^9 large molecules/sec (at 0.1% input concentration).

From Eqn. 3.11, the nanocentrifuge separates salt from seawater with \( c_2/c_1 = 300 \) across the width of the vessel (\( r_v - r_a = 150 \) nm); extracting glucose from water at 310 K, \( c_2/c_1 = 10^5 \) over 150 nm. For proteins with \( \rho_{\text{particle}} = 1500 \) kg/m^3, separation product removal ports may be spaced, say, 10 nm apart along the vessel radius while maintaining \( c_2/c_1 = 10^3 \) between each port. Vacuum isolation of the unit in an isothermal environment and operation in continuous-flow mode could permit exchange of contents while the vessel is still moving, sharply reducing remixing, vibrations, and thermal convection currents between product layers. A complete design specification of product removal ports, batch and continuous flow protocols, compression profiles, etc. is beyond the scope of this book.

Variable gradient density centrifugation may be used to trap molecules of a specific density in a specific zone for subsequent harvesting, allowing recovery of each molecular species from complex mixtures of substances that are close in density. The traditional method is a series of stratified layers of sucrose or cesium chloride solutions that increase in density from the top to the bottom of the tube. A continuous density gradient may also be used, with the.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Sed. Coeff. (sec)</th>
<th>Mol. Wt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>O_2</td>
<td>0.12 x 10^{-13}</td>
<td>32</td>
</tr>
<tr>
<td>CO_2</td>
<td>0.07 x 10^{-13}</td>
<td>44</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.18 x 10^{-13}</td>
<td>180</td>
</tr>
<tr>
<td>Insulin monomer</td>
<td>1.5 x 10^{-13}</td>
<td>6,000</td>
</tr>
<tr>
<td>Ribonuclease</td>
<td>1.75 x 10^{-13}</td>
<td>13,683</td>
</tr>
<tr>
<td>Lysozyme</td>
<td>2.03 x 10^{-13}</td>
<td>17,200</td>
</tr>
<tr>
<td>Insulin</td>
<td>1.86 x 10^{-13}</td>
<td>24,430</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>3.4 x 10^{-13}</td>
<td>43,500</td>
</tr>
<tr>
<td>Serum albumin</td>
<td>4.3 x 10^{-13}</td>
<td>68,460</td>
</tr>
<tr>
<td>Alcohol dehydrogenase</td>
<td>7.2 x 10^{-13}</td>
<td>150,000</td>
</tr>
<tr>
<td>Catalase</td>
<td>10.7 x 10^{-13}</td>
<td>250,000</td>
</tr>
<tr>
<td>β-Lipoprotein</td>
<td>5.6 x 10^{-13}</td>
<td>2,663,000</td>
</tr>
<tr>
<td>Actomycin</td>
<td>11.3 x 10^{-13}</td>
<td>3,900,000</td>
</tr>
<tr>
<td>TMV</td>
<td>175 x 10^{-13}</td>
<td>31,340,000</td>
</tr>
</tbody>
</table>

Fig. 3.2. Schematic of teragravity nanocentrifuge.
density of the suspension fluid calibrated by physical compression. For example, the coefficient of isothermal compressibility \( \kappa = - (\Delta V/V) / \Delta P \) is \( \Delta P = (\Delta \rho_{\text{fluid}} / \rho_{\text{fluid}}) / \Delta V = 4.492 \times 10^2 \text{ atm}^{-1} \) for water at 1 atm and 310 K (compressibility is pressure- and temperature-dependent). Applying \( P_1 = 12,000 \text{ atm} \) to the vessel raises fluid density to 1250 kg/m\(^3\), sufficient to partially regulate protein zoning. A multistage cascade (Section 3.2.4) may be necessary for complete compositional separation. Protein denaturation between 5000-15,000 atm\(^{485}\) due to hydrogen bond disruption may limit nanocentrifugation rotational velocity. Protein compressibility may further reduce separability. The balance between the differential densities and the differential compressibilities will determine the equilibrium radius of the protein in the centrifuge; in the limiting case of equal compressibilities for a given target protein and water, there is no stable equilibrium radius.

The nanocentrifuge may also be useful in isotopic separations. For a \( \text{D}_2 \text{O}/\text{H}_2 \text{O} \) mixture, \( c_2/c_1 = 1.415 \) per pass through the device; \( c_2/c_1 = 10^6 \) is achieved in a 40-unit cascade. Tracer glycine containing one atom of C\(^{14} \) is separated from natural glycine using a 113-unit cascade, achieving \( c_2/c_1 = 10^6 \).

### 3.3 Membrane Filtration

Filtration through a permeable membrane is closely related to the process of diffusion, since in both cases random molecular motions help carry the process to completion. However, the presence of a membrane adds a new measure of control that is not exploited in simple diffusive transport. This control may be either passive or active, unidirectional or bidirectional, as described below.

#### 3.3.1 Simple Nanosieving

Nanometer-scale isoporous molecular sieves (with ovoid, square, or hexagonal holes) are common in almost every taxonomic group of eubacteria and archaeobacteria.\(^{525}\) Other well-known examples of nanoporous structures are the 6-nm pore arrays found in reverse osmosis and kidney dialysis membranes.

Likewise, it is possible for a nanodevice to sort molecules by simple sieving.\(^{987,1177}\) In this process, a sample containing particles of various sizes suspended in water passes through a graduated series of filters perforated by progressively smaller holes of fixed size and shape. Between each filtration unit, the filtration residue consists almost exclusively of particles having a narrow range of sizes and shapes. For example, a series of \( n = 100 \) filtration units could reliably differentiate an input sample containing particles from 0.2-1.2 nm into 100 separate fractions, each fraction consisting predominantly of particles differing in mean diameter by \( \pm 0.01 \text{ nm} \). (In a practical system, several passes would be required to achieve complete discrimination; Section 3.2.4.) A \( \pm 0.01 \text{ nm} \) difference in molecular diameter corresponds to the mean contribution of \( \pm 0.01 \text{ additional carbon atom to the size of a small molecule (MW = 100 daltons), or to the mean contribution of \( \pm 0.10 \text{ additional carbon atoms to the size of a large molecule (MW = 100,000 daltons, -17,000 atoms).} \) A nanomembrane might even permit the (slow, multi-pass) sieving of oxygen from air, since the molecular diameters of \( \text{N}_2 \) and \( \text{O}_2 \) differ by \( \pm 0.01 \text{ nm} \). (Section 3.5.5.)

Two opposing forces are at work when moving water and solutes through a membrane. One is the osmotic pressure established by the presence of nonpermeating solutes; the other is the hydraulic or fluid pressure. The velocity of material movement depends on the relative values of the osmotic and hydraulic forces, and on the size of the pores in the filter.

Osmotic pressure \( P_\text{x} \) is given by the Donnan-van't Hoff formula:\(^{403}\)

\[
P_x = \left( \frac{R \times T \times (c_2 - c_1)}{\text{MW}_{\text{kg}}} \right) \left( 1 + \frac{Z^2 (c_2 - c_1)}{c_1} \right) \text{ (N/m}^2\text{)}
\]

where \( P_x = 8.31 \text{ joules/mole-K, } T \) is temperature in kelvins (K), \( c_2 \) and \( c_1 \) are solute concentrations on either side of the membrane in kg/m\(^3\) \((c_2 > c_1)\), \( \text{MW}_{\text{kg}} \) is the molecular weight of the solute in kg/mole, and the \( Z^2 \) term (dependent upon polymer-polymer interactions) is a correction factor for highly concentrated solutions which are not solvents and temperatures may equal zero; \( Z = \text{net solute charge number and } c_1 \) is the concentration in kg/m\(^3\) of a second solute, as for example when the first solute is a protein and the second solute is salt, as in human serum. Water at 310 K dissolves a maximum of \( c_2 = 370 \text{ kg/m}^3 \) of sodium chloride (a 37.0% solution, by weight), and \( \text{MW}_{\text{kg}} = 0.05844 \text{ kg/mole for NaCl, so for salt water solutions the theoretical maximum } P_x = 1.6 \times 10^7 \text{ N/m}^2 \) is 160 atm. Natural bloodstream concentrations of salt produce \( P_x = 3 \text{ atm} \). Since osmotic pressure depends on the number of molecules present, the contribution from large molecules is usually negligible. For instance, total protein concentration in human blood serum is \( c_2 = 73 \text{ kg/m}^3 \), \( \text{MW}_{\text{kg}} = 50 \text{ kg/mole} \) (50,000 daltons), so \( P_x = 0.04 \text{ atm} \).

In theory, extremely large hydraulic counterforces up to \( 10^5 \text{ atm} \) may be applied in nanomechanical systems, for example by a piston, to overcome osmotic backpressure. As a practical matter, however, rapidly pushing small molecules at high pressure through nanoscale holes is an effective method for generating significant amounts of waste heat; a design compromise is required.

Consider a simple sorting apparatus in which a square piston is used to compress solvent fluid (say, water) trapped in a chamber \( \text{h}_{\text{chamber}} \) in length and \( \text{L}_{\text{chamber}} \) in cross-sectional area, forcing the fluid to filter through pores of radius \( r_{\text{pore}} \) (target molecule radius) covering a fraction \( \alpha_{\text{filt}} \) (50%) of the surface of a square nanoscale sieve of thickness \( h_{\text{sieve}} \) and area \( \text{L}_{\text{sieve}}^2 \). Solute which is dissolved or sieved at \( \text{h}_{\text{sieve}} \) increases the linear dimension of the molecule by \( \pm 0.1-1\% \).)

The first design constraint on this system relates to its maximum operating pressure. If \( \Delta P \) is applied pressure (N/m\(^2\)), then avoiding boiling the solvent water and denaturing proteins requires at least that \( \Delta P_{\text{max}} < C_V \Delta T_{\text{boil}} \), where the heat capacity of water \( C_V = 4.19 \times 10^3 \text{ joules/m}^3\text{-K} \) and \( \Delta T_{\text{boil}} = 373 \text{ K} - 310 \text{ K} = 63 \text{ K} \) give \( \Delta P_{\text{max}} < 2600 \text{ atm} \). The designs presented below operate at 6% of this maximum (\( \Delta T = 3 \text{ K} \)) or less.

There are two major design constraints on the duration of the power stroke, or \( t_p \).

1. **Molecular Rotation Constraint** — Flow through the sieve must be slow enough to allow molecules to align with the holes. Assuming round pores, the total number of pores in the sieve is \( N_{\text{pore}} = \alpha_{\text{filt}} \text{L}_{\text{sieve}}^2 / \pi r_{\text{pore}}^2 \). The volume processing rate (m\(^3\)/sec) of the sieve \( \dot{V}_{\text{sieve}} = \dot{V}_{\text{chamber}} = \text{h}_{\text{chamber}} \text{L}_{\text{chamber}}^2 / t_p \). Hence the molecule processing rate is \( \dot{N} = \dot{V}_{\text{sieve}} / 2 \text{ pore molecules/sec} \) where \( \epsilon_{\text{target}} \) is the concentration of target molecules (molecules/m\(^3\)). At any one time, each pore channel through the sieve can hold at most \( N_{\text{channel}} = \text{h}_{\text{sieve}} / 2 \text{ pore molecules in single file; during each power stroke, at most } N_{\text{stack}} = \dot{N} \frac{t_p}{N_{\text{pore}}} \text{ molecules pass through each pore. Hence the time available for molecular rotation} \ t_r = t_p / (N_{\text{channel}} / N_{\text{stack}}) \) which assumes the layer of rotating target molecules in the vicinity of the pores approximates sieve thickness \( h_{\text{sieve}} \) a reasonable assumption as long as the typical molecular diffusion time (across
a distance \( h_{\text{sieve}} \ll r_{\text{rot}} \). Taking \( N_{\text{rot}} \approx 10 \) as the mean number of molecular revolutions needed to ensure proper pore alignment with noncircular sieve holes (the most difficult case), then, from Eqn. 3.2, \( \Delta \alpha = \left( kT \frac{r_{\text{rot}}}{4 \pi \eta r_{\text{pore}}} \right)^{1/2} \approx 2 \pi N_{\text{rot}} \), where \( \eta \) is solvent viscosity \((1.1 \times 10^{-3} \text{ kg/m-sec for plasma})\) at \( T = 310 \text{ K} \). Solving for minimum \( t_p \) gives:

\[
\frac{1}{t_p} = \frac{32 \pi^4 N_{\text{rot}}^2 \eta c_{\text{target}} h_{\text{chamber}}^2 r_{\text{pore}}^6}{kT \alpha_{\text{chamber}} h_{\text{sieve}}} \quad \text{(sec)} \quad \text{[Eqn. 3.14]}
\]

2. Pressure/Flow Constraint — Flow through the sieve must be fast enough to establish a sufficient pressure to oppose osmotic backflows. From the Hagen-Poiseuille law (Section 9.2.5), the volume processing rate through each pore is \( V_{\text{pore}} = \pi \frac{r_{\text{pore}}^4}{8 \eta} \Delta P \) and the volume processing rate through the entire sieve is \( V_{\text{sieve}} = N_{\text{pore}} V_{\text{pore}} = V_{\text{chamber}} \); solving for maximum \( t_p \) gives:

\[
\frac{1}{t_p} = \frac{8 \eta h_{\text{chamber}}^2 h_{\text{sieve}}^2}{\alpha_{\text{chamber}} N_{\text{rot}}^2 r_{\text{pore}}^2 h_{\text{sieve}}} \quad \text{(sec)} \quad \text{[Eqn. 3.15]}
\]

Equating these two bracketing constraints, \( h_{\text{sieve}} \geq 150 \text{ nm} \) for large molecules \((r_{\text{pore}} \approx 5 \text{ nm})\) but \( h_{\text{sieve}} \leq 1 \text{ nm} \) for small molecules \((r_{\text{pore}} \approx 0.32 \text{ nm})\).

As a final constraint, power released by fluid flow through the chamber and sieve must not exceed safe thermogenic limits. Given the maximum safe power density for in vivo nanomachines given in Section 6.5.3 as \( D_n = 10^9 \text{ watts/m}^3 \), then:

\[
D_{\text{device}} = \frac{P_{\text{device}}}{(h_{\text{chamber}} + h_{\text{sieve}})^2} \leq D_n \quad \text{(watts/m}^3\text{)} \quad \text{[Eqn. 3.16]}
\]

where \( D_{\text{device}} \) = device power density \((\text{watts/m}^3)\), total device power \( P_{\text{device}} = P_{\text{chamber}} + P_{\text{sieve}} \text{ (watts)} \), chamber fluid flow power \( P_{\text{chamber}} = \pi \frac{r_{\text{chamber}}^4}{128} h_{\text{chamber}} \eta \), sieve fluid flow power \( P_{\text{sieve}} = \alpha_{\text{sieve}} \frac{r_{\text{sieve}}^2}{2} h_{\text{sieve}} \Delta P_{\text{sieve}} / 8 \eta \), and \( \Delta P_{\text{sieve}} = 16 \eta h_{\text{chamber}} r_{\text{pore}}^2 \Delta P_{\text{chamber}} / \pi h_{\text{sieve}}^2 \). For small molecules such as NaCl or glucose, \( \Delta P_{\text{sieve}} \approx 160 \text{ atm} \) to overcome maximum osmotic backpressure; for large molecules such as -3 nm and small molecules diffuse -17 nm in \( 10^{-7} \text{ sec} \), just fast enough to clear the hole, so a -10 MHz sawtooth pressure profile imposed on the power stroke may ensure sufficient backflushing action to avoid serious blockages. To reduce the possibility of clogging due to surface force adhesion (Section 9.2.3), as a design criterion the work of adhesion should be reduced to \( W_{\text{adhesion}} < \Delta P_{\text{sieve}} t_p \) for small molecules and \(-5 \times 10^3 \text{ J/m}^2 \) for large molecules to come into contact with sieve pore surfaces. Clogging due to long-term random polymerizations can be minimized by periodically exchanging the entire contents of the input chamber with fresh solution, by operating the device at reduced power density, or by periodically replacing the sieve.

### 3.3.2 Dynamic Pore Sizing

A more efficient nanosieve system can be designed if pore size and shape can be actively modified during device operation, as for example by exchanging filters (from a membrane library stock- ing various pore sizes) each half cycle. Better, if pores can be reliably dilated or constricted in place during a period of time \( t_p \), then filtration cascades can be more rapidly reconfigured to match changing input feedstock characteristics or to extract varying selections of desired molecules at will. Additionally, fully differentiating sieving cascades can be collapsed into a single unit, providing more compact devices especially useful in chemical sensor systems requiring preconcentration of sample (Section 4.2.1). Control of pore shape should also provide finer discrimination among molecules of similar size but different shape, such as some isomers of nonchiral molecules.

Two or more overlapping surfaces containing regular arrays of perforations of fixed size and shape can conveniently generate a wide variety of pore geometries. Control of pore geometry is achieved by sliding or rotating one surface relative to the other by a small increment, as suggested schematically by the examples in Figure 3.3.

Circular dilating apertures can also be constructed using a matched set of overlapping segments, which may be driven either radially or tangentially to enlarge or contract the hole like an irisng camera diaphragm (Fig. 3.4), consistent with Akey’s model of the nuclear pore complex present on the cell nucleus surface. Diaphragming mechanisms may be vertically staggered to maximize areal hole density in filtration surfaces (at the cost of increased vertical rugosity). Filters constructed of hydrogen-passivated...
diamondoid can have pores with <0.1 nm feature sizes, although H-free fullerene materials might avoid any possibility of dehydrogenation shearing.

Methods of positioning surfaces to accuracies of ~0.01 atomic diameter (~0.001 nm) are discussed in Section 3.5.6. Assuming pore sizing blades require ~25 nm² of diamondoid contact surface per pore and each blade travels 25 nm at 0.01 meter/sec during one cycle, sliding friction dissipates ~0.01 zJ/pore, or ~4 x 10⁻¹⁸ watts/pore during each 2.5 microsec resizing cycle. Since fluid friction approaches kT for nanometer-size holes changing size in ~10⁻⁸ sec, maximum blade speed is ~1 m/sec and the fastest resizing cycle is ~10⁻⁸ sec.

A single sieving unit with controllable pores can be moderately efficient. Consider a design similar to that described in Section 3.3.1, except for a separate chamber and piston on either side of the filter block. Suppose that the particles desired to be extracted are of radius r, and the next smallest possible pore size is r - ∆r. The device operates in two phases. In the first phase, the sample is placed in the first chamber, pore size is set equal to r, then the first piston forces the fluid through the membrane. Particles larger than r remain behind and are flushed from the first chamber. The pores then contract to r - ∆r, and the second piston pushes the remaining filtrate back into the first chamber. After this second phase, particles of radius -(r ± ∆r) remain in the second chamber at significantly higher concentration and may be removed for further use. Analogous double-sieve editing paradigms are commonplace in biological systems.

Other designs might work equally well, such as a 3-chamber flowthrough design using a variable pore membrane with pores of size r between the first and second chamber, a membrane with pores of size r - ∆r between the second and third chamber, and a piston at either end (one pushing, one pulling), thus concentrating molecules of size r ± ∆r in the central chamber. M. Krummenacker suggests fixed chambers with a moving sieve operated as a dragnet. Filtration processes may be most useful in performing complete separations of complex mixtures. But they are inefficient in the sense that the energy expended to orient molecules passing through pores is wasted if the molecules are allowed to randomize on the other side; a eutactic mill-like molecule handling system (Section 3.4.3) might preserve this order and greatly improve energy efficiency.

3.3.3 Gated Channels

Besides controlling nanopore size and shape, individual molecular transport channels can be gated either mechanically (e.g., ligand gating) or electrically (e.g., voltage gating). Either method might usefully be employed to control molecular transport through the surfaces of medical nanodevices in a process that could very loosely be described as molecular transistor gating.

A good example of mechanical gating in biology is the nicotinic acetylcholine receptor channel, probably the best understood ligand-gated channel. Nerve impulses are communicated across neuromuscular junctions and autonomic ganglia via neurotransmitters such as acetylcholine. STM images confirm that the receptor itself is cylindrical, a bundle of 5 rod-shaped polypeptide subunits arranged like barrel staves with outside diameter ~6.5 nm. The receptor protrudes 6 nm on the synaptic side of the membrane and 2 nm on the cytoplasmic side. The water-filled channel pore lies along the symmetry axis, lined by 5 α-helices, with a 2.2 nm wide mouth on the synaptic side, a 0.65 nm waist where the structure dives through the cell membrane, and a 2 nm wide cytosolic exit.

Normally, the channel is closed and no ions may pass. In this closed state, the channel is occluded at the waist by a ridge of large residues forming a tight hydrophobic ring. Each subunit has a bulky leucine at the bend in the α-helix, a critical position. When two acetylcholine molecules bind to the receptor, these helices allosterically tilt, shifting the position of the ridges. The pore becomes open because it is now lined with small polar residues rather than by large hydrophobic ones. This conformational change allows 2.5 x 10⁵ Na⁺ ions/sec to flow through the channel, about 10% of the diffusion-limited rate. (Anions like Cl⁻ cannot enter the pore because they are repelled by rings of negatively charged residues positioned at either end of the receptor.)

Acetylcholine binding opens the gate in less than 100 microsec under physiologic conditions. Subsequent rapid destruction of acetylcholine by acetylcholinesterase, an enzyme tethered to the membrane surface by a covalently attached glycolipid group, closes the gate in ~1 millisecond. Much faster gating action (~10⁻⁸ - 10⁻⁶ sec) could be achieved by nanodevices operating variable-scale nanopores (Section 3.3.2) in response to sensor data or other control signals. Such signals could drive the insertion or retraction of diamondoid rods, wedges, or trapdoors across the channel lumen to regulate the transmission of molecules having specific sizes, shapes, and charge distributions.

Transport channels through nanodevice surfaces may also be gated electrically. In contrast to the acetylcholine receptor, which is relatively nondiscriminating and allows both inorganic and organic cations to pass, the voltage-gated calcium channel has a highly discriminating mechanism with a Ca²⁺:Na⁺ permeation ratio on the order of 1000:1. (The high specificity of the voltage-gated Ca²⁺ channel is a consequence of a single-file pore mechanism involving a pair of specific Ca²⁺ ion binding sites. Selectivity is assured if either of the two sites is occupied by Ca²⁺, as monovalent ions do not bind strongly enough to the free site or generate sufficient electrostatic repulsion to push the first Ca²⁺ ion through the channel.) Potassium channel are 100 times more permeable to K⁺ than to Na⁺, and sodium channels favor the passage of Na⁺ over K⁺ by a factor of 12. All three of these voltage-gated channels are important in the generation and conduction of neural action potentials.

A nerve impulse is an electrical signal produced by the flow of ions across the plasma membrane of a neuron. Neuron interiors
have high concentrations of K⁺ and low concentrations of Na⁺. The resting potential of a neuron is ~60 mV. An action potential may be generated when the membrane potential is slightly depolarized to ~40 mV. This opens the Na⁺ voltage-gated channels, rapidly accelerating depolarization to a peak of ~30 mV in ~1 millisecond. Then Na⁺ channels close and K⁺ channels open, allowing K⁺ ions to exit the cell, restoring the ~60 mV resting potential. Only ~1 ion of every ~10⁶ Na⁺ and K⁺ ions present in the local extracellular medium and the axoplasm participate in each such nerve impulse.

The sodium channel is a single polypeptide chain with four repeating units. Each repeating unit folds into six transmembrane alpha helices, including one that is positively charged called the S4 helix. The S4 helix is the voltage sensor that triggers the opening of the gate. Three positively charged residues on each S4 helix are paired at the resting membrane potential with negative charges on other transmembrane helices in a staircase geometry. The initial small depolarization event produces a spiral motion of each S4 accompanied by the net movement of one or two charges to the extracellular side of the membrane, essentially turning this left-handedhydrogen-bonded “molecular screw” through a ~60˚ rotation. 

This outward 0.5-nm translation of the four S4 segments opens the sodium gate by removing a steric barrier to ion flow. The energy cost of moving ~6 electrical charges (~10⁻¹⁸ coul) from the cytosolic to the extracellular side of the membrane against a ~100 mV potential (thus opening the gate in ~25 microsec) is ~100 zJ. Quantum tunneling activation of sodium channels, taking 1-1000 microsec, has been analyzed by Chancy. 

Artificial ion-gated polymer membranes were reported in 1982, and protein engineering of switchable pore-forming proteins is well-known, and “intelligent gels” are being developed that can change size and molecular porosity in response to chemical, electrical or thermal stimuli. In 1998, however, electroporation was a more commonly used method in biological research and a useful technique for “transfecting” cells in genetic studies. Electroporation employs a brief intense pulse of electricity to provide a force that opens cellular pores enabling the insertion of macromolecules like DNA into cells of interest; laser pulses reduce cell loss to 10% by using a square-wave pulse to effect rapid and reversible pore formation. Artificial pH-gated 200-nm diameter nanopores and natural pH-gating of virion pores in the cowpea chlorotic mottle virus were demonstrated in 1998.

The first true voltage-gated nanomembrane was fabricated by Charles Martin and colleagues in 1995. This membrane consists of cylindrical gold nanotubules with inside diameters as small as 1.6 nm. When the tubules are positively charged, cations are excluded and only negative ions are transported through the membrane. When the membrane receives a negative voltage, only positive ions are transported through the tubules. Nanodevices may combine voltage gating with pore size and electrostatic constraints to achieve precision transport control with moderate molecular specificity at diffusion-limited throughput rates. In 1997, an ion channel switch biosensor with sub-picomolar sensitivity and natural pH-gating of virion pores in the cowpea chlorotic mottle virus were demonstrated in 1998.

3.4 Receptor-Based Transport

The most efficient of all modes of molecular transport involves receptor sites capable of recognizing and selectively binding specific molecular species. Many receptors reliably bind only a single molecular type; others, such as sugar transporters, can recognize and transport several related sugar molecule types. In nanomechanical systems, artificial binding sites of almost arbitrary size, shape, and electronic charge may be created and employed in the construction of a variety of highly efficient molecular sortation and transport devices, described below. A discussion of binding sites and receptor engineering follows in Section 3.5.

3.4.1 Transporter Pumps

While gated channels enable ions to flow rapidly through membranes in a thermodynamically downhill direction, active pumps use a source of free energy to force an uphill transport of ions or molecules. In biology the energy supply is usually ATP or photons of light; medical nanomachines can make use of vastly more diverse energy sources (Chapter 6). Actually, the term “pump” may be something of a misnomer because the action is highly specific—only one or a very small number of molecular species are selectively transported.

Molecular pumps generally operate in a four-phase sequence: (1) recognition (and binding) by the transporter of the target molecule from a variety of molecules presented to the pump in the input substrate; (2) translocation of the target molecule through the membrane, inside the transporter mechanism; (3) release of the molecule by the transporter mechanism; and (4) return of the transporter to its original condition, so that it is ready to accept another target molecule. Such molecular transporters that rely on protein conformational changes are ubiquitous in biological systems, and are illustrated schematically in Figure 3.5.

The minimum energy required to pump molecules is the change in free energy ∆G in transporting the species from one environment having concentration c₁ to a second environment having concentration c₂ given by:

\[ \Delta G = kT \ln \left( \frac{c_2}{c_1} \right) + \frac{Z_e F \Delta V}{N_A} \]

(Fig. 3.5. Schematic of transporter molecular pump operation (uniport).)

where \( k = 0.01381 \, \text{zJ/K} \) (Boltzmann constant), \( T = 310 \, \text{K} \), \( Z_e \) is the number of charges per molecule transported (i.e., the valency), \( F = 9.65 \times 10^4 \, \text{coul/mole} \) (Faraday constant), \( \Delta V \) is the potential in volts across the membrane, and \( N_A \) is Avogadro’s number. So for example, transport of an uncharged molecule across a \( c_2/c_1 = 10^3 \) gradient (typical in biology) costs \( \sim 30 \, \text{zJ} \). An extremely aggressive \( c_2/c_1 = 10^6 \) concentration gradient costs \( \sim 60 \, \text{zJ/molecule} \), plus another \( \sim 30 \, \text{zJ/ion} \) if we are moving Ca\(^{++} \) ions against a 100 mV potential. An artificial nanopump of dimension \(-10 \, \text{nm}\) moving at a conservative \(-1 \, \text{cm/sec}\) velocity operates at MHz frequencies, transporting \(-10^6 \, \text{molecules/sec}\) for a continuous power consumption of \(-0.03 \, \text{pW}\) at \( c_2/c_1 = 10^3 \). Such a pump has a mass \(-10^{-21} \, \text{kg}\).

Transporter pumps need not be limited to the movement of a single molecular species in a single direction, which biochemists call a uniport transport mechanism. Numerous well-known biological systems are capable of moving two molecules simultaneously in one direction (symport mechanisms), two molecules sequentially in opposite directions (antiport mechanisms), and charged molecules in one direction only, thus building up an electrical charge on one side of the membrane (electrogenic mechanisms). Such pumps exist in nature for numerous ions, amino acids, sugars, and other small biomolecules. Active drug efflux systems and multidrug resistance are made possible by the expression of bacterial genes coding for molecular pumps that are constantly evolving new specificities to increasing numbers of microbial drugs.

The action of the Na\(^+\) - K\(^+\) antiporter, a familiar ion pump present in all mammalian cells, is illustrated in Figure 3.6. Three Na\(^+\) and two K\(^+\) ions are transported per 10 millisec cycle, requiring the hydrolysis of one ATP molecule to ADP to drive the conformational changes. (More than one-third of the ATP consumed by a resting animal is used to pump these two ions.) Hydrolysis of ATP liberates \(-80 \, \text{zJ/molecule} \) of free energy, so the antiporter is transporting Na\(^+\) and K\(^+\) at a cost of \(-16 \, \text{zJ/ion} \) at a 0.5 KHz frequency. Pump site density is \(-1000/\text{micron}^2 \) of cell membrane in neural C fibers. (The Na\(^+\) - K\(^+\) pump can also be operated in reverse to synthesize ATP from ADP by exposing the mechanism to steep ionic gradients.) By contrast, artificial nanomechanical antiporter and symporter devices will operate at MHz frequencies. They should be able to transport much larger molecules, and may also be fully reversible.

### 3.4.2 Sorting Rotors

Drexler’s molecular sorting rotor is a related class of nanomechanical device capable of selectively binding molecules from solution and then transporting these bound molecules against concentration gradients (Fig. 3.7). The archetypal sorting rotor is a disk with 12 binding site “pockets” along the rim exposed alternately to the external solution and interior chamber by axial rotation of the disk. Other designs may have more, or fewer, pockets. Each pocket selectively binds a specific molecule when exposed to the solution. Once the binding site (Section 3.5) rotates to expose it to the interior chamber, the bound molecules are forcibly ejected by rods thrust outward by the cam surface (or using some other means by which receptor affinity can be adjusted during the inbound transport process). In the case of protein molecules, the debinding geometry must be carefully designed to avoid denaturation during ejection. Also, the rotor in Figure 3.7 implicitly assumes that target molecules remain in the liquid or gaseous state after importation. M. Krummenacker observes that most bloodborne molecular species will precipitate as solids unless they are well-solvated; thus nanomedical sorting systems may require internal solvent or in some cases should be made completely eutactic (Section 3.4.3). The discovery of positionally disordered water molecules resident inside protein hydrophobic cavities suggests that good rotor designs may also need to include solvent drainage channels.

Molecular sorting rotors can be designed from about \(-10^5 \, \text{atoms} \) (including housing and pro rata share of the drive system), measuring roughly \( 7 \, \text{nm} \times 14 \, \text{nm} \times 14 \, \text{nm} \) in size with a mass of \(-2 \times 10^{-21} \, \text{kg}\). Rotors turn at \(-86,000 \, \text{rev/sec} \) with a conservative rim speed of \(-2.7 \, \text{mm/sec} \) and an almost negligible drag power of \(-10^{-16} \, \text{watts} \) against the fluid, sorting small molecules at a rate of \(-10^6 \, \text{molecules/sec} \) with laminar flow. From Eqn. 3.18, the energy cost of small-molecule sortation at 310 K ranges from \(-10 \, \text{zJ/molecule} \) at low pressures \((c_2/c_1 = 10)\) up to \(-40 \, \text{zJ/molecule} \) when pumping against the highest head pressures \((c_2/c_1 = 10^4, \sim 30,000 \, \text{atm}) \) for natural bloodstream sortation.

**Fig. 3.6.** Schematic of sodium-potassium antiporter ion pump operation.

**Fig. 3.7.** Molecular sorting rotor (modified from Drexler).
concentrations of salt with osmotic \( p_A < 3 \) atm, consuming 0.01-0.04 pW per device in continuous operation. Rotors are fully reversible, so they can be used to load or unload target molecules depending on the direction of rotor rotation. Cylindrical rotors with many receptor rows are somewhat more energy-efficient, and rotor lifetimes\(^{10}\) should be >10^4 sec (Chapter 13).

Typical molecular concentrations in the blood for target molecules of nanomedical interest are \(-10^{-11} - 10^{-3}\) molecules/nm^3, which should be sufficient to ensure >99% occupancy of rotor binding sites (Section 3.5.2). Rotors targeting serum hormones and other low-concentration species at the parts-per-billion level must slow to <1 rev/sec to ensure complete receptor occupancy and to avoid exceeding diffusion limits.

Sorting of ionically charged species can use binding sites that display opposite charge, effectively increasing the affinity of the binding site for the charged species. Many ionic species dissolved in water are actually more complex than their symbolic representation would suggest. As an example, a naked proton (H\(^+\)) in water is always highly hydrated, usually as H\(_2\)O\(_4^+\) or H\(_2\)O\(_3^+\)\(_9\)96, or even as H\(_2\)O\(_4^+\) in strong acid solutions.\(^{323}\) Small ions from Li\(^+\) to I\(^-\) are also found in solvent cages, bound with energies \(\geq 330\) J relative to vacuum;\(^{2338}\) for instance, Li\(^+\) and I\(^-\) are coordinated to 4-6 water molecules.\(^{1149,3235}\) As a result, the design of an appropriate binding site or filtration process for such ions can likewise become more complex. The existence of biological channels that show remarkable selectivity for specific ions (e.g., Na\(^+\) channels that largely exclude complex) could ensure in micron-scale nanosystems using at most 5 stages, starting from a dilute input substrate containing only 1 part per billion of the target molecule with each stage providing a concentration factor of \(-10^{4}\). For statistically pure extractions of more common molecules in blood and cytoplasm, a 3- or 4-stage cascade will usually suffice.\(^{326}\)

Note that the optimal receptor structure may differ at different stages in a cascade,\(^{10}\) and that each 12-arm outbound rotor can contain binding sites for 12 different impurity molecules. The first-stage receptor will likely pass only a relatively small number of different contaminant species, so the number of outbound rotors in the entire system can probably be reduced to a small fraction of the number of inbound rotors.

### 3.4.3 Internal Transport Streams

After pump mechanisms described above have reliably sorted externally-encountered target molecules into reservoirs filled with species of a single type, a medical nanodevice may require these molecules to be transported to specific internal locations for further processing. Bulk fluid flow or fluidized (solvated or suspended) transport through nanpipes may suffice for some purposes (Section 9.2.5). However, in many cases it will be necessary to present reagent molecules to other subsystems as a well-ordered stream of precisely positioned moieties transported in vacuo, especially for mecanochemical operations (Chapter 19).

For this purpose, Drexler\(^{10}\) proposes molecular mills—eutectic systems of nanoscale belts moving over rollers, with reagent-binding devices mounted on the belt surface (Fig. 3.9). This class of device can be assembled into complex molecular transportation networks using conditional switching, crossed-axis belting, and transit speed/frequency multipliers,\(^{10}\) and may also be employed to drive mecanochemical and vsound = 10^4 m/sec (~speed of sound in diamond), giving P\(_{drag}\) \(-4 \times 10^{-21}\) watts per reagent device, or P\(_{drag}\) / \(v\) \(-10^{-8}\) zJ/nm traveled per reagent device (palette). Note that volume containerization of pallet-transported molecules is least efficient at the smallest scales, where surface area per unit enclosed volume is highest, since energy usage is proportional to the surface area of the carrier. Containerization (Section 9.2.7.7) of n >> 1 molecules for large-pallet transport is more efficient.

A less energy-efficient, but far more versatile, internal molecular transport device is the 100-nm telescoping manipulator arm described in Section 9.3.1.4. This flexible \(-10^{-19}\) kg device employs a binding tip to pick and place small and large molecules alike, moving them at \(-1\) cm/sec with repeatable placement accuracy of 0.04 nm. Multiple devices can be used to establish an internal ciliary transport system (Section 9.3.4); standardized volume containerization of molecules permits rapid stereotypical handoff motions and efficient parcel routing. Conveyance through a 100-nm arc takes \(10^5\) sec consuming 0.1 pW while the arm is in motion, or \(-10\) zJ/nm traveled per reagent molecule or per container transported (vs. \(-1000\) zJ per typical covalent bond).

In molecular cytobiology, vesicles and organelles are transported throughout the interior of a cell by riding on microtubular cables crisscrossing the cytosol (Section 8.5.3.11). For example, neural vesicles show transport speeds up to 2.4 microns/sec.\(^{938}\)

### 3.5 Molecular Receptor Engineering

Molecular recognition requires a detailed surface complementarity between the target molecule and its receptor. The interplay of various molecular forces between ligand and receptor causes them to selectively bind together, typically engineered to occur in \(-10^{-6}\) sec. It is useful first to briefly review and quantify the principal physical forces at work.

#### 3.5.1 Physical Forces in Molecular Recognition

Covalent bonds, which occur when atoms share electrons, are the strongest bonds. Aside from metals and salts, most material objects are made of atoms held together by covalent bonds. The atoms comprising biological molecules like proteins, nucleic acids...
and lipids are strung together mostly by single, double, or triple covalent bonds, as are receptors and diamondoid nanomechanical structures. Interatomic bond strengths range from 181 zJ/bond for O-F up to 1785 zJ/bond for C-O. Covalent bond lengths range from 0.10-0.16 nm within CHON-atom molecules, giving a typical covalent bond rupture force of ~10 nN/bond.

But as Jean-Marie Lehn points out, “there is a chemistry beyond the molecule”—noncovalent supramolecular chemistry (Section 2.3.2). The bonds employed in molecular recognition are weak noncovalent bonds. Noncovalent bonds are largely responsible for the secondary and higher order structure of macromolecules. On a per-bond basis, noncovalent bonds are 1-3 orders of magnitude weaker than covalent bonds. However, the possibility of combining within a limited area a great number of noncovalent bonds having complementary elements allows the formation of a large specific association whose affinity may be of the same order of magnitude as a covalent bond. The high combinatorial diversity provided by many complementary elements allows numerous orthogonal specific associations, enabling self-assembly of many components; by comparison, covalent chemistry offers a poor diversity of reactivities. An additional advantage is that the formation of noncovalent bonds often is not hindered by high energy barriers. At least five types of noncovalent bonds may be distinguished: electrostatic, hydrogen, van der Waals, π aromatic, and hydrophobic.

1. Electrostatic Bond — The electrostatic bond between two charged particles (e.g., the “salt bridge” in proteins) is a dipole interaction whose energy $E_e$ is given by Coulomb’s law as:

$$E_e = \frac{e^2}{4\pi\varepsilon_0} \left( \frac{Z_1Z_2}{k_BT} \right) \exp\left(-\frac{K_{dl}r}{\varepsilon_0} \right) \text{ (joules)} \quad [\text{Eqn. 3.20}]$$

where $e = 1.60 \times 10^{-19}$ coul (elementary charge), $\varepsilon_0 = 8.85 \times 10^{-12}$ farad/m (permittivity constant), $Z_1$ and $Z_2$ are the numbers of
attractant charges, \( r \) is the distance between the charges, and \( \kappa_s \) is the dielectric constant (74.3 for pure water at 310 K, usually reduced to ~40 in a hydrophobic environment). The bond is strengthened if the charges are in a hydrophobic environment. Conversely, the presence of electrolytes weakens the bond energy due to a shielding effect, given by \( K_{\text{DH}} \), the Debye-Hückel reciprocal length parameter, which has a value of 1.25 nm\(^{-1}\) for 0.15 M NaCl (~1% solution, ~human blood). Thus two unit charges separated by 0.3 nm produce an interaction energy of \( E_e = 19 \text{ zJ} \) in diffusion.\(^{2144}\) Cell surfaces may be much more efficient than predicted by random diffusion.

Most isolated amino acids in neutral solution are zwitterionic—the molecule has no overall charge but carries both a negatively charged group (carboxyl, \( \text{CO}_2^- \)) and a positively charged group (amino, \( \text{NH}_3^+ \)). In proteins the individual amino acids are polymerized, giving a peptide backbone which is electrically neutral except for the ends of the chain. Most of the standard amino acids found in proteins have uncharged side chains, although histidine, lysine and arginine each have a positive charge at neutral pH and both glutamic and aspartic acids normally carry a negative charge.

2. Hydrogen Bond — A second important noncovalent interaction is the hydrogen bond, a dipole formed when a hydrogen atom covalently bonded to an electronegative group is shared with a second electronegative group (typically an oxygen, nitrogen or fluorine atom), such that the proton may be approached very closely by an unshared pair of electrons. Hydrogen bonds are largely responsible for the unusual thermodynamic properties of water and ice, and the DNA double-helical and protein \( \alpha \)-helical and \( \beta \)-structure conformations are extensively hydrogen bonded. Higher binding energies occur when donor and acceptor atoms are 0.26-0.31 nm apart. Typical hydrogen bond strengths in proteins are 7-50 zJ.

3. Van der Waals Interaction — A third important noncovalent force is van der Waals interactions (London dispersion forces).\(^{1149}\) There is an attractive component due to the induction of complementary partial charges or dipoles in the electron density of adjacent atoms when the electron orbitals of two atoms approach to a close distance. There is also a strongly repulsive component at shorter distances, when the electron orbitals of the adjacent atoms begin to overlap, commonly called steric hindrance. The van der Waals attractive bonding energy between two parallel plates of area \( A \) and separation \( z_{\text{sep}} \) is approximated by:

\[
E_{\text{vdW}} = \frac{HA}{12 \pi z_{\text{sep}}^2} \quad \text{(Eqn. 3.21)}
\]

where the Hamaker constant \( H = 37 \text{ zJ} \) for water, 66 zJ for glycerol, 340 zJ for diamond (Table 9.1). For \( A = 0.4 \text{ nm}^2 \) (small molecule), \( z_{\text{sep}} = 0.3 \text{ nm} \), then \( E_{\text{vdW}} = 4 \text{ zJ} \) (water) to 8 zJ (small organic molecules)—close to the mean energy of a thermally excited harmonic oscillator, \( kT \approx 4.3 \text{ zJ} \) at 310 K. While van der Waals bonds are individually very weak, they are also very numerous since they involve all pairs of neighboring atoms. For example, experimental analysis of an antigen molecule trapped in the anti-hen egg-white lysozyme monoclonal antibody Fv active binding site found 86 distinct interatomic contact points with antigen-antibody separations ranging from 0.25-0.46 nm, averaging 0.36 nm.\(^{346}\) Since intermolecular dispersion forces act on all molecules, there are probably no ligands with MW > 400 daltons which cannot be receptored (Section 3.5.5). That is, van der Waals interactions ensure that virtually all molecules of nanomedical interest are theoretically bindable noncovalently.

4. Aromatic \( \pi \) Bonds — A fourth type of interaction (\( \pi \) electron to \( \pi \) electron), called “aromatic” or “\( \pi \)” bonding, occurs when two aromatic rings (conjugated \( \pi \) systems) approach each other with the plane of their aromatic rings overlapping, with successive \( \pi \)-bonded systems stacked like layers in a cake. This results in a noncovalent attractive force with a bond strength of ~40-50 zJ. (That is, the planes of conjugated \( \pi \) systems attract each other when superimposed.) \( \pi \) bond stacking forces contribute to nucleic acid stability at least as much as the hydrogen bonds between bases.\(^{401}\)

In designing an artificial binding site, the above forces may be combined to achieve the desired level of affinity and specificity for a given ligand. All forces are not equally useful in this regard, however. For example, hydrophobicity is the major factor in stabilizing protein-protein associations.\(^{402}\) But hydrophobicity is almost entirely nonspecific, hence contributes little to ligand discrimination. By contrast, the proper formation of hydrogen bonds and van der Waals contacts require complementarity of the surfaces involved. Such surfaces must be able to pack closely together, creating many contact points, and charged atoms must be properly positioned to make electrostatic bonds. Thus van der Waals and polar interactions may contribute little to the dynamic stability of the ligand-receptor complex, but they do determine which molecular structures may recognize each other.\(^{403}\) Other design elements of binding sites, such as directed channeling of substrates into the receptor,\(^{407}\) may also prove useful.

In analyzing molecular forces, note that at the nanoscale level, surface/surface, molecule/surface, and molecule/molecule interactions may feature very complicated behaviors. Nanodevices performing work may generate both thermodynamic and mechanical local nonequilibrium conditions, so calculations based on the general forms of interactions and on macroscopic expressions valid at equilibrium conditions should be taken only as basic estimates.

3.5.2 Ligand-Receptor Affinity

For a ligand binding to a receptor in a solvent, there will be a characteristic frequency with which existing ligand-receptor complexes dissociate as a result of thermal excitation, and a characteristic frequency with which empty receptors bind ligands as a result of Brownian encounters, forming new complexes, with the frequency of binding proportional to the concentration of the ligand in solution,
...ligand (molecules/nm\(^3\)). For simple processes, the equilibrium constant \(K_d\), taken in the direction of dissociation, is:

\[
K_d = k_d / k_a \quad \text{(molecules/nm}\(^3\)) \tag{Eqn. 3.22}
\]

where \(k_d\) is the dissociation rate constant (sec\(^{-1}\)) and \(k_a\) is the association rate constant (nm\(^3\)/molecule-sec). The \(k_a\) rate constant reflects mainly the molecular weight of the ligand, and thus varies little among antibody, enzyme, or other receptor systems. For example, Delaage\(^{401}\) notes that changing the solution pH for growth hormone from 7.5 to 4.0 increases \(K_d\) by a factor of 3000, due to \(k_d\) increasing by a factor of 1600 but \(k_a\) decreasing only by a factor of 1.7. Hence it is the rate constant of dissociation, \(k_d\), which accounts for the vast bulk of affinity in receptor systems.

Thus receptor affinity is usually taken as the inverse of the dissociation rate constant, which may be placed in the context of the half-life of the ligand-receptor complex approximated\(^{401}\) by:

\[
\tau_{1/2} = \ln(2) / k_d \quad \text{(sec)} \tag{Eqn. 3.23}
\]

Observed half-lives range from <0.1 microsec (\(k_d\rightarrow10^8\) sec\(^{-1}\)) for the enzyme catalase to a few months (\(k_d\rightarrow10^{-7}\) sec\(^{-1}\)) for enzyme inhibitors such as the Kunitz inhibitor of trypsin\(^{406}\) and for avidin-biotin binding.\(^{407}\) The smaller the \(k_d\) (or the \(K_d\)), the greater the affinity and so the more firmly the receptor grasps the ligand.

The probability \(P_{\text{occupied}}\) that a receptor will be occupied\(^{10}\) is given by:

\[
P_{\text{occupied}} = \frac{c_{\text{ligand}}}{K_d} P_{\text{unoccupied}} \quad \text{[Eqn. 3.24]}
\]

where \(P_{\text{unoccupied}} = 1 - P_{\text{occupied}}\). To ensure \(P_{\text{occupied}} = 99\%\) receptor occupancy, \(K_d\) must \(= c_{\text{ligand}} / 100\). For target molecules present at the \(10^{-3} - 10^{-11}\) gm/cm\(^3\) concentrations typically found in human blood (Appendix B), \(c_{\text{ligand}} = 3 \times 10^5\) molecules/nm\(^3\) for glucose to \(c_{\text{ligand}} = 10^{-11}\) molecules/nm\(^3\) for female serum testosterone, giving a range of \(K_d \sim10^{-3} - 10^{-15}\) molecules/nm\(^3\) to achieve 99\% occupancy.

How much binding energy per receptor will this require? The free energy of dissociation \(\Delta G_d\) of a ligand-receptor complex is related to its equilibrium dissociation constant \(K_d\) by:

\[
\Delta G_d = -kT \ln(K_d / K_0) \quad \text{[Eqn. 3.25]}
\]

which refers to a standard reference state where all chemical species are 1 M (i.e., \(K_0 = 0.6\) molecules/nm\(^3\)) and attributes a free energy which refers to a standard reference state where all chemical species.

For \(T = 310\) K, the range of required \(K_d\) gives a range for \(\Delta G_d\) of 39.4 zJ for glucose (at typical serum concentrations) to 128 zJ for large molecules (MW \(\sim10^6\) daltons). For T = 310 K, the range of required \(K_d\) gives a range for \(\Delta G_d\) of 39.4 zJ for glucose (at typical serum concentrations) to 128 zJ for large molecules (MW \(\sim10^6\) daltons). Thus receptor affinity is usually taken as the inverse of the dissociation rate constant, which may be placed in the context of the half-life of the ligand-receptor complex approximated\(^{401}\) by:

\[
\Delta G_d = -kT \ln(K_d / K_0) \quad \text{[Eqn. 3.25]}
\]

For \(T = 310\) K, the range of required \(K_d\) gives a range for \(\Delta G_d\) of 39.4 zJ for glucose (at typical serum concentrations) to 128 zJ for large molecules (MW \(\sim10^6\) daltons). Thus receptor affinity is usually taken as the inverse of the dissociation rate constant, which may be placed in the context of the half-life of the ligand-receptor complex approximated\(^{401}\) by:

\[
\Delta G_d = -kT \ln(K_d / K_0) \quad \text{[Eqn. 3.25]}
\]

For \(T = 310\) K, the range of required \(K_d\) gives a range for \(\Delta G_d\) of 39.4 zJ for glucose (at typical serum concentrations) to 128 zJ for large molecules (MW \(\sim10^6\) daltons).
Each increase of \(-10\) zJ in bonding energy causes the reaction equilibrium constant to decline (hence receptor affinity to rise) by a factor of \(-10\) (Eqn. 3.25). If a difference in affinities of \(-10^{-10}\) between the target molecule and its nearest competitor likely to be present in the environment provides sufficient receptor specificity for nanomedical purposes, this requirement corresponds to a binding energy differential affinity of \(-30-50\) zJ at 310 K between target and closest-neighbor ligands.

### 3.5.4 Ligand-Receptor Dynamics

Diamondoid structures can exhibit a stiffness and rigidity one or two orders of magnitude greater than that available in protein structures. In general, stiffer structures permit greater specificity because they enhance exclusion of non-target ligands based on van der Waals overlap forces (called steric hindrance) and allow narrower tolerances in distinguishing acceptable ligands.

In less-stiff protein-based receptors, each of the atoms is engaged in relatively large, rapid jiggling movements. Experimental and theoretical work has been done on the atomic fluctuations within the basic pancreatic trypsin molecule, a small enzyme with 58 amino acids and 454 heavy (non-hydrogen) atoms. This work established that fluctuations increase with distance from the center of the molecule, with the magnitude of RMS fluctuations ranging from \(-0.04\) nm for backbone atoms to \(-0.15\) nm for the ends of long side chains (roughly one atomic diameter), and an average of \(0.069-0.076\) nm per atom over the entire molecule. A similar experimental analysis of reduced cytochrome c, a common metabolic enzyme, shows that RMS fluctuations of each of the 103 amino acid residues in the molecule averages \(-0.11\) nm with lattice disorder (\(-0.05\) nm) included, and fluctuations range from \(0.09-0.16\) nm (Fig. 3.10).

The ratio of RMS displacements for protein/diamondoid receptors is \(-10:1\), so the minimum addressable volume (hence inverse maximum specificity) of a diamondoid receptor should be \(-10^3\) smaller than for protein receptors, a \(-30\) zJ binding energy advantage for diamondoid receptors.

By contrast, in stiff diamondoid-based receptors each of the atoms is locked in a rigid crystalline structure and thus is subject to thermal displacements approximately 10 times smaller. The RMS displacement for a quantum mechanical harmonic oscillator\(^{10}\) is given by:

\[
\Delta X = \left( \frac{\hbar}{k_i} \right)^{1/2} \left( \frac{1}{\omega} + \frac{1}{e^{\omega/kT} - 1} \right)^{1/2}
\]

where \(\hbar = 1.055 \times 10^{-34}\) joule-sec, \(kT = 4.28\) zJ at \(T = 310\) K, and angular frequency \(\omega = (k_i/\mu_{\text{red}})^{1/2}\) rad/sec where \(k_i\) is the mechanical stiffness and \(\mu_{\text{red}}\) is the reduced mass = \(m_1m_2/(m_1 + m_2)\). For C-C atoms (e.g., in the receptor body), \(k_i = 440\) N/m, \(m_1 = m_2 = 2\times10^{-26}\) kg, thus \(\omega = 2.1 \times 10^{14}\) rad/sec, so the RMS displacement of each atom is only \(-0.005\) nm every \(-3 \times 10^{-14}\) sec. For C-H atoms (e.g., on the hydrogen passivated receptor surface), \(k_i = 460\) N/m, \(m_1 = 2 \times 10^{-26}\) kg (C), and \(m_2 = 1.673 \times 10^{-27}\) kg (H), thus \(\omega = 5.5 \times 10^{14}\) rad/sec, so the RMS displacement of each atom is \(-0.008\) nm every \(-1 \times 10^{-14}\) sec. Similarly, at \(310\) K the RMS thermal displacement of a 1-nm wide, 10-nm long diamondoid rod is \(-0.01\) nm, including elastic and entropic contributions.\(^{19}\)

The ratio of RMS displacements for protein/diamondoid receptors is \(-10:1\), so the minimum addressable volume (hence inverse maximum specificity) of a diamondoid receptor should be \(-10^3\) smaller than for protein receptors, a \(-30\) zJ binding energy advantage for diamondoid receptors.

### 3.5.5 Diamondoid Receptor Design

Natural enzymes and antibodies are proteins folded into highly organized, preformed shapes that present a ready-made “keyhole” into which a target ligand will fit. The enzyme is folded in such a way as to create a region that has the correct molecular dimensions, the appropriate topology, and the optimal alignment of counterionic groups and hydrophobic regions to bind a specific target molecule. Tolerances in the active sites can be narrow enough to exclude one isomer of a diastereomeric pair. For example, D-amino acid oxidase will bind only D-amino acids, not L-amino acids.

These “keyholes” are extremely floppy, yet still achieve fair specificity. This is a consequence of "induced fit" in protein binding sites. That is, the interaction of the target molecule with an enzyme induces a conformational change in the enzyme, resulting in the formation of a strongly binding site and the repositioning of the appropriate amino acids to form the active site. The receptor flexes, balloons, hinges, or contracts by 0.05-1.0 nm in just the right places to maximize specificity as the selected ligand enters the site. In some cases such as \(O_2\) and \(CO\) binding by myoglobin, ligands enter the receptor through a series of temporary voids that appear and disappear in the receptor as \(-10\) picosecond dynamic structural fluctuations.\(^{409}\) Induced fit can reduce receptivity to undesired proteins that exploit relative geometry by binding enough to bring portions of their surfaces into alignment with the same receptor sites that bind desired proteins. Folding transitions appear to be the most prevalent and to possess the most possibilities for adaptability or induced fit.\(^{1068}\)

For smaller molecules, it is likely that recognition processes will be relatively inefficient, time-consuming, and more difficult to engineer if they involve a good deal of rearrangement of the receptor's shape.\(^{382}\) Thus there is considerable interest among chemists in designing artificial receptors that have their cavities already formed into the shape appropriate for the intended substrate.\(^{1057}\) For instance, rigid-cavity “spherand” receptors are exceptionally efficient at binding metal ions.\(^{410}\) Bowl-shaped molecules such as cryptands, calixarenes, and carcerands can be lined with chemical groups along their walls and with charged groups along their rims to achieve high binding specificity.\(^{410,411,1262}\)

Self-assembling capsules or “container molecules” made of hydrogen-bonded subunits, capable of limited molecular recognition, have been synthesized.\(^{2143,2356}\) A designed receptor for creatinine was demonstrated in 1995,\(^{222}\) and in 1998, K. Suslick and colleagues\(^{2714}\) designed metalloporphyrin-dendrimeric artificial receptors that can bind straight, skinny molecules but block out bent or fat molecules. Container molecules with 0.2-0.4 nm portals control entry to their interiors using ‘French door’ and ‘sliding door’ gates.\(^{417}\)
hinges. Active binding sites for small simple molecules such as NO, CO, and C$_2$H$_4$ are well-known.

Ultimately, receptors will be designed to nanoscale precision and may be constructed using diamondoid materials. Electrostatic, hydrophobic, and hydrogen-bond forces will add immensely to artificial receptor specificity and are essential for binding small molecules. For example, using a 0.2-nm range in a saline environment, 10-40 charge contacts would be required to bind molecules of various sizes and concentrations using electrostatic force alone, which is ~0.5 charge/nm$^2$ over the entire surface of a 60,000 dalton globular protein (vs. ~1 charge/nm$^2$ for the surface of an isolated zwitterionic amino acid). Or, a 7 nm$^2$ hydrophobic cavity having the exact folded shape of the target ligand generates ~120 zJ binding energy as the molecule stuffs itself into the cavity to exclude its surface from solvent water.

But consider a theoretical receptor that employs van der Waals dispersion forces alone. Atoms comprising the typical protein or CHON target molecule in the human body have an average atomic mass of ~6 amu/atom and an average density of 1500 kg/m$^3$, giving a mean molecular volume of ~6.7 x 10^{-30} m$^3$ per atom in the target molecule. Assume for simplicity a spherical receptor surface that forms a negative image of the surface of the target molecule. The receptor surface lies ~1.5 x (minimum van der Waals contact distance) ~0.3 nm from the perimeter atoms of the target molecule, and completely encloses the target molecule (thus requiring at least one moving part). Table 3.6 shows that the theoretically available maximum binding energy $E_{vdW}$, using only dispersion forces (from Eqn. 3.21), should be sufficient to adequately bind all but the smallest target molecules according to the criteria set forth in Section 3.5.2. (Proteins are actually ellipsoidal with a much larger surface area $A_p > 0.111$ MW$^{0.5}$ nm$^2$, than if they were spherical, so Table 3.6 figures are conservative for protein binding; proteins typically have a ~60%-85% interior packing density.) Dispersion forces alone can provide the minimum required binding energy of ~120 zJ with $A_p > 6$ nm$^2$ of contact surface (MW > 400 daltons) at 0.3 nm mean range, and dispersion-force receptors can offer exceptionally high affinities for molecules >1000 atoms.

**What about specificity?** Given the ability to design diamondoid binding sites to at least localized <0.01 nm tolerances, chirality is readily detected and (purely as a design exercise) it may even be possible to distinguish diatomic nitrogen and oxygen on the basis of width (~0.01 nm). With a van der Waals energy well depth of 1.1-1.4 zJ for N$_2$ and O$_2$, in a tight receptor these gases are bound at an energy density corresponding to 3000-4000 atm of pressure.

### 3.5.6 Minimum Feature Size and Positioning Accuracy

Maximum displacement measurement accuracy in nanoscale devices is ~0.01 nm (Section 4.3.1), and RMS thermal displacement in diamondoid bonds is ~0.01 nm at 310 K (Section 3.5.4). Thermal displacements in 10-nm long diamondoid rods are ~0.01 nm at a 1-nm rod width, ~0.02 nm at 0.5-nm width, and ~0.10 nm at 0.3-nm width. However, it is possible to construct components to even narrower tolerances.

For instance, a single C-O bond inserted into a diamondoid rod in a collinear carbon chain extends rod length by 0.1402 nm, the C-O bond length. An adjacent rod into which an N-N bond is similarly inserted is extended by 0.1381 nm, the N-N bond length. By bonding these rods (aligned at one end) it is possible to build diamondoid structures having 0.002-nm features (at the other end), which at 310 K will nonetheless suffer thermal displacements of ~0.01 nm or more. Similarly tiny displacements can be induced in binding cavity surfaces by inserting a foreign atom deep inside the bulk diamondoid structure, causing dislocation strains that decline in magnitude at greater distances from the compositional disturbance.

---

*The formula for $A_p$ is valid for small and medium size monomeric proteins (50-320 residues). The ratio of actual protein surface area to the surface area of a smooth ellipsoid of equal volume increases with molecular weight, since larger proteins are more highly textured and apherical in shape. For oligomers whose monomers have 330-840 residues, surface areas ~MW and are 20%-50% greater than those given by $A_p$.**

**For even finer feature control, the ground state atomic radius $R_i$ depends on nuclear mass ($m_i$), proton mass ($m_p$) and electron mass ($m_e$) through the reduced mass of the electron $\mu_e = m_e / (m_p + m_e) = m_e / (1 + 1836 m_e / m_p)$, or $R_i = 1/\mu_e$, where $A_{nuc}$ is nuclear mass number and $m/N = 1836$. Thus, a 25-nm rod consisting of 162 planes of C atoms is ~0.1 picometer (pm) longer than a 162-plane rod of C$_2$ atoms; a single ground-state deuterium atom is measured as ~0.4 pm smaller than a hydrogen atom. SQUIDs and x-ray interferometers have been used to measure displacements of 10$^{-7}$ nm, or ~1% of the nuclear diameter.***

***J. Soreff points out that such strain fields have components at various spatial frequencies, and that at high spatial frequencies there is much design freedom but the effects decay exponentially with a short characteristic length, hence design choices are not spread uniformly throughout the constraint space but rather are clustered. As a result, it may not be possible to achieve a 0.01-nm designed receptor topography simultaneously everywhere across an entire binding surface, nor may it be possible via surface binding alone to distinguish molecules which differ only deep in their interiors.
It is also possible to translate diamondoid components through a picometer step size, much smaller than the unavoidable RMS thermal displacements, using any of several methods; for example:

A. **Levers** — Consider a 10-nm lever joined to a fixed bar by a pivot at one end, and driven axially by a ratchet interposed between lever and bar at the other end. Ratchet movements translate to smaller displacements at positions along the lever distant from the ratchet. Thus a follower rod attached to the lever 1 nm from the pivot and driven by a ratchet with 0.01-nm steps moves ~0.001 nm per ratchet step.

B. **Screws** — Consider a 3-nm diameter cylindrical screw with a 1-nm pitch. Rotating the screw through a 0.01-nm circumferential displacement causes the screw to move laterally by ~0.001 nm, which may be transmitted elsewhere in the machine by an attached follower rod. Of course, nanoscale screws or gears are sensitive to the precise cancellation of the potentials and hence cannot be perfectly smooth and circular, producing some unavoidable “knobiness” under load.

C. **Gear Trains** — Consider a 32-nm diameter worm gear with 1-nm teeth. One rotation of the gear requires 100 rotations of the worm; hence a 0.1 nm displacement applied to the worm produces a 0.001 nm displacement in the gear. More efficient (and coaxial) compound planetary gear trains commonly employed in transmissions achieve displacement ratios up to 10,000:1, a hundred times better than the above example.

D. **Hydraulics** — Consider a sealed, fluid-filled, tapered pipe. A piston 1000 nm² in area is mounted at one end; another piston 10 nm² in area lies at the other end. A displacement of 0.1 nm applied to the smaller piston produces a 0.001 nm displacement in the larger piston. (Here again the finite size of molecules may produce “knobby” performance as fluid particles slip from one stable configuration to the next.)

E. **Compression** — Consider a rod upon which a compressive force of 100 nN/nm² (near the maximum diamondoid strength) has been imposed. Affixed to the rod are two crossbars spaced 4 nm apart. If the force on the rod is increased to 101 nN/nm², the gap between the crossbars compresses by ~0.001 nm.

### 3.5.7 Receptor Configurations

Many different useful receptor configurations may be readily envisioned, of which the following brief descriptions are but a small sample. Note that most large target molecules of nanomedical interest are proteins which are probably floppy enough to permit entry into reasonably open multiply-concave rigid receptor structures; if not, hinges are easily added to the receptor design.

#### 3.5.7.1 Imprint Model

Molecular imprinting is an existing technique in which a cocktail of functionalized monomers interacts reversibly with a target molecule using only noncovalent forces. The complex is then cross-linked and polymerized in a casting procedure, leaving behind a polymer with recognition sites complementary to the target molecule in both shape and functionality (Fig. 3.11). Each such site constitutes an induced molecular “memory,” capable of selectively binding the target species. In one experiment involving an amino acid derivative target, one artificial binding site per (3.8 nm)³ polymer block was created, only slightly larger than the (2.7 nm)³ sorting rotor receptors described by Drexler (Section 3.4.2). Chiral separations, enzymatic transition state activity, and high receptor affinities up to Kd ~ 10⁻² have been demonstrated, with specificity against closely competing ligands up to ∆Kd ~ 10⁻² (~20 zJ).

Several difficulties with this approach from a diamondoid engineering perspective include:

1. A sample of the target molecule is required to make each mold.
2. It is currently unknown how to prepare diamondoid castings.
3. Once the imprint has been taken, the site cannot easily be further modified.

#### 3.5.7.2 Solid Mosaic Model

In the solid mosaic receptor model, the precise shape and charge distribution of the target molecule is already known. Working from this information, a set of diamondoid components could be fabricated which, when fitted together like a Chinese puzzle box, create a solid object having a cavity in the precise shape of the optimum negative image of the target molecule (Fig. 3.12A). The mosaic may contain point charges, voids, stressed surfaces, or dislocations to achieve fine positional control. Mosaic components may be as small as individual atoms, so this model is conceptually similar to 3D printing or raster-scan techniques in which the desired cavity formation is constructed atom by atom inside a nanofactory (Fig. 3.12B; Chapter 19). This model, like the imprint model, cannot easily be reconfigured once it has been constructed because each of the many unique parts may contribute to the entire structure. The construction of receptors from parts of fixed size and shape is crudely analogous to members of the heterodimer receptor class (a two-component receptor) such as GABAB.

---

*Fig. 3.11. Imprint model for creating artificial molecular receptors (redrawn from Ansell, Ramstrom and Mosbach).*
M. Reza Ghadiri has designed a protein mosaic model using cyclic peptides that assemble spontaneously into nanotubes of pre-defined diameter; incorporation of hydrophobic amino acid side chains on the outside of these tubes leads to spontaneous insertion into bilayers, allowing the tubes to function as transmembrane ion channels. Other examples of mosaic model receptors are mesoporous silica filters with functionalized organic monolayers forming 3-6 nm sieve-like pores and zeolites and zeolite-like molecular sieves. Zeolites are artificial crystal structures with precise and uniform 0.4-1.5 nm internal void arrays which can also be used as shape-selective catalysts able to favor one product over another that differs in size by as little as 0.03 nm, such as p-xylene and o-xylene. By 1998, rational de novo computational design of artificial zeolite templates and crystal engineering had begun.

### 3.5.7.3 Tomographic Model

In the tomographic receptor model, the receptor engineer again starts with a known target molecule topography and designs a series of thin planar sections which, when stacked together in the correct order (using positionally-coded docking pins) and bonded, create a solid object containing the desired optimum binding cavity (Fig. 3.13). As in the mosaic model, point charges or dislocations in each planar segment can be used to manipulate cavity features and dimensions to precise tolerances. Unlike the mosaic model, a tomographic receptor can be reconfigured by partial disassembly and replacement.
of specific planar segments, each of which contributes only locally to the total receptor structure. Hybrid or modular artificial enzymes and two-dimensional sheetlike hydrogen-bonded networks are crude analogs in current research.

### 3.5.7.4 Pin Cushion Model

The pin cushion receptor, suggested independently by K.E. Drexler [personal communication, 1996], is a hemispherical or hemiellipsoidal shell through which a number of rods protrude, each of which may be moved radially (Fig. 3.14). When inserted through the shell to varying depths, the endpoints of the rods define a negative image surface which may be made to mirror the topography and charge distribution of a known target molecule. Rods may be tipped with positive, negative or no charge, or they may terminate in any number of functionalized surface segments designed to optimally match parts of the target molecule shape. Other configurations such as a rectangular box, hinged plates with protruding rods, counterrotating rollers, or time-varying rod positioners are readily conceivable. Pin cushion receptors are easily reconfigured to bind different target molecules, hence may be regarded as fully programmable "universal" binding sites. The principal difficulty with the pin cushion receptor is its excessive size (compared to other receptor models) and its greater complexity (since each rod may be controllable individually).

Pin cushion receptors can also be used to discover the shapes of unknown molecules (Section 3.5.8): A target molecule is placed in the central cavity with all rods fully retracted, and the rods are slowly slid forward using nanopistons with force reflection feedback, until all pistons register zero force, indicating balance between attractive and repulsive van der Waals interactions, at which point all rod positions are recorded. Rods of differing end tip charge may then be tested for additional attractive potential. The final result is a precise mapping of the target molecule, which data may be stored or transmitted elsewhere for future use.

### 3.5.7.5 Construction Costs

The active binding site of a receptor consisting of \((2.7 \text{ nm})^3 \cdot 19 \text{ nm}^3\) of structural atoms and constructed with 0.001-nm feature sizes in theory requires information from \(2 \times 10^{10}\) voxels (volume pixels) for complete description. However, there are only \(N_{\text{atom}} \approx 1000\) atoms involved in the structure and their locations cannot be arbitrarily chosen. Atomic scale (-0.1 nm) resolution would require \(-19,000\) voxels; each voxel minimally requires an index number (\(-\log_{10}(19,000) \approx 14\) bits), an atomic identifier (\(-\log_{10}(92) \approx 7\) bits), and a charge identifier (\(-\log_{10}(3) \approx 2\) bits), for a total of \(4 \times 10^3\) bits/receptor at atomic scale resolution. Drexler estimates the number of configurational options per atom \(N_{\text{opt}} \approx 150\); hence a description of the receptor could require as few as \(N_{\text{atom}} \log_2(N_{\text{opt}}) \approx 7 \times 10^3\) bits.* Assuming energy dissipation of \(-5 \text{ zJ}/\text{bit}\) for rod logic register reading [17] or \(-kT \ln (2)\) (Eqn. 7.1), then each time the receptor description is retrieved, stored or processed may require a minimum energy dissipation of \(-10^{-4} - 10^{-6}\) zJ. Reversible computing may reduce this energy requirement by a factor of 10-100 or more (Section 10.2.4.1). Construction of a receptor containing \(-1000\) atoms using the nanomanipulator arm mentioned in Section 3.4.3 requires \(-10^{-2}\) sec and consumes \(-0.001\) picojoule \((-10^6\) zJ) of mechanical energy.

### 3.5.7.6 Receptor Durability

Receptor durability is difficult to estimate ab initio. Molecularly imprinted polymer sites can be stored at least several years without loss of performance, but these receptors have only been tested to \(\approx 100\) cycles of use without any detectable loss of memory. The lifetime in vivo for metabolic enzymes often exceeds \(-10^5\) sec (-1 day), and taste cell receptors survive \(-10^6\) sec (-2 weeks), which suggests operational lifetimes for natural receptors on the order of \(10^6 - 10^{12}\) cycles despite the relative fragility of protein structures. Diamondoid structures should be even more durable because of their superior physical strength, affirmative forcible ejection of bound ligands each cycle, and high resistance to chemical degradation thus reducing susceptibility to poisoning.

### 3.5.8 Ligand-Receptor Mapping

To achieve a fully general-purpose receptor capability in nanomedical systems, two classes of analytic function are essential. First, presented with an arbitrary molecule, the system or user must be able to infer from the molecule’s structure the shape and electronic configuration of an optimal receptor geometry that will

---

*Higher level descriptions may allow considerable additional information compaction, as for instance where a 30-bit word indexes a single receptor structure stored in a library containing perhaps \(10^9\) distinct designs.
efficiently bind it, with a particular affinity and specificity (as a
design specification). This discovery procedure may involve a process
akin to molecular imprinting (Section 3.5.7.1), fluorescent dye
affinity matching on testing chips (Chapter 20), Structure-Activity
Relationships (SAR) by NMR techniques,424 or pin cushion receptor
mapping (Section 3.5.7.4).

Second, presented with an arbitrary protein-built binding site
embedded in living tissue, the system or user must be able to infer
the molecule(s) which the given receptor could bind, and to compute
the affinity and specificity of that activity. This capability may
require a rather diverse steric toolkit. Biological receptors are
constructed in one of ~500-1000 distinct shapes or “domains,” such
as the well-known Y-shaped antibody immunoglobulin domain
(~100 residues) and other assorted clefts, folds, kringles and coils.
Several hundred distinct fold families are known, though it is believed
there are only ~20 major domain types,414 and Chothia2604 believes
that “the large majority of proteins come from no more than one
thousand families.”

One mapping technique would employ a series of rodlike probes
inserted into the receptor cavity like a pick gun or other lockpicking
tool. After physically securing the receptor, the first crude probes
quickly map the cavity to nanometer scale. Based on this preliminary
information, subsequent probes having ever-finer discrimination
chart smaller features as well as charge distributions by inserting a
predetermined set of test rods with functionalized tips in a standard
sequence, to rapidly prune the huge configurational space down to
a single unique electrophysical shape using the minimum possible
number of tests. (A 1 nm³ volume of multi-element diamondoid
receptor structure has ~10¹⁴⁸ possible distinct configurations, by
one conservative estimate,10 requiring at least 492 binary tests to
eliminate all but one configuration. Antibody domains contain ~10⁵⁰
possible configurations, requiring 166 binary tests.) Cavity Stuffer,
a software package comprising an experimental design tool to
investigate automated cram-packing of predefined cavities using
randomly branched polymers, is a preliminary effort in this general
direction,425 although the algorithmic task of discovering an
unknown receptor contour may be considerably more challenging.

Another approach to receptor mapping would be the reversible
chemical or mechanical denaturation of the receptor protein
followed by precise nondestructive amino acid sequencing (Chap-
ter 20), from which tertiary structure and activity could then be
computationally inferred. Algorithms to perform such computations
are the subject of intense current research interest.952 Even imperfect
tertiary structure predictions should greatly reduce the search space,
so that only a partial residue sequencing may be necessary for
unambiguous identification from a library of possible proteins. Once
the target ligand structure has been inferred, the subsequent design
and manufacture of receptor-specific agonists and antagonists,
including catalysts and cofactors, activators and inhibitors, promoters
and repressors, should be comparatively easy.

3.5.9 Large Molecule Binding, Sorting and Transport

Is there any size limit for target molecules to be transported?
Natural receptors have already been found for large molecules
including low-density lipoproteins (LDLs) > 1,000,000 daltons426
and high-density lipoproteins (HDLs).1038

The methods described in earlier Sections can be adapted for
binding large molecules (>1000 atoms; Fig. 3.15), including molecules
far wider than the binding device itself (e.g., ~200-nm diameter
virus particles and larger). Making a binding site for a large molecule
should be physically easier (albeit computationally more challenging)
than making a binding site for a small molecule because of the greatly
increased area of interaction. For example, a binding energy of 400 zJ
may be realized by creating a dispersion-force binding area covering
only ~25% of the surface of a 10,000-atom target molecule (Table
3.6) or a mere ~0.02% of a 200-nm virus particle.
This makes possible the concept of binding pads—small surfaces with dimples (concave or convex), each dimple consisting of precisely-placed nanometer-scale features that are complementary to specific patches (e.g., epitopes) on the surface of the large target molecule. (Specificity is lost for portions of the molecule outside the particular patches.) Each dimple could effectively grasp the side of the large molecule without having to fully enclose it—a capability useful in nanorobot foot pads during cell walking and anchoring (Section 9.4.3), in handles for nanociliary or nanomanipulator transport functions (Section 9.3.2), and for chemotactic sensing (Section 4.2.6). For example, macrophage receptors for LDLs employ a “pad” consisting of three globular cysteine-rich domains. The large compliance of target protein molecule subunits should prove curative for any misalignment problems caused by cumulative small errors in bond lengths across large diamondoid receptor structures.

Rather than using sorting rotors, which become unwieldy when large molecule binding pockets must be used, the shuttle pump illustrated schematically in Figure 3.16 may provide reasonably efficient large molecule sortation and transport. The shuttle pump consists of a diamondoid tube within which a receptor ring moves between iris diaphragms at either end (shuttling mechanism not shown). The receptor ring is constructed as two or more binding pad segments. For molecule pickup, the ring is pressed together, forming an annular binding region for the target large molecule, which binds and is shuttled to the other side. The receptor ring is then fragmented, destroying binding affinity and unlocking the target molecule, which escapes via diffusion. The shuttle returns to the pickup side, the receptor is pressed together again, and the cycle repeats. A biocompatible solvent environment is maintained during large-protein manipulation tasks.

Assuming a roughly spherical large molecule and laminar fluid flow at 1 atm forcing pressure (Section 9.2.7), a 10-nm diameter molecule moves through a 20-nm long pump (~10^{-20} kg, ~10^9 atoms) in ~10^{-6} sec at ~0.02 m/sec, consuming ~0.02 pW during transfer. A 200-nm virus-size target molecule moves through a 400-nm long pump (~10^{-17} kg, ~10^9 atoms) in ~10^{-2} sec at ~60 microns/sec, consuming ~10^{-16} watts during transfer; at ~0.0002 atm, release time is diffusion limited. The transfer force exerted on a 10-nm molecule is ~1 pN, ~600 pN on a 200-nm virion; a binding energy of 400 zJ at a 0.2-nm contact distance gives a binding force of ~2300 pN, sufficient to hold a particle of either size firmly during transport and release.

J. Soreff points out that as protein size increases, so does the energy available for local minima in the binding. Desired proteins may become stuck in incorrect positions, or undesired proteins may become partially adhered to a receptor. Besides designing to minimize these possibilities, using a multireceptor cascade with different combinations of binding patches at each stage should allow complete exclusion of undesired large-molecule species.
4.1 Nanosensor Technology

Medical nanorobots need to acquire information from their environment to properly execute their assigned tasks. Such acquisition is achieved using onboard nanoscale sensors, or nanosensors, of various types. Nanosensors allow for medical nanodevices to monitor environmental states at three different operational levels:

1. Internal nanorobot states,
2. Local and global somatic states (inside the human body), and
3. Extrasomatic states (sensory data originating outside the human body).

The general physical limits to sensory perception are reviewed by Block and Bialek.

The specific nanosensor technologies required include sensors to detect chemical substances (Section 4.2), displacement and motion (Section 4.3), force and mass (Section 4.4), and acoustic (Section 4.5), thermal (Section 4.6), and electromagnetic (Section 4.7) stimuli. Typical sensor device mass, volume, and sensitivity limits are summarized in each Section. In vivo bioscanning is briefly described in Section 4.8, followed by external macrosensing in Section 4.9. Methods of getting sensors in and out of the body are explained in Chapter 7. A discussion of sensor biocompatibility is deferred to Chapter 15.

4.2 Chemical and Molecular Nanosensors

The area of chemical microsensors is well developed, and there is increasing research interest in nanoscale chemical sensors. "chemosensors" and "biosensors". The most common nanomedical application of chemical sensors will be to measure the concentration of specific molecules and biopolymers in aqueous solvent—whether in blood serum, interstitial fluid, or cytosol. This may be accomplished using ordered arrays of receptors or by using sorting rotors to directly count molecular populations in known sample volumes. Spatial and temporal concentration gradient sensing is also essential in navigation, communications, and in mediating rapid response to environmental stimuli. Chemotactic sensors may be used to sample the chemical composition of surfaces.

4.2.1 Broadband Receptor Arrays

One simple broadband concentration sensor, shown schematically in Figure 4.1, consists of a graduated series of receptors having uniformly high specificity but engineered with progressively greater affinities (successively smaller equilibrium dissociation constants $K_d$ (molecules/nm$^3$); Section 3.5.2) for the target molecule, which is present in the test sample at concentration $c_{\text{ligand}}$. Exposure of the receptor array to the test sample for a time $t_{\text{EQ}}$ necessary to reach diffusion-driven equilibrium gives a probability of receptor occupancy $P_{\text{occ}}$ in a receptor with $K_d ~ 0.1 c_{\text{ligand}}$, $P_{\text{occ}} ~ 0.50$ in a receptor with $K_d ~ c_{\text{ligand}}$, and $P_{\text{occ}} ~ 0.09$ in a receptor with $K_d ~ 10 c_{\text{ligand}}$.

During each measurement cycle, all steric probes are simultaneously extended into their associated receptor volumes at a time $t_{\text{EQ}}$ after presentation of sample to the array. Probes which reach full extension register an empty receptor, those which cannot fully extend register an occupied receptor. After registration, the probes are retracted and ejection rods are thrust into all receptors (typically requiring ~1 nanonewton (nN) of force for occupied receptors) to empty them while the test chamber is flushed clear in preparation for the next cycle to begin. Ideally, the probe rod pushing into the binding site should push the bound molecule further into the binding site to avoid complications that may arise due to competing kinetic barriers, and ejection rods should push molecules far enough to ensure that the next resample is independent.

Consider a series of $N$ steric probe units as proposed by Drexler. Each probe measures $8 \times (2.5 \, \text{nm})^2 ~ 50 \, \text{nm}^3$ and has mass $\approx 2 \times 10^{-22} \, \text{kg}$. The ratio of the dissociation constants of adjacent units is $\kappa = K_{d_i} / K_{d_{i+1}} > 1$, with $K_{d_i} ~ 10^{-4} \, \text{molecules/nm}^3$ and $K_{d_N} ~ 10^{-13} \, \text{molecules/nm}^3$, typically in the human body, and so $N = 1 + \log_{10}(K_{d_i}/K_{d_N}) / \log_{10}(\kappa)$.

Minimum measurement error occurs when $c_{\text{ligand}}$ exactly matches the $K_d$ of a probe unit, that is, when $c_{\text{ligand}} = K_d$ (i.e., $P_{\text{occupied}} = 0.5$). Maximum measurement error occurs when $c_{\text{ligand}}$ lies exactly midway between two probes such that $K_{d_i} > c_{\text{ligand}} > K_{d_{i+1}}$, or, more specifically, at the geometric midpoint $c_{\text{ligand}} = (K_{d_i} K_{d_{i+1}})^{1/2} = K_d \kappa^{-1/2}$. In this case, Eqn. 3.24 becomes

$$P_{\text{occupied}} = \left( \frac{c_{\text{ligand}}}{K_d} \kappa^{1/2} \right) P_{\text{unoccupied}}$$

where $P_{\text{unoccupied}} = 1 - P_{\text{occupied}}$. Additionally, sampling error is $-N_m^{-1/2}$ when $N_m$ independent measurements are taken (i.e., $N_m$ measurement cycles employing all $N$ probe units during each cycle). For $N_m > 1$, establishing an error bound on the probability of receptor occupancy of $P_{\text{occupied}} \pm N_m^{-1/2}$. If two concentrations $c_1 = P_{\text{occupied}} / (1 - P_{\text{occupied}})$ and $c_2 = (P_{\text{occupied}} + N_m^{-1/2}) / (1 - (P_{\text{occupied}} + N_m^{-1/2}))$ may be distinguished, where $P_{\text{occupied}}$ is given by Eqn. 4.1, then the minimum detectable concentration differential $\Delta c / c$ is

$$\Delta c / c = \left[ c_2 / c_1 \right] - 1 = \frac{1 + \kappa^{-1/2} N_m^{1/2}}{1 - (1 + \kappa^{-1/2})^{-1} N_m^{1/2}} - 1$$

[Nanomedicine, Volume I: Basic Capabilities, by Robert A. Freitas Jr. ©1999 Landes Bioscience.]
For common small molecules, $t_{EQ} \approx 10^{-6}$ sec; for common large molecules, $t_{EQ} \approx 0.2 \times 10^{-6}$ sec, but a sample chamber (~25 nm$^3$) in size is diffusion-limited to a minimum $t_{EQ} \approx 10^{-5}$ sec for $c_{ligand} \approx 10^4$ nm$^{-3}$. For the rarest small molecules in the human body, $t_{EQ} \approx 800$ sec, ~200 sec for rare large molecules, and a larger chamber may also be required. Preconcentration of the test sample by dehydration using sorting rotors (Section 3.4.2) to rapidly remove water from the sample chamber reduces $t_{EQ}$ to 10-40 sec for very rare molecules.

Steric probes using a driving force$^{10}$ of -100 pN, over 2 nm of motion per cycle require 200 zJ/measurement. If 1000 measurement cycles of a 10-probe sensor suffice to determine concentration to ~2 significant figures, this information costs ~0.002 pJ to acquire, a power consumption of from ~10 pW for large common molecules to ~10$^{-9}$ pW (~0.6 kT/sec) for small rare molecules while the sensor is working. Power dissipation may run up to ~10 times higher, depending upon the number of occupied receptors that must be cleared via ejection rods during each cycle.

4.2.2 Narrowband Receptor Arrays

If the concentration of the target ligand varies only slightly within known limits, similar results may be obtained using a narrowband sensor that employs a repeating array of a single receptor type tuned to the maximum expected concentration $c_{max}$ of the target ligand (Fig. 4.2). If the sensor consists of $N_r$ receptors, choosing $K_d \approx c_{max}$ $N_r^{-1}$ / (1 - $N_r^{-1}$) ensures that all receptors are occupied at $c = c_{max}$ thus avoiding array saturation at concentrations below $c_{max}$. At minimum sensitivity, $K_d = c_{min} (1 - N_r^{-1}) / N_r^{-1}$, and so $c = c_{max}$ / $c_{min}$ for the entire array. Thus the minimum detectable concentration differential $\Delta c / c$ for the array is

$$\Delta c / c = \left( \frac{N_r - 1}{2N_r} \right)$$

{Eqn. 4.4}

If $A = $ receptor unit active area (~50 nm$^2$), $L = $ receptor unit length (~10 nm), and receptor arrays are packed in parallel sheets with separation $\Delta x$ (~10 nm, for small-molecule diffusion in ~10$^{-6}$ sec) to allow sample access, sensor volume $V_s = AL N_r (1 + (\Delta x / L))$ and sensor scale $V_s^{1/3}$. Thus a sensor with $N_r = 100$ receptor units achieves $\Delta c / c = 0.10$ (10%) with a (46 nm)$^3$ sensor; $N_r = 1400$ receptor units achieves $\Delta c / c = 0.01$ (1%) with a (112 nm)$^3$ sensor. A dedicated micron-scale nanodevice employing a (1.44 micron)$^3$ sensor array having 3 million receptor cells could at best distinguish 0.001% concentration differentials in ~3 sec (from Eqn. 4.6) for common molecules like serum glucose.

The ion channel switch (ICS) biosensor (aka. the “Australian sensor”) can detect small-molecule concentrations as low as ~10$^{-14}$ nm$^{-3}$ in a ~600 sec measurement time.$^{3039,3040}$

4.2.3 Counting Rotors

Concentration may be determined by varying exposure time to the sample. One energy-efficient design (Fig. 4.3) uses an input sorting rotor running at varying speeds (according to target concentration) synchronized with a counting rotor (linked by rods and ratchets to a data storage register device) to assay the number of molecules of the desired type that are present in a known volume of fluid. The counting rotor uses a steric probe, moved up and down in synchrony with the arrival of a binding site under the test probe, and ratchets to a data storage register device) to assay the number of ligands striking the receptor surface per second.$^{434}$ or

$$t_{EQ} = \left( \frac{N_{encounters}}{A_{ligand}} \right) \left( \frac{\pi MW_{kg}}{2kT N_A} \right)^{1/2}$$

{Eqn. 4.3}

where $A = $ active receptor cross-sectional area ~0.1 nm$^2$ for small molecule (MW ~100 gm/mole) receptors to ~10 nm$^2$ for large molecule (MW ~10$^7$ gm/mole) receptors; $c_{ligand} = 10^{-2}$ nm$^{-3}$ for the most common molecules to 10$^{-12}$ nm$^{-3}$ for rare molecules in the human body; MW$_{kg} = $ target ligand molecular weight in kg/mole, $k = 0.01381$ zJ/K (Boltzmann constant), $T = 310 K$ (human body temperature), and $N_A = 6.023 \times 10^{23}$ molecules/mole (Avogadro’s number).

Fig. 4.1. Schematic of broadband chemical concentration sensor array using 5 receptor units with steric probes.
Basic Capabilities • Nanosensors and Nanoscale Scanning

pistons (Section 9.2.7; not shown) operated slowly enough to avoid damaging molecules in the sample volume. This sensor, which measures 45 nm x 45 nm x 10 nm comprising ~500,000 atoms (~10^{-20} kg), should count \( \hat{N} \sim 10^4 \text{ - } 10^5 \) molecules/sec of small common molecules present at typical human serum concentrations of \( 10^{-3} \text{ - } 10^{-2} \text{ nm}^{-3} \), such as glucose, with the rotor spinning at ~1% of normal speed (Section 3.4.2). Clogging of intake apertures in vivo is addressed in Sections 3.3.1 and 9.2.3, and in Chapter 15.

For rare molecules with \( c \sim 10^{-11} \text{ nm}^{-3} \), the count rate for a single sensing rotor falls to \( \hat{N} \sim 10^{-4} \) molecules/sec. However, a \((340 \text{ nm})^3\) reservoir (a \(10^4\) larger chamber volume) serviced by a bank of \(10^4\) rotors (\(4 \times 10^{-18} \text{ kg}\) additional mass) improves the count rate for rare molecules to \( \hat{N} \sim 1 \) molecule/sec; that is, \( \dot{N} \sim 2500 \text{ c ligand reservoir}^3 \) (molecules/sec) \{Eqn. 4.5\}

Sample preconcentration may allow further improvement by a factor of up to 10-100. Energy requirements are comparable to those for sorting rotors (Section 3.4.2) except for the need to retrieve ~all target molecules from a known sample volume, which slows the sorting process, requiring ~10 zJ/count for common molecules and ~100 zJ/count for the rarest molecules in the human body. Large-molecule concentrations may be measured using a shuttle pump (Section 3.5.9) and a counting device.

4.2.4 Chemical Assay

It may also be useful to design a sensor that is capable of detecting as many different chemical species as possible, in contrast to detecting the precise concentration of a single species, for use in chemomessaging (Section 7.4.2.4), chemonavigation (Section 8.4.3), cellular diagnosis (Chapter 21), and other in vivo assay work. Counting rotors may be used for this purpose. Each 12-receptor sorting rotor/counting rotor pair has a minimum \( 400 \text{ nm}^3 \) volume.\(^{10}\) Allowing an additional ~800 nm\(^3\) per pair to account for power, control, mechanical attachments and housings, which includes ~300 nm\(^3\) per pair for sample chamber volume and sample access, requires ~100 nm\(^3\) per receptor, roughly the same as for the steric probe units described in Section 4.2.1. A dedicated 1 micron\(^3\) chemical assay nanorobot, which includes ~0.25 micron\(^3\) sample volume, with ~\(10^7\) pairs (receptors) could thus continuously scan for as many as ~\(10^7\) different chemical species, counting \( \hat{N} \sim 1 \) large molecule/sec at concentrations of \( c_{\text{ligand}} \sim 10^{-10} \text{ molecules/nm}^3\).

4.2.5 Chemical Nanosensor Theoretical Limits

Berg and Purcell\(^{337}\) estimated that an ideal concentration sensor, limited only by diffusion constraints and drawing through a spherical boundary surface of radius \( r_s \), provides a minimum detectable concentration differential \( \Delta c / c \) of

\[
\Delta c / c = (1.61 \Delta t D/c r_s)^{1/2} \quad \{\text{Eqn. 4.6}\}
\]

where \( \Delta t = \text{time (sec)} \), \( D \) is diffusion coefficient \( 10^{-9} \text{ m}^2/\text{sec} \) for small molecules (Table 3.3), and \( c \) = concentration (molecules/m\(^3\)). Using a sample chamber with \( r_s = 10 \text{ nm} \), a concentration of \( c = 3 \times 10^{-3} \text{ nm}^{-3} \) for serum glucose could be measured with 1% uncertainty in \( \Delta t = 260 \) microsec, or with 0.01% uncertainty in 2.6 sec. A dedicated micron-scale nanodevice using a 10-sec sampling time could at best distinguish a 0.0005% concentration differential for small common molecules like serum glucose but only a 27% concentration differential for large rare serum molecules like somatotropin.

Fig. 4.2. Schematic of narrowband chemical concentration nanosensor array using receptors of a single type.

Fig. 4.3. Chemical concentration sensor using counting rotors.
Berg and Purcell\textsuperscript{337} also found that the capture of small molecules at nanobobot surfaces can be surprisingly efficient. Specifically, for a spherical nanodevice of radius $R$, across whose surface are uniformly distributed $N_r$ receptor spots of radius $r_r$, the maximum diffusion current absorbed by the receptor array, $J_{array}$ is

$$J_{array} = J \left( \frac{N_r r_r}{N_r r_r + \pi R} \right)$$

(Eqn. 4.7)

where $J$ = maximum diffusive intake current, given by Eqn. 3.4. For $r_r = 1$ nm and $R = 1$ micron, the $J_{array}$ is fully 50% of $J$ using only $N_r = 3100$ surface receptors, which occupy a mere 0.1% of the total device surface area.

For a purely sensory spherical nanodevice, a uniform distribution of a relatively small number of chemoreceptors confers optimal sensitivity. However, in applications where sensors (which attract and then release) are in competition with pumps (which strongly attract and then remove) for the same target molecules at the nanodevice surface, greater sensing efficiency may be achieved by clustering the sensors, as in a "nose" organ.\textsuperscript{437}

Proteins may be modified as they age within biological cells. For instance, under steady state conditions, 10% of protein molecules may exhibit carbonyl (oxidation) modifications,\textsuperscript{2138} glycosylation sites that are likely to be subject to such alterations. Proteins may be modified as they age within biological cells. For instance, under steady state conditions, 10% of protein molecules may exhibit carbonyl (oxidation) modifications,\textsuperscript{2138} glycosylation (non-enzymatic), and dysfunctional proteins may be ubiquitinated (Chapter 13) and phosphorylated under enzyme control. Heat shock also modifies proteins. Chemosensor receptors must be carefully designed to accommodate changes in target ligand structures or sites that are likely to be subject to such alterations.

\subsection*{4.2.6 Spatial Concentration Gradients}

Medical nanodevices can detect both extremely steep and extremely shallow spatial concentration gradients. For example, a receptor array sensor (Section 4.2.1) consisting of 10 steric probe units has a total volume of -1000 nm$^3$ including its prorated share of drive systems, housings, etc. In just a few measurement cycles, adjacent sensors can detect a maximum concentration change from $10^{-2}$ nm$^3$ to $10^{-11}$ nm$^3$ across a mean nanosensor separation of 10 nm on a nanodevice surface, a steep chemical spatial gradient of $10^{-10}$% per nanometer. Soreff points out that maintaining a steep concentration gradient of a freely diffusing species implies a power dissipation of $-kT$ per molecule per time to diffuse a concentration-halving distance. For a common, small, tethered molecule with $\text{ligand} = 3 \times 10^{-3}$ molecules/nm$^3$ that diffuses $-1$ nm in $\tau_{diff} = 1$ nanosec., implied power density $D_{mol} = kT \text{visc} / \tau_{diff} = 10^{-11}$ watts/m$^3$, a power density that might possibly be encountered in the vicinity of chemoelectric transducers (Section 6.3.4.5). (Such high power densities never occur on macroscale dimensions, and heat rapidly diffuses at nanoscale dimensions; Sections 4.6.1.)

Similarly, counting rotor sensors (Section 4.2.3) for small common serum molecules like glucose deployed at opposite ends of a 1-micron nanodevice can detect a concentration differential of $-10^{-6}$ over a -1 sec measurement period, a shallow chemical spatial gradient of only $-10^{-5}$% per nanometer. Even lower gradients may be detected by communication (Chapter 7) between physically separated cooperating nanodevices.

\subsection*{4.2.7 Temporal Concentration Gradients}

Koshland and Macnab\textsuperscript{435} have shown that bacteria possess a mechanism for sensing small temporal concentration gradients. A bacterium detects a spatial gradient of attractant not by comparing the concentration at its head and tail, but rather by traveling through space and comparing its observations over time.\textsuperscript{E. coli} can measure a $-0.01\%$ gradient\textsuperscript{436} during the time it travels one body length (~2 microns) at top speed (20–40 microns/sec), a minimum temporal gradient of 0.1% per second. This appears very efficient, since the counting rotor sensor proposed in Section 4.2.3 is not much better, achieving $-0.01\%$ per second. As for peak gradients, the 10-probe receptor array sensor described in Section 4.2.1 can detect a maximum concentration change (e.g., from an already-known rare concentration) in $-10^{-4}$ sec, a peak temporal gradient of $-10^{13}$ % per second.

\subsection*{4.2.8 Chemotactic Sensor Pads}

Besides concentration sensing, chemical sensors may be used to determine the chemical characteristics of surfaces. A binding pad coated with an array of reversible, possibly reconfigurable, artificial receptors (Section 3.5.9) may be pressed against test surfaces such as cell membranes, or the perimeters of very large molecules, in order to detect the presence of specific chemical ligands. If pad receptor specificity is high for the target ligand, then the degree of pad "stickiness" to the test surface is a good measure of the population of target ligands on that surface. Probing two-dimensional surfaces with receptors may be constrained by orientation effects, since binding may require a specific relative orientation of ligand and receptor. Compliant pad receptors can reduce this problem but may impair lateral resolution—there is a three-way tradeoff among angular tolerance, lateral resolution, and multioriented multicopy receptors. A physical contact sensor prevents confusion of surface-bound and free-solvated ligands of the same species. The selective and reversible stickiness of such pads can also be exploited for locomotion (Section 9.4.3.1).

Molecular adhesion experiments in chemical force microscopy\textsuperscript{440,1066} and interfacial force microscopy\textsuperscript{441} using AFM probes with functionalized tips (Section 2.3.3) typically detect maximum adhesive noncovalent force differentials of 40–160 piconewton (pN)/ligand for competing ligands at a peak-force distance of ~0.2 nm. Forces $>10$ pN are readily measured by nanosensors (Section 4.4.1), so binding pad sensors should give an exact count of the number of target ligands present per unit area of test surface if (receptor areal number density) > (ligand areal number density) > (1 ligand / pad area), and if contact time is adequate. Receptors for many different ligands can be mounted on moving rotors or belts, or employed as interchangeable tool tips (Section 9.3.2), thus allowing a single sensor pad to reconfigure its specificity or sensitivity under external control.

\subsection*{4.2.9 Receptor Sensors}

Ligand mimics may also be used to allow a medical nanodevice to detect a particular receptor in its environment. Each such "receptor sensor" is a protruding physical structure that mimics all or part of the electrostatic structure of the ligand to which the target receptor has maximum specificity. For example, probes consisting of artificial structures that simulate antigen peptides (8-15 amino acids long) bound in an MHC-like structure\textsuperscript{439} (Section 8.5.2.1) could be used to sense the presence of specific T-cell receptors.

Sensing pads displaying ligand mimic structures are brought into contact with the test surface. Any target receptors present will bind to a mimic structure, thus may be counted as the sensor sweeps slowly across the surface. Immediately after recognition has occurred, the mimic structure may be reversibly fragmented, destroying receptor/mimic affinity and allowing the receptor to release the mimic. Mimics are mounted on the sensor pad with an appropriately elastic compliance to prevent structure detachment during a measurement scan.

Receptor sensing is most useful in detecting receptors for large molecules. In the case of receptors for small molecules, there may
sometimes be insufficient space in the target molecular volume to match the surface electrosteric structure, to accommodate ligand fragmentation control rods or cables, or to build internal boundaries sufficient to permit ligand fragmentation without creating undesirable reactive species. Some of these problems with small-molecule receptors may be partially overcome by employing an annular repulsive ring through which a mimic ligand may be forcibly retracted without fragmentation, but receptor orientation effects will present additional constraints.

### 4.3 Displacement and Motion Sensors

The ability to precisely measure position and displacement, velocity and acceleration, rotation and orientation is essential requirements for nanoscale locomotion, physical manipulation, and navigation. A comprehensive analysis is beyond the scope of this book, but the following approximations should prove useful in nanomedical design.

#### 4.3.1 Displacement Sensors

Diamondoid parts may contain features, or be positioned, in picometer (1 pm = 10^-12 m) steps, much smaller than the unavoidable RMS thermal vibrations (Section 3.5.6). The nanomanipulator robot arm described in Section 9.3.1.4 may be moved in picometer increments, and the RMS thermal longitudinal displacements at the tip of a 20-nm long, 10-nm wide diamondoid lever is -1 pm. 10 Displacement sensitivity of 10 pm is routinely achieved in STMs. It is claimed this can be improved down to the -1 pm level; 450 STM resolution of 2 pm has been demonstrated experimentally. 10 A C_{60} molecule used as the active element in an electromechanical amplifier transmits -100 times more electrical current when physically compressed by 100 pm, 301 suggesting a minimum deformation detection limit of a few picometers. A (strain gauge and a vibration sensor using a (17,0) carbon nanotube (-20000 atoms) has been proposed. 2908

Thermally-induced positional uncertainty in the displacement of nanoscale components is approximated by the classical value

\[ \Delta x = \frac{(kT/k_s)^{1/2}}{10^{12}} \quad (\text{Eqn. 4.8}) \]

where T is temperature (K) and k_s is the Hooke’s law spring constant (N/m). At 310 K, this classical approximation is accurate to within \( \pm 10\% \) of the quantum mechanical treatment for harmonic oscillators with RMS displacements

\[ \Delta x \approx 10^{10} \text{ pm} \].

Spring stiffness k_s is -0.1 N/m for nonbonded (noncovalent) interatomic interactions, -30 N/m for covalent bond angle bending, -400 N/m for covalent bond stretching, and -1000 N/m for solid 1 nm^3 diamondoid blocks. 10 At 310 K, springs of such stiffness produce minimum displacement uncertainties of 200 pm, 10 pm, 3 pm and 2 pm, respectively.

The log ratio of detection energy to noise energy, the signal/ noise ratio (SNR), for a displacement sensor is derived from the harmonic potential (1/2) k_s x^2 as

\[ \text{SNR} = \ln \left( \frac{\text{signal energy}}{\text{noise energy}} \right) = \ln \left( \frac{k_s \Delta x^2}{2kT} \right) \quad (\text{Eqn. 4.9}) \]

Assuming a very stiff k_s = 600 N/m, a minimal SNR = 1 gives a minimum detectable displacement of 6 pm, or 10 pm at a more reasonable SNR = 2 (20 dB). The conservative conclusion is that \( \Delta x_{\text{min}} \approx 10 \text{ pm} \) displacements should be reliably detectable by medical nanosensors in vivo, in a measurement time \( t_{\text{meas}} \approx 10^{-9} \text{ sec} \) (Section 4.3.2). This compares favorably with the stereocilia of the inner ear, which can only detect 100 pm displacements in \( 10^{-5} \text{ sec} \). 446

#### 4.3.2 Velocity and Flow Rate Sensors

Typical thermal velocities of nanoscale diamondoid components at 310 K, from Eqn. 3.3, are -60 m/sec for 1 nm^3 objects, -2 m/sec for (10 nm)^3 objects, 0.06 m/sec for (100 nm)^3 objects, and 0.02 m/sec for 1 micron^3 objects. These thermal motions are in degrees of freedom whose range of motion is largely confined, by design, to sub-nanometer distances (Section 4.3.1) within nanoscale measuring devices. However, objects of nanomedical interest may have velocities ranging from \(-10^{-9} \text{ m/sec} \) (Section 9.4.4.2) to \(-10^{-3} \text{ m/sec} \) (e.g., the speed of sound in water), and may include biological elements in the environment or internal machine components. Thus it is useful to explore the fundamental limitations involved in measuring the full range of this physical variable.

There are at least two simple ways to determine a velocity. First, velocity may be measured by transferring the kinetic energy of a moving object or fluid volume to a sensor element. The movement of the sensor element is then clocked as it traverses a known distance. Second, and more efficiently, the passage of the moving object may sequentially trip two latches located a known distance apart, again revealing the moving object’s velocity; alternatively, detecting the presence or absence of a protruding feature on a passing body using probe rods (e.g., Fig. 4.5) can determine velocity without any requirement for large kinetic energy transfers. The need for physical gating movements by sensor components and computational register-shift operations using diamondoid rods moving at GHz clocking speeds 10 suggests a conservative minimum stable sensor cycle time of \(-10^{-9} \text{ sec} \) (Section 10.1).

The two principal restrictions on the maximum detectable velocity \( v_{\text{max}} \) are the maximum linear dimension of the sensor element, or \( L_{\text{sensor}} \) and the minimum event time that may be accurately discriminated, or \( t_{\text{min}} \). For a sensor element with thermal noise energy \( \Delta E = 10kT = 43 \text{ eV} \) at 310 K, the quantum mechanical limit is \( \Delta t_{\text{min}} \approx \hbar / (2 \Delta E) = 10^{-15} \text{ sec} \). \( \Delta t_{\text{min}} \) is also limited by the thermal variation in length of the individual latches. If the minimum detectable latch displacement is \( \Delta x = 10 \text{ pm} \) (Section 4.3.1), and the maximum speed of latch displacement is approximately the speed of sound in the latch material (e.g., \( v_{\text{sound}} = 17300 \text{ m/sec for diamond} \), then \( \Delta t_{\text{min}} \approx \Delta x / v_{\text{sound}} \approx 10^{-15} \text{ sec}, near the quantum mechanical limit. \( \Delta t \) is further restricted by the minimum detectable phase variance between latches restricted at either end of the sensor. An acoustically-transmitted clock pulse emitted from a centrally-located source that is \( \Delta x \) closer to one of the two latches can discriminate a phase variance of \( \pm \Delta \phi = \Delta x / L_{\text{sensor}} = 10^{-5} \text{ rad} \) taking \( L_{\text{sensor}} = 1 \text{ micron} \), so the minimum discriminable time is again \( \Delta t_{\text{min}} = 10^{-5} L_{\text{sensor}} / \Delta v_{\text{sound}} \approx 10^{-15} \text{ sec} \); hence \( v_{\text{max}} = \Delta x / t_{\text{meas}} = 10^{-9} \text{ m/sec} - c = 3 \times 10^8 \text{ m/sec} \) (the speed of light).

However, unlike photons, physical objects moving in an aqueous medium are normally limited to the speed of sound in that medium, e.g., \( v_{\text{sound}} = 1500 \text{ m/sec} \) in water at 310 K (Table 6.7). If \( v_{\text{max}} \) is taken as \(-10^{-3} \text{ m/sec} \), then for a sensor of size \( L_{\text{sensor}} = 1 \text{ micron} \) the required time discrimination is only \( \Delta t_{\text{min}} = L_{\text{sensor}} / v_{\text{max}} = 10^{-9} \text{ sec} \).

The minimum detectable velocity \( v_{\text{min}} \) is limited by the minimum measurable linear displacement \( \Delta x \approx 10 \text{ pm} \) and the limits of our patience in making the measurement. For a maximum measurement time of \( t_{\text{meas}} = 1 \text{ sec} \), \( v_{\text{min}} = \Delta x / t_{\text{meas}} = 0.01 \text{ nm/sec} \) in vacuo. In a nanomedically-relevant aqueous medium, the systematic

*Crystal dislocations,* 3042 high-speed bullets, detonation waves and other nanomedically-relevant phenomena can exceed the local speed of sound, but supersonic operations are hardly conservative and a turbulent medium adds additional sources of error to the measurement.
motion of a free-floating (untethered) object of radius \( R \) is masked by added Brownian noise as defined in Eqn 3.1, giving the limiting \( \nu_{\text{Brownian}} = \Delta X / \tau \approx 260 \text{ nm/sec} \) for \( R = 10 \text{ micron} \) and \( \nu_{\text{Brownian}} \approx 2.6 \text{ micron/sec} \) for \( R = 100 \text{ nm} \), where \( \nu = \tau/\text{meas} = 1 \text{ sec} \). The thermal motions of tethered objects moving along surfaces (e.g., vesicles transported along microtubules by kinesin motors; Section 9.4.6) are more restricted, hence Brownian velocity masking is reduced.

Modest accuracy may be achieved using acoustic reflection in a Doppler velocimeter. Employing acoustic waves of frequency \( \nu \), the minimum measurable velocity change \( \Delta \nu = \nu_{\text{sound}} / (\nu / \text{meas}) = 150 \text{ microns/sec} \) in an aqueous medium at 310 K with \( \nu_{\text{sound}} = 1500 \text{ m/sec} \), \( \nu \approx 10 \text{ MHz} \), and \( \text{meas} = 1 \text{ sec} \), corresponding to the shifted frequency \( \nu (1 + (\Delta \nu / \nu_{\text{sound}})) \) having enough time to make one extra cycle during \( \text{meas} \). Even at such long wavelengths (~150 microns), significant backscattering of ultrasound from ~7.82-micron red cells is observed experimentally.

Velocity sensors may also be used to estimate volumetric flow rates in fluidic channels and blood vessels by measuring the speed at which fluid passes a nanodevice resting on the vessel wall. The parabolic velocity profile of \( \nu \) is a function of radial distance \( r \) from the center of a cylindrical vessel of radius \( R \) for Hagen-Poiseille flow (Section 9.2.5):

\[
\nu = \frac{(R^2 - r^2) \Delta P}{4 \eta L_v}
\]

where absolute viscosity \( \eta = 1.1 \text{ centipoise} (1.1 \times 10^{-3} \text{ kg/m-sec}) \) for plasma at 310 K and \( \Delta P \) is the pressure change along the length of a vessel of length \( L_v \). A 1-micron nanodevice obtains measurements of \( \nu \) at various distances from the vessel wall, then fits these data to Eqn. 4.10 to find the choice of \( R \) that minimizes the variation in estimated \( \Delta P / L_v \), which should be roughly constant under nonaccelerative conditions. Once \( R \) is determined, \( \Delta P / L_v \) is known as well as the cross-sectional area (\( \pi R^2 \)) of the vessel lumen, and a simple integration of the radial velocity distribution provides the flow rate in \( \text{m}^3/\text{sec} \). Unfortunately, the usefulness of this technique is limited by boundary layer effects (Section 9.4.2.6).

4.3.3 Acceleration Sensors

Accelerations inside nanoscale devices are often measured in the trillions of \( g \)-s (\( g = 9.81 \text{ m/sec}^2 \)). For example, an \( L = 1 \text{ nm} \) diamondoid block (\( m = 3.5 \times 10^{-24} \text{ kg} \)) with thermal energy \( -kT \) experiences random thermal accelerations of \( a_{\text{thermal}} - kT / \text{g m L} = 10^{11} \text{ g/s} \); a 10-nm logic rod with vibrational energy \( -kT \) experiences \( 10^{12} \text{ g} \) oscillations; a 200-nm wide hoop spinning axially with 1000 \text{ m/sec} rim speed has rotational acceleration of \( -10^{12} \text{ g/s} \).

4.3.3.1 Box-Spring Accelerometers

As one member of a class of devices, a simple nanoscale omnidirectional accelerometer is conceptualized as a reaction mass suspended by three pairs of reaction springs (one pair per orthogonal directional accelerometer is conceptualized as a reaction mass) and compresses others, which is measured by force or displacement sensors attached to each spring. The signal/noise ratio for this sensor is

\[
\text{SNR} = \ln \left( \frac{E_{\text{spring}}}{E_{\text{noise}}} \right) = \ln \left( \frac{m^2 a^2}{2 kT k_s} \right)
\]

For a sensor mass \( m = 10^{-17} \text{ kg} \) (~78 nm-wide platinum* or ~140 nm-wide diamond cube) suspended with \( k_s = 1 \text{ N/m} \) floppy springs at \( T = 310 \text{ K} \), the minimum detectable acceleration \( a_{\text{min}} = 1.6 \times 10^6 \text{ g/s} \) for \( \text{SNR} = 1 \) and \( 2.6 \times 10^6 \text{ g/s} \) for \( \text{SNR} = 2 \) (undamped resonant acceleration ~ \( L \text{ k/m} - 10^9 \) g/s for \( L = 100 \text{ nm} \)). Undamped resonance frequency \( \nu_{\text{res}} = (k/m)^{1/2} = 3 \times 10^8 \text{ radians/sec} \), which is measurable (Section 4.3.4.3).

Sensor mass exchanges reciprocally with acceleration sensitivity in Eqn. 4.11, so the lower limit even for a dedicated micron-scale platinum-core box-spring accelerometer is \( a_{\text{min}} = 10 \text{000} \text{ g/s} \) at \( \text{SNR} = 2 \) (undamped resonant acceleration ~ \( L \text{ g/s} \) for \( L = 2 \text{ microns} \)). The maximum detectable acceleration requires \( m = 10^{-23} \text{ kg} \) (~1.4 nm diamond cube) and \( k_s = 1000 \text{ N/m} \), giving \( a_{\text{min}} = 5 \times 10^3 \text{ g/s} \) at \( \text{SNR} = 1 \).

If the accelerometer depicted in Figure 4.4 is installed near the center of mass of a nanorobot, the sensor may avoid confusing centrifugal and translation forces. Otherwise, information from multiple sensors must be combined to obtain an unambiguous measurement. Multiple measurements can improve the SNR for constant accelerations of long duration.

In 1998, ~100-nm vibrating-beam silicon accelerometers were under development by Analog Devices and Daimler-Benz. Tunneled accelerometers had been fabricated with ~10^-9 g resolution at a 350 Hz bandwidth or ~10^-8 g/Hz^{1/2} 3041 Silicon-based capacitive accelerometers (5-1000 g full scale) manufactured by Silicon Designs (Issaquah, WA) were widely used in automotive airbag safety systems; in 1998, Motorola shipped its ten-millionth accelerometer, and the global market for automotive silicon accelerometers was expected to reach $463 million by 2002.

---

*Platinum is probably the ideal choice for medical nanodevice components requiring maximum density. Depleted uranium, a relatively abundant element, is more common in macroscale nonmedical devices. However, pure \(^{238}\text{U}^{100}\) has a density of 19,050 kg/m³, somewhat lower than the density of Pt (21,450 kg/m³) or even gold (19,320 kg/m³). Uranium and its compounds are highly toxic, and finely divided (e.g., micron-scale) uranium metal is pyrophoric (spontaneously combustible) in air. By comparison, Pt metal is almost completely inert in air and in the body, thus is nontoxic, and is mechanically ~50 times stronger than gold (Table 9.3). Only two elements have higher density than Pt: Os (22,480 kg/m³) and Ir (22,420 kg/m³). Osmium is brittle even at high temperatures and the powdered metal oxidizes in air, giving off osmium tetroxide which has a strong smell and is highly toxic (as little as 100 nanograms per m³ of air causes lung congestion, skin and eye damage). Iridium metal is also very brittle and, like osmium, is at least five times less abundant than Pt.\(^{102,362}\)
4.3.3.2 Displacement Accelerometers

Formed elements in the human bloodstream (e.g., red cells, white cells, platelets) typically are buffeted by much smaller accelerations. For example at 310 K, instantaneous random thermal accelerations of \( a_{\text{thermal}} \approx 500 \text{ g's} \) are experienced by a 0.2-micron virus particle, \(-0.05 \text{ g's} \) by a 2-micron platelet, \(-10^{-4} \text{ g's} \) by a 10-micron neutrophil—and \(-0.4 \text{ g's} \) by a 1 micron¹ spherical nanorobot. Force-based sensors cannot easily detect such low accelerations, since the smaller the sensor, the larger the \( a_{\text{min}} \) for a given SNR (Eqn. 4.11). An alternative approach uses displacement sensors to determine an object's velocity twice in rapid sequence, allowing acceleration to be computed from the difference.

Consider a constantly accelerating object that triggers three clocking latches at positions \( x_1, x_2, \) and \( x_3, \) in sequence. Time and position of the object are measured as \( (x_1, t_1), (x_2, t_2), \) and \( (x_3, t_3). \) The measured constant acceleration is then given by

\[
a = \frac{v_{32} - v_{21}}{t_{32} - t_{21}} \quad \text{(m/sec}^2) \quad \text{[Eqn. 4.12]}
\]

where \( v_{32} = (x_3 - x_2) / (t_3 - t_2), \) \( v_{21} = (x_2 - x_1) / (t_2 - t_1), \) \( t_{32} = (t_3 + t_2) / 2 \) and \( t_{21} = (t_2 + t_1) / 2. \) If adjacent latches are an equal distance \( x_{\text{latch}} \) apart so that \( x_{\text{latch}} = (x_3 - x_2) = (x_2 - x_1), \) and taking \( t_1 = 0, \) then Eqn. 4.12 reduces to

\[
a = \frac{2 x_{\text{latch}}}{t_1 (t_3 - t_2)} \quad \text{(m/sec}^2) \quad \text{[Eqn. 4.13]}
\]

Assume that a slowly accelerating external object passes three clocking latches that are fixed to the exterior of a nanorobot of radius \( R. \) Measurements of the object's velocity incur an unavoidable error due to nanorobot thermal motion. The object is observed to pass between latch pairs during a measurement time \( t_{\text{meas}}. \) During this same interval, from Eqn. 3.1 the nanorobot translates a distance \( \Delta X = (kT t_{\text{meas}}/3 \pi \eta R)^{1/2}, \) incurring a maximum velocity measurement error of \( -\Delta X / t_{\text{meas}} \) and thus from Eqn. 4.12 a maximum acceleration measurement error of \( -2 \Delta X / t_{\text{meas}}^2 \) if \( (t_3 - t_2) - (t_2 - t_1) \) \(- t_{\text{meas}}. \) Thus the minimum detectable acceleration is

\[
a_{\text{min}} = \frac{2 x_{\text{latch}}}{t_1 (t_3 - t_2)} \quad \text{(m/sec}^2) \quad \text{[Eqn. 4.14]}
\]

Taking \( T = 310 \text{ K}, \eta = 1.1 \times 10^{-3} \text{ kg/m-sec} \) and \( R = 1 \text{ micron}, \) then \( a_{\text{min}} \approx 10^{-7} \text{ g's} \) for \( t_{\text{meas}} = 1 \text{ sec}; a_{\text{min}} \approx 0.004 \text{ g's} \) for a more reasonable \( t_{\text{meas}} = 10^{-3} \text{ sec measurement time in the in vivo environment.} \)

Onboard reference oscillators with the requisite accuracy are described in Section 10.1.

4.3.3.3 Fluid Acceleration Sensors

Biological fluid accelerations may also be inferred from measurements of spatial pressure gradients. Ignoring frictional and gravitational forces and assuming laminar flow, fluid acceleration \( a_{\text{fluid}} \) is given by

\[
a_{\text{fluid}} = -\frac{\Delta P}{\rho \Delta x} \quad \text{(m/sec}^2) \quad \text{[Eqn. 4.15]}
\]

where \( \rho \) = fluid density (e.g., blood plasma is an almost incompressible Newtonian fluid, with \( \rho = 1025 \text{ kg/m}^3 \) at 310 K) and \( \Delta P \) is the measured change in pressure over a distance \( \Delta x. \) The mitral and aortic valves in the heart \( (\Delta P / \Delta x = 10^{-2} \text{ N/m}^2), a_{\text{fluid}} = 10^{-2} \text{ g's} \) (a few cm), as well as valves in the veins and lymphatics, operate as bistable analog decelerometers employing this principle.⁴⁴⁴ Microscale medical nanorobots traveling across blood vessel surfaces could use pressure sensors capable of \( 10^{-6} \text{ atm resolution} \) (Section 4.5.1) to infer local blood flow accelerations to an accuracy of \( \pm 0.01 \text{ g} \) from sequential measurements taken 1 mm apart.

4.3.3.4 Pivoted Gyroscopic Accelerometers

Consider a cylindrical gyroscope spinning at an angular velocity \( \omega \) (radians/sec), with radius \( r \) and thickness \( h, \) whose center of gravity lies a distance \( L_p \) from a frictionless pivot point along the rotational axis. The pivot point defines the center of a spherical sensor grid of radius \( R \approx (r^2 + 4h^2)^{1/2} \) having distinguishable grid elements located a distance \( \Delta x = 1 \text{ nm} \) apart. Under uniform acceleration \( a, \) the gyroscope precesses around its spin axis at an angular velocity \( \omega_a = 2 a L_p / r^2 \Delta x / R t_{\text{meas}}, \) where \( t_{\text{meas}} \) is measurement duration. Hence the minimum detectable acceleration is

\[
a_{\text{min}} = \frac{\Delta x r^2 \omega}{2 L_p R t_{\text{meas}}} \quad \text{(m/sec}^2) \quad \text{[Eqn. 4.16]}
\]

by comparing precessional motion against the grid.

In choosing an \( \omega \) with which to measure \( a_{\text{min}}, \) there is an upper and a lower limit. The upper bound is the bursting strength condition

\[
a = \omega = \omega_{\text{max}} = \sqrt{\frac{2 \sigma_a}{\rho r^2}} \quad \text{(rad/sec) \quad [Eqn. 4.17]}
\]

where \( \sigma_a \) is a safe working stress \( = 10^{10} \text{ N/m}^2 \) (−0.2 tensile strength) for diamond⁴³ and \( \rho \) is the density of the gyroscope material \( (\approx 3510 \text{ kg/m}^3 \) for diamond). The lower bound is the minimum spin angular velocity below which the pivoted gyroscope cannot spin stably about the vertical axis and begins to wobble.⁴⁴⁸ and thus ruining the measurement. The lower bound is given by

\[
\omega = \omega_{\text{min}} = \sqrt{\frac{4 a_{\text{min}}}{L_p} \left(r^2 \frac{1}{2} + \frac{h^2}{3}\right)} \frac{2 a}{\pi h \rho r^6} \quad \text{[Eqn. 4.18]}
\]

If \( \omega_{\text{max}} / \omega_{\text{min}} > 1, \) then nonwobbling spin velocities are available below the bursting speed of the gyroscope cylinder. Since this ratio scales with \( R, \) there is a minimum sensor size \( R_{\text{min}} \) below which no useful spin velocities are available. Combining Eqns. 4.16 and 4.18 and solving the quadratic in \( \omega \) gives

\[
R_{\text{min}} = \frac{2 L_p}{\pi h t_{\text{meas}}^2 (2 \sigma_a / \rho)^{1/2}} \quad \text{[Eqn. 4.19]}
\]

For \( t_{\text{meas}} = 10^{-3} \text{ sec}, R_{\text{min}} \approx 69 \text{ microns (optimum} r = 48 \text{ microns,} \) \( h = 25 \text{ microns); for} t_{\text{meas}} = 1 \text{ sec,} R_{\text{min}} \approx 6.9 \text{ microns (optimum} r = 4.8 \text{ microns,} h = 2.5 \text{ microns).} \) Thus it appears that pivoted gyroscopic accelerometers likely will be employed only in devices >> 10 microns in size (but see Section 4.3.4.1). Microscale gyroscopes implemented in silicon have been discussed by Greiff⁴⁸³ and others.⁴⁴⁷

4.3.3.5 Accelerative Onset

Aside from pure acceleration, the rate of accelerative onset—sometimes called jolt, surge, or jerk—can produce additional damaging effects on macroscale mechanical and biological structures. For example, the effects of accelerative onset (à) on the human body were studied by Col. John Paul Stapp,⁴⁷ in a series of aggressive rocket sled experiments conducted during the 1950s.⁴⁷ For
human subjects, cardiovascular shock was entirely absent for $a \equiv 500$-600 g/sec but began to appear at $a \geq 1100$-1400 g/sec, at 1-40 g accelerations. Human tolerance to linear decelerative force is determined both by rate of onset and peak acceleration; the endurance limit for well-restrained healthy young males has been given as $\leq 500$ g/sec and $\leq 50$ g for $0.2$ sec durations.\(^{1,17}\)

By 1998, the effects of accelerative onset on nanomechanical structures had not been extensively studied, but similar limits might apply to nanomachines. For example, a diamond rod of length $L_{rod} = 1$ micron, cross-sectional area $S_{rod} = 0.01$ micron^2, mass $m_{rod} = 3.5 \times 10^{-17}$ kg, and working strength $W_{rod} \leq 5 \times 10^{10}$ N/m^2 (Table 9.3) fails when subjected to a static acceleration $a_{static} \geq W_{rod} S_{rod} / m_{rod} = 10^{12}$ g. But accelerative shocks can produce dynamic amplification in the response spectrum of elastic systems. Dynamic stress rises to a peak of roughly twice the corresponding static value when $a_{peak} \geq 4 \nu_{res} a_{max}$ for a trapezoidal pulse shape, where $\nu_{res}$ is the frequency of natural vibration of the system,\(^{363}\) because the stresses from the free oscillations of the system, induced by the changes in the acceleration of the system, can add to the static stresses from the acceleration, with damage done by the total stress. Taking $k_s = 1$ N/m for the rod described above, then $\nu_{res} = 30$ MHz (Section 4.3.3.1) and $a_{peak} = 10^{20}$ g/sec at $a_{max} = 10^{12}$ g. For comparison, the teragray nanoendotruff (Section 3.2.5) has mass $m = 10^{-15}$ kg and is spun up to $a_{max} = 5 \times 10^{11}$ g/sec in $\approx 6$ sec, giving $a \leq 5 \times 10^{17}$ g/sec << $a_{peak}$.

### 4.3.4 Angular Displacement

There are circumstances in which it is necessary for a medical nanodevice to measure, establish or maintain a preferred orientation, as for example nanorobots with a clearly defined “top” and “bottom” or nanodevices to measure, establish or maintain a preferred orientation, e.g., a nanorobot with a clearly defined “top” and “bottom”.\(^{4,48}\) The minimum radius, then spinning it slower than $\omega_{crit} = (1/4) m r^2 + (1/12) m h^2$, the moment of inertia around $\psi$, and $m = \pi r^2 h \rho$. The maximum value for $p_{\psi}$ is given by the burst-strength condition $\omega < \omega_{max}$ (Eqn. 4.17) as $p_{\psi} = I_{3} \omega$, where $I_{3} = (1/2) m r^2$, the moment of inertia around $\psi$. Hence:

$$\Delta \theta_{\text{nutate}} = \frac{\left( kT \left( r^2 + \frac{1}{3} h^2 \right) \right)^{1/2}}{2 \pi \sigma_w h^2 \tau} \quad \text{(radians)} \quad \text{[Eqn. 4.20]}$$

For $T = 310$ K, $r = 0.5$ micron, $h = 1$ micron, $\sigma_w \equiv 10^{-10}$ N/m^3, $\Delta \theta_{\text{nutate}} = 0.8$ microradian.

Second, the gimbal bearings have small frictional losses, exerting small torques on the gyro and causing a small precession away from the original orientation. The angular velocity of precession $\omega_{precess} = - I_{gimbals} / (I_{3} \omega)$ (rad/sec) where $I_{gimbals}$ is the frictional torque caused by an imperfect gimbal bearing. Assume that the gimbal is driven by external forces to oscillate at some frequency $\nu_{gimbals}$ and that a gimbal bearing dissipates power at the rate of $P_{gimbals}$ (watts), so that each oscillatory motion of the gimbal dissipates $E_{gimbals} = P_{gimbals} / \nu_{gimbals}$ (joules). If the angular amplitude of the gimbal motion is $\alpha_{gimbals}$ (radians) and the gimbal radius is $r_{gimbals}$, then the gimbal bearing travels $X_{gimbals} = \alpha_{gimbals} r_{gimbals}$ (meters) per gimbal oscillation and the force applied in dissipating an energy $E_{gimbals}$ is $F_{gimbals} = E_{gimbals} / X_{gimbals}$. Thus, the torque applied by each gimbal oscillation event is $T_{gimbals} = F_{gimbals} r_{gimbals} = P_{gimbals} / \nu_{gimbals} / (r_{gimbals} \omega_{gimbals})$. The torque applied by each gimbal oscillation event is $\Delta \theta_{\text{osc}} = \theta_{\text{precess}} / \nu_{gimbals}$. The number of gimbal oscillations $N = \tau / \nu_{gimbals}$, where $\tau$ is elapsed time since the gyro was last calibrated. If gimbal oscillations are independent and randomly distributed, then the total change in orientation angle is very roughly $\Delta \theta_{\text{precess}} = N^{1/2} \Delta \theta_{\text{osc}}$; assuming $\nu_{gimbals} = \tau$, then:

$$\Delta \theta_{\text{precess}} = \frac{2 k_p \alpha_{gimbals} r_{gimbals} \nu_{gimbals}^{1/2}}{\pi r^2 h \rho \omega} \quad \text{(radians)} \quad \text{[Eqn. 4.21]}$$

From Drexler\(^{10}\), $k_p = 2.7 \times 10^{-14}$ watt-sec^2/m^2 as a conservative value for a small (~2 nm) stiff bearing. Random hydrodynamic flows induced by thermal fluctuations inside biological cells have a characteristic duration $t_{\text{hyst}} \equiv 10$ millisecond,\(^{1009}\) which suggests $\alpha_{gimbals} \approx 0.06$ radians (-3') from Eqn. 3.2 for an R = 1 micron in cyto nanorobot. Taking $r = 0.5$ micron, $h = 1$ micron, $\rho = 3510$ kg/m^3 for diamond (giving $m \approx 3 \times 10^{-15}$ kg), $\omega = 1 \times 10^{9}$ rad/sec (just below the $\omega_{max} = 4.8 \times 10^{9}$ rad/sec bursting speed), and assuming $\nu_{gimbals} = t_{\text{hyst}}^{-1} = 100$ Hz (to which the results are not terribly sensitive), then $\Delta \theta_{\text{precess}} = 10^{-8}$ t^{1/2} (radians). Thus a $\Delta \theta_{\text{precess}} = 1$ microradian orientation shift takes ~1 hour; a $\Delta \theta_{\text{precess}} = 4$ microradian shift takes ~1 day; and a $\Delta \theta_{\text{precess}} = 70$ microradian (~10 arcsce) shift takes ~1 year. (Total angular error $\Delta \theta_{\text{total}} = \Delta \theta_{\text{precess}} + \Delta \theta_{\text{nutate}}$.) Thus a gimbaled nanogyroscope may be sufficiently stable to serve as a readily storable, easily transportable, and highly accurate onboard orientation standard.

### 4.3.4.2 Nanopendulum Orientation Sensing

A nanopendulum may also be employed for orientational sensing. As an example of a large class of related devices, consider a simple rigid spherical nanopendulum of radius $r$ and mass $m$, with a large hemispherical bob free to swing through the entire solid angle about a central pivot point (Fig. 4.5), with a gimbaled housing to allow active avoidance of the pivot support beam. The nanopendulum bob moves along a spherical surface located far enough from a central pivot point such that $L_p = 1$ nm between the pivot point and the effective center of gravity along the rotation axis, then $\omega_{min} = 1$ micron, cross-sectional area $S_{rod} = 0.01$ micron^2, mass $m_{rod} = 3.5 \times 10^{-17}$ kg, and working strength $W_{rod} \leq 5 \times 10^{10}$ N/m^2 (Table 9.3) fails when subjected to a static acceleration $a_{static} \geq W_{rod} S_{rod} / m_{rod} = 10^{12}$ g. But accelerative shocks can produce dynamic amplification in the response spectrum of elastic systems. Dynamic stress rises to a peak of roughly twice the corresponding static value when $a_{peak} \geq 4 \nu_{res} a_{max}$ for a trapezoidal pulse shape, where $\nu_{res}$ is the frequency of natural vibration of the system,\(^{363}\) because the stresses from the free oscillations of the system, induced by the changes in the acceleration of the system, can add to the static stresses from the acceleration, with damage done by the total stress. Taking $k_s = 1$ N/m for the rod described above, then $\nu_{res} = 30$ MHz (Section 4.3.3.1) and $a_{peak} = 10^{20}$ g/sec at $a_{max} = 10^{12}$ g. For comparison, the teragray nanoendotruff (Section 3.2.5) has mass $m = 10^{-15}$ kg and is spun up to $a_{max} = 5 \times 10^{11}$ g/sec in $\approx 6$ sec, giving $a \leq 5 \times 10^{17}$ g/sec << $a_{peak}$.
Basic Capabilities • Nanosensors and Nanoscale Scanning

concentric spherical sensor grid of radius \( R = r + z_{\text{sep}} \) to minimize van der Waals interactions (Sections 3.5.1 and 9.2), with bob and sensor surfaces electrically neutral or uniformly mutually repulsive. Much like the pin cushion receptor model (Section 3.5.7.4), the sensor grid consists of a spherical array of diamondoid sensor rods \( \Delta x \approx 1 \text{ nm} \) wide that are rapidly and simultaneously extended into the evacuated sensor cavity to determine bob position during a measurement cycle of duration \( t_{\text{meas}} \). Radial sensor rod velocity should exceed typical tangential bob velocity by at least 1-2 orders of magnitude. For the example given below, a sensor stud on a 500-nm bob that is turning at \( v_{\text{turn}} = v_{\text{thermal}} = 3.6 \text{ mm/sec} \) passes a \(-1 \text{ nm}\) sensor rod in \(-280 \text{ nanosec}\); a sensor rod displacing \(-1 \text{ nm}\) in the radial direction in \(-1 \text{ nanosec}\) travels at \( v_{\text{rod}} \approx 1 \text{ m/sec} \). This nanopendulum may serve both as an orientation sensor and as a rotation rate sensor, using principles similar to vestibular (semicircular canal) mechanics.

If a simple nanopendulum spherical sensor housing with bob initially at rest in a uniform vertical gravity field \( g \) is rotated through the smallest detectable angle \( \theta_{\text{min}} \) (radians) around the center pivot, then ignoring friction, the bob’s position over the sensor grid moves a distance \( \Delta x \), hence

\[
\Delta \theta_{\text{min}} = \Delta x / r \quad \text{[Eqn. 4.22]}
\]

For \( \Delta x = 1 \text{ nm} \) and \( r = 500 \text{ nm} \), \( \Delta \theta_{\text{min}} \approx 2 \text{ milliradians} \) (-0.1 deg). For a practical sliding interface (e.g., between bob and sensor surface) of contact area \( S \) and turning velocity \( v_{\text{turn}} \), Drexler\(^{10}\) estimates frictional dissipation, primarily driven by shear-reflection drag and band-stiffness scattering, as \( P_{\text{drag}} = k_1 v_{\text{turn}}^2 S \), where \( k_1 = 400 \text{ kg/m}^2\text{sec} \), hence the energy dissipated between the bob and the nearest sensor rod (the one interacting with the bob) is \( E_{\text{drag}} = k_1 v_{\text{turn}} S \Delta x \).

Bob gravitational energy \( E_{\text{bob}} = mg r \sin \Delta \theta_{\text{min}} = mg r \Delta \theta_{\text{min}} \) for small displacements. The bob swings freely at the minimum displacement if \( E_{\text{bob}} / E_{\text{drag}} > 1 \). If \( m = 10^{-15} \text{ kg} \), \( r = 500 \text{ nm} \), and \( S = 1 \text{ nm}^2 \), then \( v_{\text{turn}} < 2.5 \text{ m/sec} \) or \( \omega < 49 \text{ megaradians/sec} \) for the bob to swing freely (vs. \( v_{\text{thermal}} = 3.6 \text{ mm/sec} \) (Eqn. 3.3) and \( \omega_{\text{res}} = 9900 \text{ rad/sec} \) for this pendulum).

The bob-up/bob-down energy differential is \(-2 \text{ m g} r = 9.8 \text{ zJ} = 2.3 \text{ kT} \) for \( m = 10^{-15} \text{ kg} \) and \( r = 500 \text{ nm} \). To distinguish opposite orientations \( (\Delta \theta = \pi) \) requires \( N_{\text{meas}} = (kT / 2 \text{ m g} r)^2 \approx 1 \) measurement for this sensor. To distinguish \( \Delta \theta_{\text{min}} \approx 2 \text{ milliradians} \) requires averaging \( N_{\text{meas}} = (kT / E_{\text{bob}})^2 \approx 190,000 \) independent measurements (of \(-0.002 \text{ kT/measurement} \) or \( t_{\text{meas}} = 190 \text{ microsec} \) using a sensor rod clock cycle of \(-1 \text{ nanosec} \). Thus, after subtracting out the effects of translation displacements as determined by linear accelerometers (Section 4.3.3), a nanodevice can measure angular displacements in its orientation to within \( \pm \Delta \theta_{\text{min}} \) in a measurement time \( t_{\text{meas}} \approx 10^{-4} \text{ sec} \). With sensor rods moving at \(-1 \text{ m/sec}, energy dissipation for each rod sliding in its housing is \(-0.01 \text{ zJ/cycle} \) (Section 10.2.1). If only \(-100 \text{ test rods} \) need to be triggered to continuously track the moving stud, then a sensor rod clock cycle of \(-1 \text{ nanosec} \) implies a continuous nanopendulum power dissipation of \(-1 \text{ pW} \); additional drag power loss between stud and contacting rod is only \( P_{\text{drag}} \approx 0.0004 \text{ pW} \).

Many other configurations of this basic device are readily conceived and may be more volume-efficient, including nested spheres or ovoids, two-dimensional rocking disk sections for in-plane rotational measurements (three devices required for monitoring all angular degrees of freedom in a rigid rotator), and multiaxial circumferential tracks or tubes containing rolling balls or sliding plugs. J. Logajan suggests that a pair of linear three-dimensional
translational accelerometers should be able to unambiguously extract the 3 translation and 3 rotation components of all possible acceleration vectors. Rotation about any axis will result in distinctly different mass displacements while accelerated translation along any axis will result in identical displacements.

4.3.4.3 Nanopendulum Tachometry

The smallest angular velocity detectable using a simple nanopendulum is given approximately by \( \omega_{\text{min}} \sim \Delta x / \tau \) milliradian/sec (\(-0.1\) deg/sec) for \( \Delta x = 1\) nm, \( \tau = 500\) nm, and \( \tau \) is the minimum time required to count one complete rotation of the sensor; the smallest \( t_{\text{max}} \sim 2 \pi / \Theta_{\text{max}} \) nanosec for \( \Theta_{\text{max}} \sim 49\) megaradians/sec—or \(-5\%\) of the bursting speed (Eqn. 4.17) for an \( \tau = 1\) micron spherical diamondoid nanodevice.

What rotational displacements will normally be encountered by an in vivo medical nanorobot? From Equation 3.2, a 1-micron device experiences instantaneous Brownian rotations of \( \Theta_{\text{tot}} = (kT / 4 \pi \eta R^2) \) radians/sec, where \( \tau = (M \omega_{\text{tot}} / 32 \pi kT N_{\text{AV}} c_{\text{H}_2\text{O}} R^3) \). The mean time between molecular collisions with the device (mostly by water molecules), from Equation 4.3, where \( c_{\text{H}_2\text{O}} = 3.3 \times 10^{28} \) molecules/m^3 for R = 0.5 microns and \( T = 310\) K, \( \tau = 6 \times 10^{-20} \) sec/collision and \( \Theta_{\text{tot}} = -7 \times 10^9 \) rad/sec, with each displacement of \( -\tau \Theta_{\text{tot}} = 4 \times 10^{-10} \) radian, none of which are detectable. However, the net of all Brownian rotational displacements is one rotation every 16 sec, \(-0.4\) radian/sec (Section 3.2.4).

The largest measurable angular displacements in vivo are caused by tumbling due to differential shear forces in the bloodstream. Normal vessel wall shear rates in physiological bloodflow range from 100-1400 radians/sec in the larger arteries to 50-4000 radians/sec in the smaller arteries and capillaries. The measured rolling velocity of white cells during vascular “tank tread” locomotion (Section 9.4.3.6) on venule walls is \(-40 \text{ micron/sec, or } 8\pi/\tau\) radians/sec, which may be encountered by nanodevices employing cytocarriage (Section 9.4.7). The human vestibular mechanism has a frequency response range of 0.048-260 radians/sec.

Thus, the maximum range of physiologically relevant rotation rates is \( 10^4 \) to \( 10^8 \) radians/sec. Since the natural harmonic frequency of a simple nanopendulum is \( \omega_{\text{res}} = \sqrt{K / m} \) radian/sec for \( m = 500\) nm, the possibility of forced oscillations and resonance need to be faced only when measuring the highest likely physiological rotation rates, in which case the addition of a damping force \( F_{\text{damp}} = m \nu_{\text{vibr}} \omega_{\text{res}} \) during resonance events should eliminate the problem. \( F_{\text{damp}} \) for critical damping peaks at \(-mg = 0.01\) pN for the \( m = 10^{-15} \) kg, \( \tau = 500\) nm case (\( \nu_{\text{vibr}} = 2 \) mm/sec).

4.4 Force Nanosensors

Piconewton forces may be applied and must be detected by medical nanodevices. Pendulum- and cantilever-based sensors may be used to probe nanoscale forces, to monitor the ambient gravitational field, to measure molecular-scale masses to single-proton accuracy, and to discriminate among molecules having different isotopic composition.

4.4.1 Minimum Detectable Force

Extremely small forces have been measured using macroscale instrumentation, particularly the atomic force microscope (AFM). In 1998, experimental forces of 1 nN in the contact regime and 1 pN in the noncontact regime were routinely measured. In 1994, D. Rugar achieved a force sensitivity of \( 5 \times 10^{-14} \) N using an AFM with a 90-nm thick cantilever at room temperature, and in 1998 Rugar reported a force resolution of \( 7 \times 10^{-14} \) N using a magnetic resonance force microscope (MRFM) with a 230-micron long, 60-nm thick, silicon cantilever. The theoretical limit on the force sensitivity of the AFM has been estimated as \( 10^{-18} - 10^{-19} \) N for temperatures near absolute zero using fiber optic interferometry or other techniques. However, many of these methods are impractical or impossible to employ in medical nanodevices.

What is the minimum force \( F_{\text{min}} \) likely to be detectable by nanorobot sensors? Several lines of reasoning produce consistent answers. For example, force may be measured by the displacement \( \Delta x \) of a beam against the known resistive pressure provided by a spring of minimum stiffness \( k_s \), that is, \( F_{\text{min}} = k_s \Delta x \). However, the energy stored in the spring must exceed thermal noise energy, with \( S/N > k_b T / 2 \). According to Equation 4.9, giving \( F_{\text{min}} = 2kT \exp(-\Delta x / 2kT) \). For \( T = 310\) K and \( S/N = 2 \), \( F_{\text{min}} = 100\) pN at \( \Delta x = 0.6 \) nm, 10 pN at 6 nm, and 1 pN at 60 nm. Indeed, the natural molecular motors (kinase, dynein, and actin-myosin) have isotropic force outputs on the order of 1-14 pN per protein molecule, which are routinely observed using macroscale cantilever spring sensors measuring total forces of 10-100 pN. Breaking antigen-antibody or receptor-ligand interactions generally requires rupture forces in the 50-300 pN range. A force of \( -160 \) pN separates biotin that is seated in its avidin receptor; \(-85\) pN pulls iminobiotin out of the avidin receptor. A force of \(-300 \) pN ruptures a hydrogen bond between two isolated water molecules, but \( >3000 \) pN are likely to be needed to rupture a single C-C covalent bond.

The rupture strength of single covalent bonds has been directly measured experimentally by AFM as, for instance, 2000 ± 300 pN for Si-C (528 zJ/molecule dissociation enthalpy, 0.185 nm bond length) and 1400 ± 300 pN for Au-S.

Force sensing may also require using measurement devices that involve sliding interfaces of two mutually inert periodic surfaces. This motion may be characterized by the potential energy of interaction, \( -0.1-1 \) zJ as determined by Monte Carlo simulations of typical diamondoid material at 0.2-0.5 nm separations. Assuming a minimum 0.5 zJ energy barrier, for small sliding contacts the typical force amplitude \( F_{\text{min}} \approx 1.7 \times 10^{10} \Delta x \) for the entire mechanism. Thus as a conservative estimate, the smallest force that is reliably detectable (and applicable) by nanodevices is probably \( -10 \) pN.

Finally, a nanosensor of dimension \( L \) able to distinguish two forces differing by \( \Delta F \) newtons has a probability of erroneous measurement of:

\[
P_{\text{err}} = \exp(-L \Delta F / 4 kT)
\]

[Eqn. 4.23]

For \( \Delta F = 10 \) pN and \( T = 310\) K, \( P_{\text{err}} = 1\% \) for \( L = 8 \) nm, or 0.01% for \( L = 16 \) nm, so reliable nanoscale picoforce sensors may be \(-10-20 \) nm in size.

4.4.2 Nanogravimeters

Research in space medicine has discovered that micron-scale elements of the immune system have extraordinary sensitivity to gravity. One study found that bone marrow-derived (B6MP102) macrophage single cells respond after 8 seconds to exposure to a \( 10^{-2} \) g hypogravity environment as indicated by increased cell spreading, a well-known marker of cell activation. Individual immune system cells placed in microgravity exhibit enhanced growth but depressed differentiation, decreased activation by concanavalin
4.4.1) is detectable by the cytoskeleton, and this force is applied to force generated by a natural molecular motor (F ~ 1 pN; Section 4.4.3). For instance, if a change in gravitational loading equivalent to the low loading, gas springs, diamagnetic traps and actively controlled electrostatic traps [J. Soreff, personal communication, 1997].

As of 1998, the exact mechanism by which the leukocyte “gravity sensor” operates remained unknown. It has been speculated that differences in densities of several organelles, such as the nucleus, the ribosomes, and the centrioles, could generate detectable pressures on the structure of the cytoskeleton under gravitational loading [468,471] mediated through mechanosensitive stretch-activated ion channels or the extracellular matrix. While indirect gravitational effects that modify the cellular environment (such as reduced sedimentation or decreased thermal convection due to hypogravity) could also be responsible, a direct interaction of the gravitational force with cellular structures [472,473] appears to be the most likely sensor mechanism. For instance, if a change in gravitational loading equivalent to the force generated by a natural molecular motor (F ~ 1 pN; Section 4.4.1) is detectable by the cytoskeleton, and this force is applied to the entire mass of a (10 micron)$^3$ cell (m ~ 10$^{-12}$ kg), then the minimum detectable gravitational acceleration is g ~ F/m ~ 1 m/sec$^2$. This suggests that a nanopendulum may be at least 100,000 times more sensitive than a traditional quartz crystal microbalance. In Figure 4.6, which is analogous to the size of the average human tissue cell (Section 8.5.1), a bob mass of density $\rho$, giving $g_{\text{min}} = \frac{kT e^{\text{SNR}}}{m L^2}$ for a bob mass of density $\rho$, and size $L^3$, $N_{\text{meas}}$ independent measurements, and assuming $g_{\text{min}} = 0.1 g$, which should be adequate to detect the onset of hypogravity.

A nanomechanical gravity sensor of size $A < \mu$m may be used to measure accurate gravitational forces. For example, the energy of a tethered mass swinging freely in a uniform gravity field is $E_g = mgL$, where $L$ is the maximum vertical amplitude of the motion. To be detectable, oscillator energy must exceed thermal noise energy, so SNR = $m g L_{\text{meas}}^{1/2}/kT$, giving $g_{\text{min}} = \frac{kT e^{\text{SNR}}}{m L^2}$. For $L$ = 1 micron, $g_{\text{min}}$ is approximately $E_g = m g L$, where $L$ is the maximum vertical amplitude of the motion. To be detectable, oscillator energy must exceed thermal noise energy, so SNR = $m g L_{\text{meas}}^{1/2}/kT$, giving $g_{\text{min}} = \frac{kT e^{\text{SNR}}}{m L^2}$ for a bob mass of density $\rho$ and size $L^3$, $N_{\text{meas}}$ independent measurements, and assuming $g_{\text{min}} = 0.1 g$, which should be adequate to detect the onset of hypogravity.

Regarding nanogravimeter design, for a pendulum of length $L$ in a gravity field $g$, the resonant period of the motion $T_{\text{res}} = 2\pi \sqrt{L/g}$ is $2 \times 10^{-3}$ sec for $L$ = 1 micron, $9 \times 10^{-3}$ sec for $L$ = 20 microns. If $\Delta t_{\text{min}} = 1$ nanosecond is the minimum detectable decrease in resonant period caused by an increase in gravity from $g$ to $g + \Delta g$, then the minimum detectable change in gravitational force is

$$\Delta g / g = 1 - \left( \frac{L}{g k_{\text{t}}^2} \right) - 10^{-9} \left( \frac{L}{g} \right)^{1/2}$$

(4.24)

for a medical nanorobot, where $k_{\text{t}} = \Delta t_{\text{min}} / 2 \pi + (L/g)^{1/2}$.

4.4.3 Single-Proton Massometer

A nanopendulum may also be used to measure molecular-scale masses. Torsion, helical spring, and vibrating rod mechanisms all work, but one of the most efficient designs for weighing the smallest masses may be the coiled suspension-spring cantilever massometer illustrated schematically in Figure 4.6, which is analogous in principle to the traditional quartz crystal microbalance. In this design, a beam of length $L_{\text{beam}}$ and circular cross-sectional radius $R_{\text{beam}}$ is hinged at one end and carries a sample holder of solid volume $V_{\text{sample}}$ at the other end. The beam is supported (a variable distance $L_{\text{spring}}$ from the hinge) by a spring consisting of a movable suspension rod of total length $l_{\text{rod}}$, circular cross-sectional radius $r_{\text{rod}}$, and spring constant $k_s = \pi^2 r_{\text{rod}}^4 E G / (G l_{\text{rod}}^3 + E \pi r_{\text{rod}}^4 l_{\text{rod}})$ using spring (rod) material of density $\rho$, Young’s modulus $E$, and shear modulus $G$. The natural resonance period for this unloaded system is $T_0 = (2\pi L_{\text{beam}}/l_{\text{spring}}) (m_0/k_s)^{1/2} = Z m_0^{1/2}$, where $Z = (2 L_{\text{beam}} / l_{\text{spring}} r_{\text{rod}})$, and $m_0 = \rho V_{\text{sample}} + \pi L_{\text{beam}} R_{\text{beam}}^2 + \pi l_{\text{rod}} r_{\text{rod}}$. Gravitational loading is negligible for suspension system mass $m \ll (k_s l_{\text{spring}} / 2 g L_{\text{beam}})^{-1} 10^{-13}$ kg; $m = 10^{-20}$ kg in the exemplar system.

After adding a test mass $\Delta m$ to the sample holder, the resonance period increases to $T_m = Z (m_0 + \Delta m)^{1/2}$, which may be detected if $T_m - T_0 \approx \Delta t_{\text{min}} - 1$ nanosecond, the minimum convenient time interval (Section 10.1). Hence the minimum detectable molecular weight $\Delta M_W$, expressed in proton masses ($m_p = 1.67 \times 10^{-27}$ kg), is

$$\Delta M_W = \Delta m/m_p = m_p^{-1} \left[ \frac{m_0^{1/2} + \Delta t_{\text{min}} / Z}{m_0} \right]^{1/2}$$

(4.25)

In Figure 4.6, beam and rod are coiled in two dimensions (although three-dimensional coiling, more difficult to depict, provides maximum compactness). $L_{\text{beam}}$ and $l_{\text{rod}}$ are chosen to minimize total coil area $A_{\text{coil}} = 2 R_{\text{beam}} L_{\text{beam}} + \pi r_{\text{rod}} l_{\text{rod}}$. Using aluminum for the rod and beam (E = 6.9 $\times$ 10$^{10}$ N/m$^2$, G = 2.6 $\times$ 10$^{10}$ N/m$^2$, $\rho = 2700$ kg/m$^3$), $L_{\text{beam}} = 1200$ nm, $l_{\text{rod}} = 1920$ nm, $r_{\text{rod}} = 0.5$ nm, $V_{\text{sample}} = 1$ nm$^3$, and $\Delta t_{\text{min}} = 10^{-9}$ sec, then $\Delta M_W = 1$ proton for $L_{\text{spring}} = 1$ nm, $\Delta M_W < 0.01$ nm, the minimum measurable displacement (Section 4.3.1). For $L_{\text{spring}} > 1$ nm, physiological 1 g accelerative loads on a 1 micron nanorobot in which the sensor is embedded produce undetectable acceleration beam displacements < $\Delta t_{\text{min}}$.

Fig. 4.6. Single-proton massometer sensor using coiled suspension-spring cantilever.

*The $t_{\text{res}}$ dependence of $k_s$ is not strictly valid for a folded structure because the forces at the end of a folded spring have a smaller mechanical advantage, so the structure acts somewhat stiffer. Further investigation is warranted. However, a large number of alternative low-stiffness structures are available including van der Waals contacts with low loading, gas springs, diamagnetic traps and actively controlled electrostatic traps [J. Soreff, personal communication, 1997].
Moving the suspension rod closer to the sample holder reduces device sensitivity. For \( L_{\text{spring}} = L_{\text{beam}} \), minimum mass difference \( \Delta M \approx 1200 \text{ protons} \), permitting moderately large molecules to be weighed in minimum time (~10^{-5} \text{ sec}). Using a very thin sensor configuration (\( h_{\text{sensor}} \approx 20 \text{ nm} \)) and coils occupying -50% of sensor cavity space, sensor dimensions are (79 nm)^2 x 20 nm giving a massometer volume \( -10^3 \text{ nm}^3 \) or -0.01% of the total volume of a 1 micron^2 nanorobot.

Sensor mass is \( 3 \times 10^{-19} \text{ kg} \), including diamondoid housing and prorated share of support mechanisms. If one measurement cycle including sample loading/unloading operations can be completed every -10 microsec, then the device processes its own mass of glucose molecules in -10 sec. Note that biological molecules weighed in air or in vacuo may lose water of hydration, whereas weighing in water presents other complications that merit further investigation.

J. Soreff points out that for a molecule on a spring, the precision of the measurement is linked to the sharpness of the resonance, which is optimized to the degree the vibrational frequency of the spring/molecular mass system is kept well below the vibrational resonance frequencies of the molecule itself. The following discussion reports the results of a simple lumped-element analysis of the effects of internal vibration on the massometer completed for the author by J. Soreff, which confirms the previously described performance estimates for the massometer.

Suppose that it is desired to weigh some mass \( M \) to a precision \( \Delta M \) by attaching it to a weak measuring spring (spring constant \( k_{\text{weak}} \)) and measuring the shift in resonant frequency. The shift in frequency due to coupling to an internal mode of the sample must be less than \((1/2) (\Delta M / M) (1 / 2 \pi) (k_{\text{weak}}/M)^{1/2}\). If the sample has some lowest internal mode frequency \( \nu_{\text{int}} \), then approximating the sample as two \( M/2 \) masses separated by \( x_{\text{strong}} \) connected by a strong spring \( (k_{\text{strong}}) \) gives \( \nu_{\text{int}} = (1/\pi) (k_{\text{strong}}/M)^{1/2}\). If this sample is attached to a weak measuring spring of length \( x_{\text{weak}} \), which is in turn attached to a rigid support, then the equations of motion are:

\[
\frac{M}{2} \omega_{\text{strong}}^2 \delta x_{\text{strong}} = - k_{\text{weak}} \delta x_{\text{weak}} + k_{\text{strong}} \delta x_{\text{strong}} - \delta x_{\text{weak}} \tag{Eqn. 4.26}
\]

\[
\frac{M}{2} \omega_{\text{weak}}^2 \delta x_{\text{weak}} = k_{\text{strong}} \delta x_{\text{weak}} - \delta x_{\text{strong}} \tag{Eqn. 4.27}
\]

Setting \( x_{\text{weak}} = e^{\delta t} \) and \( x_{\text{strong}} = R e^{\delta t} \) where \( t = \text{time}, \ s = \text{time constant}, \ R = -(1 + (2 k_{\text{weak}} / M s^2)) \), and since the angular frequencies of the two springs are \( \omega_{\text{strong}} = 2 \pi \nu_{\text{int}} = 2 (k_{\text{strong}}/M)^{1/2} \) and \( \omega_{\text{weak}} = (k_{\text{weak}}/M)^{1/2} \), then the lower frequency solution for \( \omega_{\text{weak}} \) \( < \omega_{\text{strong}} \) may be approximated by \( s = 1 \omega_{\text{weak}} (1 - \omega_{\text{weak}}^2 / 2 \omega_{\text{strong}}^2) \) and so the (undesired) fractional change in frequency due to the flexibility of the object to be weighed is:

\[
\omega_{\text{weak}} / \omega_{\text{strong}} \leq (\Delta M / M)^{1/2} \tag{Eqn. 4.28}
\]

For a molecule sample size of \( L_{\text{sample}} \), \( \nu_{\text{int}} = \nu_{\text{sound}} / L_{\text{sample}} = 300 \text{ GHz} \) for \( \nu_{\text{sound}} = 1500 \text{ m/sec} \) and \( L_{\text{sample}} = 3 \text{ nm} \); \( M = \rho \ L_{\text{sample}}^3 = 2.7 \times 10^{-15} \text{ kg} \) for \( \rho = 1000 \text{ kg/m}^3 \). To detect the mass difference of a single proton in a sample molecule, sensor \( Q = M / m_p = 16000 \); from Eqn. 4.28, the sensor may operate at a factor of \( Q^{1/2} \) below \( \nu_{\text{int}} \) or \( \nu_{\text{spring}} = \nu_{\text{int}} / Q^{1/2} = 4 \text{ GHz} \). A period of 0.25 nanosec is slightly too short to allow a phase shift of \( \Delta \nu_{\text{int}} = 1 \text{ nanosec} \), since the period is only one-quarter of this. However, a phase shift of \( \pi \) radians should accumulate in \( \Delta \nu_{\text{int}} = Q / 2 \) \( \nu_{\text{spring}} = 2 \) microsec and should be detectable by looking at magnitudes of forces. Interestingly, an SPM microprobe of circular cross-section and exponential profile with a 0.1 nm^2 tip has been proposed that could also reach -proton mass detection sensitivity at room temperature and 1 atm pressure with an eigenmode frequency of 10 GHz; an electrically-driven carbon nanotube-based resonant-beam balance demonstrated -10^{-17} kg (-10^{-16} proton or -0.01 micron^3 mass) sensitivity in 1999.

Existing surface acoustic wave (SAW) devices are chemical sensors that can measure areal mass densities as small as 80 picogram/cm^2, or 0.01 N_2 molecule per nm^2 of surface, though current devices are limited to -micron^2 areas. A proposed quartz microresonator microbalance could detect a 10^{-6} molecular monolayer, the equivalent of 5 x 10^{-7} molecule per nm^2. Detection of -400 million protons uses an MRFM vibrating rod -60 nm thick was demonstrated by Ruggeri in 1998.

### 4.4.4 Isotope Discrimination

It will often be useful in nanomedicine to distinguish molecules containing different isotopes of the same chemical elements, for example to count the number of organic molecules in a sample that contain radioactive C^{14} atoms in place of stable C^{12} atoms or to make direct assays of heavy radiometal-contaminated tissues. Nanodevices may also sort isotopically heteronuclear molecules into homonuclear fractions, for example to separate deuterocompounds from hydrochemicals or to acquire isotopically pure materials with which to construct:

- a. small, high-speed rotors with precisely balanced mass distributions;
- b. components with extremely fine size differences (Section 3.5.6);
- c. nanostructures with maximum thermal conductivity (Table 4.1), or mixed isotopes to reduce thermal conductivity (Section 6.3.4.4 (D)); or
- d. long-lived isotopically pure structures that generate no internal radioactivity, thus promoting longer device lifetimes (Chapter 13).

The most reliable sensor for discriminating among isotopes within target molecules is the single-proton massometer (Section 4.4.3). Isotopes of any chemical element differ in mass by at least -1 neutron (-protonic) increments. Thus a massometer with single-proton resolution can resolve all isotopes starting from chemically pure samples. Isobars (atoms having the same mass number) and isotones (atoms having the same number of neutrons) may be finally resolved chemically; nuclear isomers differing in mass by -0.001 amu cannot be resolved. Teragravity centrifugation (Section 3.2.5) and time-of-flight particle-beam deflection techniques (Section 4.7.1) may also prove useful.

Isotopes vary in many subtle ways due to their mass differences, although it is not yet clear how a potential isotope nanosensor might exploit some or all of these distinctions:

- A. Line Shifts — The electronic energy states of atoms or molecules depend on the reduced nuclear mass, causing a spectral line shift of ~1.0005 between hydrogen and deuterium, less among heavier isotopes, detectable spectroscopically.
- B. Hyperfine Structure — Isotopes differ in electrical quadrupole moment, nuclear spin and magnetic moment, giving rise to hyperfine structure in the optical spectra, especially in the heavier elements. For example, the magnetic moments of hydrogen and deuterium (2.79268 and 0.857387 nuclear magnetons, respectively) are distinguishable via NMR (Section 4.8.3). The NMR frequency in a 10 kilogauss magnetic field is 42.5759 MHz for H, 6.53566 MHz for D. Of course, not all isotopes are NMR sensitive.
Table 4.1 Isotopic Variation in Bulk Properties Possibly Useful in Nanomedicine

<table>
<thead>
<tr>
<th>Bulk Property</th>
<th>Units</th>
<th>Common Isotope</th>
<th>Rarer Isotope</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute density</td>
<td>kg/m³</td>
<td>997.044 (H₂O)</td>
<td>1104.45 (D₂O)</td>
<td>at 25°C</td>
</tr>
<tr>
<td>Critical density</td>
<td>kg/m³</td>
<td>325 (H₂O)</td>
<td>363 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Critical pressure</td>
<td>atm</td>
<td>218.5 (H₂O)</td>
<td>218.6 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Critical temperature</td>
<td>°C</td>
<td>374.2 (H₂O)</td>
<td>371.5 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Temperature of maximum density</td>
<td>°C</td>
<td>3.98 (H₂O)</td>
<td>11.2 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Freezing point</td>
<td>°C</td>
<td>0 (H₂O)</td>
<td>3.82 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Boiling point</td>
<td>°C</td>
<td>100 (H₂O)</td>
<td>101.42 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Triple point (solid/liquid/vapor)</td>
<td>K</td>
<td>13.96 (H₂)</td>
<td>18.73 (D₁)</td>
<td></td>
</tr>
<tr>
<td>Triple point (Ice I/Ice III/liquid)</td>
<td>K</td>
<td>251.2 (H₂O)</td>
<td>254.4 (D₂O)</td>
<td>at 2054/2181 atm (H₂O/D₂O)</td>
</tr>
<tr>
<td>Heat of fusion</td>
<td>kcal/mole</td>
<td>0.028 (H₂)</td>
<td>0.047 (D₁)</td>
<td></td>
</tr>
<tr>
<td>Heat of vaporization</td>
<td>kcal/mole</td>
<td>9.70 (H₂)</td>
<td>9.96 (D₁)</td>
<td></td>
</tr>
<tr>
<td>Heat of formation</td>
<td>kcal/mole</td>
<td>-68.32 (H₂O)</td>
<td>-70.41 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Formation reaction constant</td>
<td>--</td>
<td>41.553 (H₂O)</td>
<td>42.664 (D₂O)</td>
<td>logarithm at 25°C, liquid state</td>
</tr>
<tr>
<td>Electrochemical reaction potential</td>
<td>volts</td>
<td>0 (2H⁺ + 2e⁻ = H₂)</td>
<td>-0.044 (2D⁺ + 2e⁻ = D₂)</td>
<td>at 25°C</td>
</tr>
<tr>
<td>Thermodynamic entropy</td>
<td>cal/K</td>
<td>16.72 (H₂O)</td>
<td>18.16 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Lower limit of inflammability</td>
<td>% conc.</td>
<td>4.65% (H₂)</td>
<td>5.0% (D₁)</td>
<td></td>
</tr>
<tr>
<td>Upper limit of inflammability</td>
<td>% conc.</td>
<td>93.9% (H₂)</td>
<td>95.0% (D₁)</td>
<td></td>
</tr>
<tr>
<td>Surface tension</td>
<td>joule/m²</td>
<td>0.07725 (H₂O)</td>
<td>0.0678 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Absolute viscosity</td>
<td>kg/m-sec</td>
<td>0.001002 (H₂O)</td>
<td>0.00126 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Dielectric constant</td>
<td>--</td>
<td>80.36 (H₂O)</td>
<td>79.755 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Refractive index</td>
<td>--</td>
<td>1.33300 (H₂O)</td>
<td>1.32844 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Solubilities</td>
<td>gm NaCl/liter</td>
<td>359 (H₂O)</td>
<td>305 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Aqueous dissociation constant</td>
<td>--</td>
<td>1.79 x 10⁻¹⁵ (NH₄OH)</td>
<td>1.1 x 10⁻¹⁵ (ND₄OH)</td>
<td>Kₑ, 25°C, 0.1-0.01N</td>
</tr>
<tr>
<td>Thermal conductivity</td>
<td>watts/m-K</td>
<td>-3000 (pure C₁₂)</td>
<td>2000 (1% C₁³)</td>
<td>diamond at room temperature</td>
</tr>
<tr>
<td>Thermal diffusivity</td>
<td>cm²/sec</td>
<td>18.5 (0.1% C₁³)</td>
<td>12.3 (1% C₁³)</td>
<td>diamond at room temperature</td>
</tr>
<tr>
<td>Isotermal compressibility</td>
<td>kg/m³</td>
<td>1250 (H₂O)</td>
<td>1390 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Ionization energy</td>
<td>--</td>
<td>13.9965 (H₂O)</td>
<td>14.869 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Ion mobilities</td>
<td>(mole/liter)²</td>
<td>1.008 x 10⁻¹⁶ (H₂O)</td>
<td>1.54 x 10⁻¹⁵ (D₂O)</td>
<td>logarithm at 25°C</td>
</tr>
<tr>
<td>Molar magnetic susceptibility</td>
<td>cm³/mole</td>
<td>-12.97 x 10⁻⁶ (H₂O)</td>
<td>-12.76 x 10⁻⁶ (D₂O)</td>
<td>liquids at 20°C</td>
</tr>
<tr>
<td>Superfluidity transition temperature</td>
<td>K</td>
<td>2.12 (He¹)</td>
<td>0.003 (He¹)</td>
<td>near absolute zero</td>
</tr>
<tr>
<td>Superconductivity transition temperature</td>
<td>K</td>
<td>9 (PdH)</td>
<td>10 (PdD)</td>
<td></td>
</tr>
<tr>
<td>Reversible glass/liqu. transition, Ice Ih</td>
<td>K</td>
<td>129 (H₂O)</td>
<td>134 (D₂O)</td>
<td></td>
</tr>
<tr>
<td>Acoustic wave speed temp. dependence</td>
<td>m/sec-K</td>
<td>+2.2 (H₁)</td>
<td>+1.6 (D₁)</td>
<td></td>
</tr>
<tr>
<td>Isotope-exchange chemical reaction rates</td>
<td>m/sec</td>
<td>6.05 x 10⁻⁴⁶ (O¹⁺ + O¹⁻)</td>
<td>9.26 x 10⁻⁴⁶ (O¹⁺ + O¹⁻)</td>
<td>ozone formation</td>
</tr>
<tr>
<td>Relative crystal lattice constant</td>
<td>--</td>
<td>1 (natural Ge)</td>
<td>1.00002 (Ge¹)</td>
<td>at 27°C</td>
</tr>
<tr>
<td>Tunnelling vibrational voltage</td>
<td>millivolts</td>
<td>358 (C₁₁), 355 (C₁₁)</td>
<td>266 (C₁₁)</td>
<td>8 K</td>
</tr>
</tbody>
</table>

C. Vibrations and Rotations — Molecular spectra are especially sensitive to isotopic changes, since the quanta of vibrational and rotational energies are directly dependent upon the masses of the atoms involved, and there are also some subtle effects from differences in nuclear spin. Separations are commonly achieved by laser excitation.

D. Cross-Sections — Isotopes have different thermal neutron capture cross sections; for instance, the cross-section of C₁₂ is 3.7 x 10⁻²⁵ barns (1 barn = 10⁻²⁴ cm²) compared to 9 x 10⁻²⁴ barns for C₁³ and <10⁻²⁶ barns for C₁⁴. (Electron scattering might be better, since the electron is less likely to cause the nucleus to fission and it is generally difficult to use neutrons to probe samples on the molecular scale.) Similarly, the neutron cross-section for H₂O is 0.6 barns compared to 9.2 x 10⁻²⁴ barns for D₂O.

E. Other Properties — Isotopes differ slightly in various bulk properties, many of which could be measured in submicron-scale aliquots by nanosensors. The ability of both large mammals and micron-size photosynthetic plant cells to preferentially concentrate C₁₂ or C₁³ in organic matter, and the predicted ability of carbon nanotubes to selectively extract tritium from low-concentration H₂-T₂ mixtures, constitute proofs of principle that at least some of the bulk properties listed in Table 4.1 may become useful in achieving nanoscale isotopic separation or purification.

4.5 Pressure Sensing
This Section describes the ability of medical nanodevices to detect and measure fluid pressure variations while functioning inside the human body. Generation of acoustic signals for power transmission (Section 6.3.2), communication (Section 7.2.1), and navigation (Section 8.3.3) are considered elsewhere. Acoustic macrosensing is described in Section 4.9.1.

4.5.1 Minimum Detectable Pressure
A change in the ambient pressure in a fluid environment (ΔP) may be determined by measuring the change in volume (ΔV) of a fixed quantity of gas at constant temperature, relative to a known sensor volume at a known reference pressure. (Thermal sensors elsewhere in the nanorobot provide readings by which corrections can be made for temperature variations.) If the sensor volume consists of a cylinder of fixed cross-sectional area, the change in volume is converted into a linear displacement of a piston located at the interface between gas and fluid environment. The change in Gibbs free energy due to this displacement is given by

ΔG = -RT ln(V₂/V₁)
energy is \( \Delta G = \Delta P \Delta V \), while thermal noise is \( -kT \). Pressure sensor signal/noise ratio SNR \( \sim \ln (\Delta P / \Delta V / kT) \), so at the minimum \( \Delta G \) the minimum detectable pressure \( \Delta P_{\text{min}} \) is measured at the maximum possible volume change, or \( V_{\text{sensor}} \) (the sensor volume), given approximately by

\[
\Delta P_{\text{min}} = kT \frac{e^{\Delta P}}{V_{\text{sensor}}} \quad \text{[Eqn. 4.29]}
\]

using the classical ideal gas formulation. At SNR = 2, a (22 nm)^3 = 10^{-20} nm^3 sensor (-5 \times 10^{10} atoms or -10^{20} kg for a cube with 1 nm thick walls) can detect a minimum pressure change \( \Delta P_{\text{min}} \sim 0.03 \) atm; a (680 nm)^3 = 0.3 micron^3 sensor (-3 \times 10^{10} atoms, -6 \times 10^{-17} kg) detects \( \Delta P_{\text{min}} \sim 10^{-6} \) atm, probably near the practical limit for mobile in vivo medical nanodevices. Note that \( \Delta P_{\text{min}} \) scales approximately as the inverse cube of mean sensor dimension, and immunity to thermal noise scales exponentially with sensor volume. Current silicon micromachined pressure sensors are \(-1 \) mm^3 with \( 10^{-3} \) atm sensitivity, a factor of \( 10^{-12} \) poorer sensitivity than the theoretical minimum for their size. (A field-effect chemical sensors of size \( \sim L \) with minimum spring constant \( k_s \sim 2 kT / L^2 \) for a maximally sensitive spring having an energy of \(-kT \) when stretched by \( L \), and piston mass \( m \), is

\[
\nu_{\text{res}} = \left( \frac{1}{2\pi} \right) \left( \frac{k_s}{m} \right)^{1/2}
\]

where \( m = \rho L^2 h \) for a piston of area \( L^2 \), height \( h \), and density \( \rho \). Assuming \( h/L = 0.1 \) and \( \rho = 1000 \text{ kg/m}^3 \), \( \nu_{\text{res}} = 21 \text{ MHz} \) for an \( L = 22 \) nm sensor (\( k_s = 1.8 \times 10^{13} \text{ N/m} \)) and \( \nu_{\text{res}} = 4 \text{ kHz} \) for an \( L = 680 \) nm sensor (\( k_s = 1.9 \times 10^{18} \text{ N/m} \)). Resonance frequency may be adjusted by selecting appropriate \( k_s \) and \( m \). In practical devices resonance will occur at a slightly lower frequency than \( \nu_{\text{res}} \) and with some broadening of the peak, because \( m \) should also include the masses of the shank of the piston plus a time-varying portion of the fluid mass occupying the piston’s chamber (depending upon the time-varying position of the piston).

Consider the surface of a piston of height \( h \), area \( L^2 \) and mass \( m \) to which a pressure spike of amplitude \( P \) (N/m^2) and frequency \( \nu_{\text{driver}} \) is applied, producing a stroke length \( X_{\text{stroke}} \). The solution for the equation of motion of an undamped forced harmonic oscillator has a maximum amplitude of \( X_{\text{stroke}} = F_{\text{spike}} / (\rho \nu_0^2 \sigma m) \) where \( \sigma \approx 2 \pi \nu \), \( F_{\text{spike}} = P L^2 \) and \( m \) is defined above. The driving frequency that produces stroke length \( X_{\text{stroke}} \) is \( L \) is given by:

\[
\nu_{\text{driven}} = \left( \frac{1}{2\pi} \right) \left( \left( \frac{P}{\rho L h} \right) - 4 \pi^2 \nu_{\text{res}}^2 \right)^{1/2}
\]

For an \( L = 22 \) nm sensor driven to full throw by a minimum detectable \( \Delta P_{\text{min}} \sim 0.03 \) atm pressure spike (again taking \( h/L = 0.1 \) and \( \rho = 1000 \text{ kg/m}^3 \)), \( \nu_{\text{driven}} = 34 \text{ MHz} \); \( \nu_{\text{driven}} = 45 \text{ GHz} \) when the sensor is driven to full throw by the maximum possible pressure spike \( P_{\text{max}} = 39,000 \text{ atm} \) consistent with water remaining in the liquid state (see below). The piston may be driven at these frequencies, or slower, using triangular-wave pulse trains of pressure spikes of these magnitudes.

For an \( L = 680 \) nm sensor driven to full throw by a minimum detectable \( \Delta P_{\text{min}} \sim 10^{-6} \) atm pressure spike, \( \nu_{\text{driven}} \sim 6 \text{ kHz} \); \( \nu_{\text{driven}} \sim 1.5 \text{ GHz} \) when the sensor is driven to full throw by the maximum pressure spike \( P_{\text{max}} = 39,000 \text{ atm} \).

What is the maximum tolerable pressure spike in aqueous media? The requirement of subsonic piston motion in water gives the highest maximum: \( P_{\text{max}} < (1/2) \rho \nu_{\text{sound}}^2 / (\text{N/m}^2) \), if \( \nu_{\text{sound}} \sim 350 \text{ m/s} \) for a solid diamondoid piston and the speed of sound \( \nu_{\text{sound}} \sim 1450 \text{ m/sec} \) in water (Eqn. 4.30) for bulk modulus \( B_{\text{water}} \sim 2.2 \times 10^{10} \text{ N/m}^2 \), \( \rho_{\text{water}} \sim 1000 \text{ kg/m}^3 \), giving \( \nu_{\text{max}} \sim 39,000 \text{ atm} \). However, rapid dissipative pressure spikes \( P_{\text{max}} > 26,000 \text{ atm} \) add sufficient energy to boil 310 K water, and isothermal static compression of pure water at 310 K causes crystallization into ice VI near \(-11,500 \text{ atm} \). Additional restrictions on (and features of) high-frequency pressure transducers are described in Sections 6.3.3, 6.4.1, and 7.2.2.
4.6 Thermal Nanosensors

The ability of medical nanodevices to measure both absolute temperature and changes in temperature is crucial for monitoring in vivo physiological thermoregulatory mechanisms and intracellular energy transactions. Precision thermal sensing is also important within nanoscale devices to provide corrective input for pressure, chemical, and displacement sensors, and to improve the stability of onboard clocks. General limits of thermal sensing are considered in Section 4.6.1, followed by several proposals for specific implementations.

4.6.1 Minimum Detectable Temperature Change

A sensor system consisting of N atoms has (3N-6) internal coordinates and thus -(3N-6) oscillators, each with an average energy of -kT assuming no modes are high enough to be frozen out. However, the standard deviation of the energy of the oscillators around the mean energy is -kT, and these energy fluctuations are uncorrelated (ignoring coupling between the oscillators), so (3N-6) of them will have an average fluctuation of kT (3N-6)\(^{1/2}\). (By comparison, the mean energy is ~kT, and these energy fluctuations are uncorrelated) so (3N-6) of them will have an average fluctuation of kT (3N-6)\(^{1/2}\). As a result, an instantaneous measurement of the total energy of all N atoms gives a minimum detectable temperature change \(\Delta T_{\text{min}}\) of

\[
\Delta T_{\text{min}} / T = (3N - 6)^{1/2} - (3N_{\text{meas}} n_d V_{\text{sensor}})^{1/2}
\]

[Eqn. 4.34]

for a set of \(N_{\text{meas}}\) independent temperature measurements using a sensor of volume \(V_{\text{sensor}}\) constructed with material of atomic number density \(n_d\), which might be maximized using diamond \((n_d = 1.76 \times 10^{29} \text{ carbon atoms/m}^3)\).

Thermal equilibration time \(t_{\text{EQ}} = L^2 C_V / K_i\) (Eqn. 10.24) for a sensor of size L, of heat capacity \(C_V\) \((1.8 \times 10^6 \text{ joules/m}^3\cdot\text{K}\) for diamond) and of thermal conductivity \(K_i\) \((2000 \text{ watts/m}\cdot\text{K}\) for diamond\(466\)). Thus, \(t_{\text{EQ}} \sim 10^{15} \text{ sec}\) for a sensor of size \(L = 10 \text{ nm}\), \(10^{11} \text{ sec}\) for \(L = 100 \text{ nm}\), and \(10^9 \text{ sec}\) for \(L = 1 \text{ micron}\), so thermal sensors up to 1 micron in size should easily achieve thermal equilibration within a typical measurement cycle time \(\Delta t_{\text{min}} = 10^{-9} \text{ sec}\). Sensor measurement time \(t_{\text{meas}} = N_{\text{meas}} \Delta t_{\text{min}}\).

A (57 nm)\(^3\) sensor can detect \(\Delta T_{\text{min}} / T = 10^{-4} (\sim 31 \text{ millikelvins at } 310 \text{ K})\) in a single measurement \((N_{\text{meas}} = 1)\), with measurement time \(t_{\text{meas}} = 1 \text{ nanosec}\). A sensitivity of \(\Delta T_{\text{min}} / T = 10^{-6} (\sim 310 \text{ microkelvins at } 310 \text{ K})\) may be achieved using either a -1 micron\(^3\) sensor with a single measurement \((N_{\text{meas}} = 1, t_{\text{meas}} = 1 \text{ nanosec})\) or a (124 nm)\(^3\) sensor with \(N_{\text{meas}} = 1000\) independent measurement cycles and \(t_{\text{meas}} = 1 \text{ microsec}\). Finally, a 1 micron\(^3\) sensor can detect \(\Delta T_{\text{min}} / T = 3 \times 10^{-9} (\sim 1 \text{ microkelvin at } 310 \text{ K})\) with \(N_{\text{meas}} = 100,000\) independent sensor cycles, giving a measurement time \(t_{\text{meas}} = 100 \text{ microsec}\).

These figures are confirmed by experimental estimates of detector noise temperature, such as \(\Delta T_{\text{min}} / T = (k / L^3 C_V)^{1/2} \sim 10^{-6}\) for a silicon nitride detector\(467\) with \(C_V = 5.2 \times 10^6 \text{ joules/m}^3\cdot\text{K}\) and \(L = 1 \text{ micron}\). It has been calculated that sensitivity of \(-1 \text{ microkelvin} is theoretically possible using quartz electronic microresonators as precision thermometers.\(462,1699\) Thermocouple probes with 100 nm tips have already shown -100 microkelvin sensitivity,\(463\) tunneling thermometers capable of measuring thermolectric potential localized to atomic-scale dimensions have been proposed,\(464\) and experiments leading toward “yoctocalorimetry” were being pursued in 1998.\(928\)

For comparison, heat sensors in human skin have \(\Delta T_{\text{min}} / T \sim 10^{-4} (\sim 90 \text{ millikelvins})\), and the infrared sensor pit of the rattlesnake is sensitive to an energy intensity of -0.8 pJ/micron\(^2\) in a measurement time \(t_{\text{meas}} \sim 35 \text{ millisec}\) and achieves \(\Delta T_{\text{min}} / T \sim 3 \times 10^{-6}\). The microkelvin piston-based coiled-cylinder temperature sensor.

4.6.2 Piston-Based Temperature Sensors

Consider a coiled cylinder of cross-sectional area \(A\) filled with gas at pressure \(P\), with a piston at one end whose movement is resisted by a constant-force spring (Fig. 4.7). Increasing gas temperature from the coldest temperature at which the sensor will operate (piston at maximum ingress), \(T_0\), to some warmer temperature \(T_1\) causes the piston to move from position \(x_0\) to position \(x_1\), while the spring holds pressure constant at \(P\) inside the cylinder during the measurement. If \(\Delta x\) is the smallest measurable piston displacement and \(\Delta T = T_1 - T_0\), then the poorest accuracy is

\[
\Delta T / T_0 = \frac{x_1 - x_0}{x_0} = \frac{\Delta x}{x_0} = N_{\text{meas}}^{1/2} (N + 1)^{-1/2}
\]

[Eqn. 4.35]

because \(\Delta x\) has a thermal noise component as well,\(10\) and the number of gas molecules \(N = n_d V_{\text{sensor}}\) where \(n_d \sim 10^{28} \text{ gas molecules/m}^3\) at \(P = 1000 \text{ atm}\) (Table 10.2). Sensitivity is maximized at the largest feasible \(x_0\); for \(\Delta T / T_0 = 10^{-6} (\sim 300 \text{ microkelvins at } 310 \text{ K})\) and \(\Delta x = \Delta x_{\text{min}} = 1 \text{ nm}\) (Section 4.2.1), then \(x_0 = 1000 \text{ microns}\), \(N = 10^9\) gas molecules at 1000 atm (using the van der Waals equation of state) and 310 K taking \(N_{\text{meas}} = 1000\), giving \(V_{\text{sensor}} = 10^{-19} \text{ m}^3\) and thus a cylinder cross-sectional area of \(A = V_{\text{sensor}} / x_0 = (10 \text{ nm})^2\). There are -1000 gas molecules per nanometer of cylinder length, each traveling with mean thermal velocity \(v_t \sim 500 \text{ m/sec}\) (Eqn. 3.3). Tightly coiled into a cubical volume, the folded sensor size is \(L_{\text{sensor}} = V_{\text{sensor}}^{1/3} = (464 \text{ nm})^3\) and measurement time \(t_{\text{meas}} = N_{\text{meas}} \Delta t_{\text{min}} = 1 \text{ microsec}\). Sensor mass is \(-10^{-10} \text{ kg}\). Fig. 4.7. Two-dimensional representation of a three-dimensional micro-kelvin piston-based coiled-cylinder temperature sensor.
The coiled tube design, though not strictly necessary, is nonetheless convenient because it allows a continuous sensor element to be arbitrarily distributed through the nanorobot volume, to be concentrated in multiple specific interior regions, or to be placed inside nanorobot components that will be required to flex such as metamorphic protuberances. If instead the working gas is placed in a cubical box of side $L_{box}$ with a sensor piston of area $A_p$, any change in temperature causes the piston to move a distance $\Delta x = (L_{box}^3 / A_p) (\Delta T / T_0) - (kT_0 / k_b)^{1/2}$, where $k_b = (A_p/L_{box}^2) (N kT_0 / L_{box})$, from Drexler,\(^\text{10}\) hence

$$\Delta T / T_0 - \left( A_p / N L_{box}^2 \right)^{3/2}$$

(Eqn. 4.36)

which argues for a small piston, a large box, and a high operating pressure to achieve maximum sensitivity. For $A_p = 100 \text{ mm}^2$, $N = 10^9 \text{ molecules}$, and $L_{box} = 464 \text{ nm}$, $\Delta T / T \sim 10^{-9}$.

### 4.6.3 Thermal-Expansion Temperature Sensors

Consider a thin rod of length $x$ and width $w$, constructed as a sandwich of two dissimilar materials each of thickness $h$, having coefficients of linear expansion $\alpha_1$ and $\alpha_2$, respectively (Fig. 4.8). When heated, the two materials expand differentially to lengths of $x (1 + \alpha_1 \Delta T)$ (neglecting the quadratic and higher-order terms in $\alpha_1$), producing a cantilever deflection of $\theta = (x/h) (\alpha_1 - \alpha_2)$ (radians) from simple geometry. The minimum detectable temperature change $\Delta T_{\text{min}} = 2 \pi \Delta x_{\text{min}} / x$ or

$$\Delta T_{\text{min}} = 8 \pi h^3 w^2 \Delta x_{\text{min}} \left( \alpha_1 - \alpha_2 \right) V_{\text{rod}}^{1/2}$$

(Eqn. 4.37)

where volume of the sandwich rod $V_{\text{rod}} = 2 \text{ h w x}$. Taking $\Delta x_{\text{min}} = 10 \text{ pm}$, $\alpha_1 = 2.38 \times 10^{-5} / K$ for aluminum,\(^\text{460}\) $\alpha_2 = 8 \times 10^{-7} / K$ for diamond at $310 \text{ K}$, $h = w = 5 \text{ nm}$, and assuming sensor volume $V_{\text{sensor}} = 2 V_{\text{rod}}$ in a tightly packed three-dimensional coil configuration, $\Delta T_{\text{min}} = 1 \text{ microkelvin}$ for an $L^3 - (227 \text{ nm})^3$ sensor - 1% of the volume of a 1-micron nanorobot. Use of a suitable negative-coefficient material for $\alpha_2$, such as the ceramics zirconium tungstate\(^\text{025}\) and cordierite, can further reduce $L$ by 5%-10%.

Measuring the volume of the sensor is essentially like taking an inventory of total stored thermal energy, which has a $kT$ $N^{1/2}$ fluctuation term from adding the $N$ uncorrelated $kT$ fluctuations; hence

$$\Delta T / T_0 - \left( N_{\text{mean}} n_d V_{\text{sensor}} \right)^{-1/2}$$

(Eqn. 4.38)

Hence, to achieve $\Delta T \sim 1 \text{ microkelvin}$ at $T_0 = 310 \text{ K}$ with this sensor may require $N_{\text{meas}} \sim 47 \text{ million independent measurements}$ of $\Delta t_{\text{min}} = 1 \text{ nanosec}$ each, or $t_{\text{meas}} \sim 47 \text{ millisec}$.

Micromechanical silicon cantilever heat sensors (microcalorimeters) of -20,000 micron$^3$ volume have already achieved $\Delta T_{\text{min}} = 10 \text{ microkelvins}$ at room temperature ($\Delta T_{\text{min}} / T \sim 3 \times 10^{-8}$, responding to -1 pJ of heat in a measurement time $t_{\text{meas}} \sim 1 \text{ millisec}$).\(^\text{461}\) Thermal-expansion nanosensors might also exploit the temperature sensitivity of viscoelastic materials (e.g., modulus of relaxation).

### 4.6.4 Mechanoechemical Temperature Sensors

Since binding rates are temperature sensitive, a chemical concentration sensor using steric probes immersed in a sample containing a precisely known concentration of target ligands can be used as a temperature sensor. Similarly, enzyme reaction rates increase with a change in temperature according to

$$\Delta T = \frac{10 \ln (k_i / k_j)}{\ln (Q_{10})}$$

(Eqn. 4.39)

where $\Delta T = T_2 - T_1$, $k_1$ and $k_2$ are the reaction rate constants at temperatures $T_1$ and $T_2$, and $Q_{10}$ is the well-known temperature coefficient—the ratio of the rates of reaction measured at two temperatures $10^\circ C$ apart. $Q_{10} \sim 2$ for most enzymes, but may range up to 4 for some enzymes that catalyze reactions with unusually high activation energies. Assuming $Q_{10} \sim 4$, the $(112 \text{ nm})^3$ sensor with $\Delta c / c = 0.01(1\%)$ from Section 4.2.2 can distinguish temperatures $0.07 \text{ K}$ apart; the huge $(1.44 \text{ micron})^3$ sensor with $\Delta c / c = 10^{-5}$ (0.001%) resolves $\Delta T \sim 72 \text{ microkelvins}$. Because of their relatively volumetric inefficiency, such sensors are best used in large $\Delta T$ applications. Similar considerations apply to thermal sensors based on the temperature sensitivity of protein folding (as in heat shock protein mediators of cellular responses to a 5 K change),\(^\text{465,466}\) protein thermal contraction,\(^\text{1201}\) and thermal conformation rate constants for protein $\alpha$ helices that typically coil in $10^{-6} \text{ sec}$.\(^\text{467,2324}\) The quantum tunneling transmission coefficient for ion channels is also exquisitely temperature-sensitive\(^\text{679}\) and thus could serve as the basis of a nanoscale thermal sensor.

### 4.6.5 Spatial Thermal Gradients

Two $(200-400 \text{ nm})^3$ thermal sensors located at either end of a 1-micron nanorobot can detect temperature changes as small as $10^{-6}$ K per micron, a -1 kelvin/m spatial gradient. The maximum detectable gradient in nanomedical systems would involve adjacent measurements differing by -100$^\circ C$, a peak spatial gradient of $10^{-4}$ kelvins/m. The detection event must be completed before the two sensors can thermally equilibrate, e.g., $t_{\text{meas}} \sim L^2 C_v / K_t$ (Eqn. 10.24). If the material between the sensors is mostly diamondoid, then $t_{\text{EQ}} \sim 10^{-9} \text{ sec}$; if mostly water, then $t_{\text{EQ}} \sim 10^{-7} \text{ sec}$, for $L = 1 \text{ micron}$.

### 4.6.6 Temporal Thermal Gradients

On the timescale of nanomachine operations, typically $10^{-9} - 10^{-6}$ sec, the human body provides an essentially isothermal operating regime. However, medical nanorobots circulating in the bloodstream may encounter $\Delta T$ of 3 K during one circulation time -60 sec under normal circumstances (Section 8.4.1), a temporal thermal gradient of -0.1 kelvin/sec, or $\sim 10$ kelvin/sec at 1 mm tissue depth for a fingertip placed on a 600 K stove. Thermal nanosensors can take independent measurements differing by as much as -100$^\circ C$ in $10^{-6} \text{ sec}$ (e.g., Section 4.6.2), giving a maximum detectable thermal...
gradient of 10^{11} \text{ kelvins/sec}, far in excess of physiological rates.\textsuperscript{a} Detection of the minimum -1 microkelvin change between two measurements taken 1 sec apart probably represents the minimum practical temperature gradient, -1 microkelvin/sec. For comparison, the record-holder in the insect world is the eyeless cave beetle Speophyes lucidulus, whose antennae can detect thermal change rates as small as -3000 microkelvins/sec.\textsuperscript{b,12}

The extraordinary thermal conductivity of diamond ensures that endogenous nanorobot waste heat is rapidly conducted to the surrounding aqueous medium. For example, a 1 micron\textsuperscript{3} diamondoid nanorobot generating 10 picowatts of onboard power at its core experiences a temperature differential of -10\textsuperscript{-8} K between core and perimeter (using thermal conductivity K\textsubscript{r} = 2000 watts/m-K for diamond\textsuperscript{460}), well below minimum detectable limits.

### 4.7 Electric and Magnetic Sensing

#### 4.7.1 Electric Fields

The electric field E (volts/m) caused by a single point charge q (e.g., an electron or singly-charged ion) is given by Coulomb's law as

\[ E = \frac{q}{4 \pi \varepsilon_0 \kappa_r r^2} \]  

\{Eqn. 4.40\}

where q\textsubscript{e} = 1.60 x 10\textsuperscript{-19} coul (one charge), \varepsilon_0 = 8.85 x 10\textsuperscript{-12} farad/m (permittivity constant), \kappa_r = dielectric constant of the matter traversed by the electric field (\kappa_r = 74.31 for pure water at 310 K, 5.7 for diamond, 1 for vacuum or air), and r = distance from the charge, in meters. The field at a distance of 1 nm from a point charge is -10\textsuperscript{-9} volts/m in air, 2 x 10\textsuperscript{-10} volts/m in water. The force between two unit charges F = q\textsubscript{e} E = 230 pN in air, 3 pN in water, at 1-nm separation.

What is the magnitude of electric fields likely to be encountered in nanomedical situations? Electric fields in the human body may be generated from internal or external sources. Internally, ions and individual molecules may carry static electric fields. For example, the surface of an isolated charged amino acid (Section 3.5.5) has -1 charge/nm\textsuperscript{2}. Debye-Huckel shielding due to counterion flow in salty fluids (as are found in the human body) reduces these fields very rapidly with distance. From Eqn. 3.20, the static field from a single charged amino acid floating in human plasma falls from -6 x 10\textsuperscript{-6} volts/m at a range of 1 nm to 0.7 volts/m at r = 10 nm. In plant cells, photosynthetic “receivers” detect gradients in the electrochemical potential of the order of 1 eV across distances on the order of 1 nm, an instantaneous -10\textsuperscript{-9} volts/m.

The most important internal electrical sources are the electrochemical gradients caused by gated channel and transporter molecular pump operations at the intracellular level; muscular, membrane, digestive, and neural activity at the intercellular and organ level (see also Section 4.9.3.1); and piezoelectric fields generated at the membrane, digestive, and neural activity at the intercellular and molecular pump operations at the intracellular level; muscular, chemical gradients caused by gated channel and transporter

External sources may also contribute to detectable nanomedical electric fields within the human body, though usually to a lesser degree. Placing a hand flat between two plates charged to \( E_p = 3 \times 10^8 \text{ volts/m} \) (near the breakdown voltage for air) induces an instantaneous field of \( E - E_p / \kappa_{\text{water}} = 42,000 \text{ volts/m} \) transversely through the limb. (A DC field induces ion current flows which accumulate surface charge and cancel out the internal field after -one RC time.) Electric fields directly beneath high-tension power lines are 2,000-11,000 volts/m, or 50-1000 volts/m at a lateral displacement of 25 meters along the ground.\textsuperscript{67} Household appliances produce 200-300 volt/m fields;\textsuperscript{67} 20 amperes flowing through standard Romex copper wiring in the walls of a house produces 100-600 volts/m in the middle of a room, up to 24,000 volts/m at the wall (\( E = B / \mu_0 \kappa_c \); \kappa_c = the speed of light, B from Eqn. 4.44); static sparks from carpet walking are -10,000 volts/m, and -10\textsuperscript{-5} volts/m makes hair stand on end. Quiescent atmospheric charge is -100 volts/m, rising to 5000 volts/m during severe thunder-storms oscillating from positive to negative over a -10,000 volt/m range. Lightning bolts (-20 coul 0.002-sec discharge) produce 60,000 volt/m spikes at a distance of 10 meters from the strike, 600 volts/m at a range of 1 kilometer. Electron-hole avalanches in semiconductors are initiated by fields = 5 x 10\textsuperscript{7} volts/m.\textsuperscript{125} Penal electrocutions use -1500 volts/m. (See also Section 7.2.3.)

Direct force sensing of these electric fields is ineffective; from Eqn. 4.40, the minimum detectable field \( E_{\text{min}} = F_{\text{min}} / q = 6 \times 10^7 \text{ volts/m} \) using \( F_{\text{min}} = 10 \text{ pN} \) (Section 4.4.1) and a single-charge sensor. Particle deflection seems more efficient. Consider a collimated stream of singly-charged particles, each of mass m and initial velocity \( V_0 \) entering an evacuated test chamber of length L and experiencing a uniform electric field \( E \), causing a lateral displacement \( \Delta x = q E L^2 / 2 m V_0^2 \) measured by the sensor. For this measurement, \( \text{SNR} = -\ln (m V_0^2 / 2 k T) \), so the minimum detectable electric field is

\[ E_{\text{min}} = \frac{4 k T \Delta x_{\text{min}} \text{e}^{\text{SNR}}}{q L^2} \]  

\{Eqn. 4.41\}

For \( \Delta x_{\text{min}} = 10 \text{ pm} \) (Section 4.3.1), \( T = 310 \text{ K} \) and \( \text{SNR} = 2 \), \( E_{\text{min}} = 10,000 \text{ volts/m} \) for a sensor of size \( L = 28 \text{ nm} \), or 100 volts/m for \( L = 280 \text{ nm} \), allowing measurement of all nanomedically relevant fields that are homogeneous across the entire device. Using the larger 280-nm sensor, time of flight is -10\textsuperscript{-12} sec for electrons or -10\textsuperscript{-9} sec for U\textsuperscript{238} ions, permitting \( t_{\text{meas}} = 10^{-9} \text{ sec} \) and dissipating at least -30 pW waste heat (due to kinetic impact) in continuous operation. Time-varying electric fields can be measured up to \( t_{\text{meas}}^{-1} - \text{GHz} \) frequencies, and electric current may be determined using standard techniques because current density is scale-invariant.\textsuperscript{10}

Care must be taken to eliminate measurement error due to exogenous magnetic fields. Piezoelectric or other electromechanical transducers can convert electricity into mechanical signals (Section 6.3.5) appropriate for all-mechanical computation and control systems.

Another class of electric field detector of similar size and sensitivity is the electrosensitive gel sensor. For example, the negatively charged heparan sulfate proteoglycan network found in the secretory granule matrix expands 50% in volume within milliseconds of exposure to a -2 volt/micron field, and the response appears linear with voltage.\textsuperscript{48} Thus the length of a 400-nm long gel-filled sensor of fixed cross-sectional area should expand a minimally-detectable 10 pm (\( = \Delta x_{\text{min}} \)) in a 100 volt/m field. Similar sensitivities to small periodic fields, analogous to shark electroreceptors (maximum sensitivity -10\textsuperscript{7} volts/m,\textsuperscript{813}), are found in ion-channel embedded

\textsuperscript{a} Heating and cooling rates during ultrasonic cavitation in water exceed ~10\textsuperscript{10} kelvins/sec.\textsuperscript{625}

\textsuperscript{b} Heating and cooling rates during ultrasonic cavitation in water exceed ~10\textsuperscript{10} kelvins/sec.\textsuperscript{625}
artificial membranes that selectively amplify electric signals in noisy environments via stochastic resonance.\textsuperscript{511} Stochastic resonance in neurons permits detection of electric fields as weak as 10 volts/m.\textsuperscript{566}

The theoretical limits have been analyzed by Weaver and Astumian\textsuperscript{814} and by Block,\textsuperscript{810} who conclude that the minimum detectable static field for a hollow spherical sensor of radius $r_{sen}$ and shell membrane thickness $d_{mem}$ (capacitance $C_e = \frac{4 \pi \epsilon_0 \kappa_e}{d_{mem}}$) is

$$E_{\text{minDC}} = \frac{2}{3} \frac{kT d_{mem}}{4 \pi \epsilon_0 \kappa_e} r_{sen}^{1/2} \text{ (volts/m)} \quad \text{(Eqn. 4.42)}$$

Using $\kappa_e = 74.31$ for water at $T = 310 \, \text{K}$, $r_{sen} = 0.5 \, \text{micron}$ and $d_{mem} = 10 \, \text{nm}$ gives $E_{\text{minDC}} = 200 \, \text{volt/m}$. Increasing $r_{sen}$ to 5 microns reduces $E_{\text{min}}$ to -2 volts/m. To detect a periodic field of frequency $\nu_e$ by sampling over $t_{\text{meas}}$ cycles for a measurement time $t_{\text{meas}}$ gives

$$E_{\text{minAC}} = \frac{E_{\text{minDC}}}{(\nu_e t_{\text{meas}})^{1/2}} \text{ (volts/m)} \quad \text{(Eqn. 4.43)}$$

Using $\nu_e = 1 \, \text{MHz}$ and $t_{\text{meas}} = 1 \, \text{sec}$ gives $E_{\text{minAC}} = 0.2 \, \text{volt/m}$. Cylindrical sensors are 10-100 times more sensitive, and sensitivity may possibly be further enhanced by coupling the periodic electric potential to a Michaelis-Menten type enzyme or a similar nanomechanism embedded in the sensor shell to achieve tuning.\textsuperscript{814} Fields as small as $10^{-4}$ volts/m may be detectable by ~20-micron mammalian living cells;\textsuperscript{814} such detection may trigger biochemical observables which may be monitored or cavedropped by resident nanomedical devices. Submicron-scale electrometers capable of -0.1 electrons/Hz$^{1/2}$ (charge per unit bandwidth) have been demonstrated experimentally, and also single-electron transistors (SETs) to $10^{-4}$ electron/Hz$^{1/2}$, with the potential to reach $10^{-6}$ electron/Hz$^{1/2}$.\textsuperscript{2926} Galvanotrophic bacteria have been reported in the literature.\textsuperscript{3351}

Radiofrequency receivers for power transmission and communication in vivo are described in Sections 6.4.2 and 7.2.3. Antenna radiation patterns may also be useful in diagnostics and macrosensing applications (Section 7.2.3). Piezoelectric macrosensing of bone loads is briefly described in Section 4.9.3.3.

### 4.7.2 Magnetic Fields

The most significant source of biomagnetic fields, generated by electric charges moving as currents through the body, is neural impulses. The magnetic field generated by these currents is given by Ampere's law as

$$B = \frac{\mu_0 \kappa_m i}{2 \pi r} \quad \text{(Eqn. 4.44)}$$

where $B$ is magnetic flux density (teslas), $\mu_0$ is the permeability constant ($-4\pi \times 10^{-7}$ henry/m), relative permeability $\kappa_m$ = 1 for most biological materials and in vacuo, $i$ is electric current (-0.2 microamperes for human nerve impulses), and $r$ = distance from the current-carrying conductor. Hence $B \sim 0.04$ microtesla for $r = 1$ micron, adjacent to an axonal conductor. For comparison, magnetic fields directly beneath high-tension power lines are 10-40 microtesla, or 2-8 microtesla at a lateral displacement of 25 meters along the ground.\textsuperscript{577} Household appliances produce 1-1000 microtesla at a distance of 3 cm and 0.01-2 microtesla at a 1-meter range\textsuperscript{477} (electric shavers produce 0.4-60 microtesla at a range of 15 cm),\textsuperscript{309} 20-amp Romex wiring in the walls of a house can produce 2 microtesla in mid-room, up to 80 microtesla at the wall, though in 1998 the average strength in 98% of U.S. homes from artificial sources was 0.05-0.09 microtesla, typical exposures throughout the day were 0.1-1 microtesla, and the federal workplace standard was a maximum of 100 microtesla.\textsuperscript{1099} Magnetotactic marsh bacteria,\textsuperscript{561} birds and insects\textsuperscript{502,503,1089} can detect the geomagnetic field vector (identifying Earth's magnetic poles), which is 30-70 microtesla, depending on geographic position. Behavioral thresholds have been measured as 0.005-0.02 microtesla for pigeons and 0.001-0.01 microtesla for honeybees,\textsuperscript{815,816} and no more than 1000-2000 microteslas for human subjects.\textsuperscript{1130} Some biochemical changes have been observed in human cells exposed to ~10-100 microtesla 60 Hz fields,\textsuperscript{877,1099} although the data are controversial; it appears the fields pumip iron atoms through cell membranes, potentially causing peroxidation or other damage to occur. In 1998, transcranial magnetic stimulation was used to expedite recovery of motor skills for stroke patients, and was being explored as a treatment for depression and a means of altering patient mood. Lightning bolts ($i \sim 10,000$ amp/discharge) produce 200 microtesla spikes at a distance of 10 meters from the strike, 2 microtesla at a range of 1 kilometer.

Computer simulations of a fully heterogeneous human body divided into 37,000 (1.31 cm)$^3$ voxels produced a maximum induced current of 8.84 amp/m$^2$ (average ~1 amp/m$^2$) when the body was exposed to a uniform 60 Hz 1-tesla magnetic field.\textsuperscript{930} Patients undergoing Echo-Planar MRI scanning report "tickling" and "pain" when exposed to field changes exceeding 60 tesla/sec; ventricular fibrillation may result from myocardial current densities of 4 amp/m$^2$, which is induced at 250 tesla/sec.\textsuperscript{490}

The simplest magnetosensor is a freely pivoting permanent magnet that experiences a restoring force when the magnet's axis is not aligned with the external field. This "compass effect" is also employed by magnetotactic bacteria, which have ~0.1 micron ferromagnets in magnetosomes in their cells that allow them to detect and track the geomagnetic field vector.\textsuperscript{502,2390} Permanent magnets of cross-sectional area $A_{\text{magnet}}$ and length $l_{\text{magnet}}$ made of Alnico or Alcomax ($\rho \sim 8000 \, \text{kg/m}^3$) can provide $B_{\text{magnet}} \sim 1.4$ tesla. (In 1998, the record flux density was 13.5 tesla for dipole magnets,\textsuperscript{697} ~27 tesla for resistive magnets.\textsuperscript{1026}) The force on such a magnet immersed in an external field of flux $B$ is $F = B B_{\text{magnet}} A_{\text{magnet}} / \mu_0$, so the minimum detectable flux $B_{\text{min}}$ is

$$B_{\text{min}} = \frac{\mu_0 \kappa_m F_{\text{min}}}{B_{\text{magnet}} A_{\text{magnet}}} \quad \text{(Eqn. 4.45)}$$

If the minimum detectable force is $F_{\text{min}} = 10$ pN (Section 4.4.1), $B_{\text{min}} = 50$ microtesla for $A_{\text{magnet}} = 0.4$ micron$^2$, about right for magnetotactic bacteria, and $B_{\text{min}} = 2$ microtesla for $A_{\text{magnet}} = 2$ micron$^2$.

The limitations of force sensing may be avoided by suspending the bar magnet from its center of mass and allowing it to oscillate in vacuo about its stable equilibrium position in the external field $B$, making a simple harmonic oscillator with period

$$t_{\text{magnet}} = \frac{2\pi}{B B_{\text{magnet}} A_{\text{magnet}}} \frac{\mu_0 \kappa_m l_{\text{magnet}}}{B_{\text{magnet}} A_{\text{magnet}}}^{1/2} \quad \text{(Eqn. 4.46)}$$
where the moment of inertia $I_{\text{mag}} = m_{\text{magnet}} \left( \frac{A_{\text{magnet}}}{4} + \frac{I_{\text{magnet}}^2}{12} \right)$. Oscillator energy $E_m = B \frac{A_{\text{magnet}} I_{\text{magnet}}}{\mu_0 m_{\text{magnet}}}$ so the minimum detectable external field is

$$B_{\text{min}} = \frac{\mu_0 K_{\text{m}} kT e^{\Delta \text{SNR}}}{B_{\text{magnet}} A_{\text{magnet}} I_{\text{magnet}}} \quad \{\text{Eqn. 4.47}\}$$

For $B_{\text{magnet}} = 1.4$ tesla, $A_{\text{magnet}} = (660 \text{ nm})^2$, $I_{\text{magnet}} = 660$ nm, $T = 310$ K and $SNR = 2$, then $B_{\text{min}} = 0.1$ microtesla and $I_{\text{magnet}} = 6 \times 10^4$ sec $^{-1}$, allowing a $\pm$KHz sampling frequency. Magnet mass $m_{\text{magnet}} > 2 \times 10^{-15}$ kg. Several classes of magnetic magnets have been described in the literature.

Another class of magnetosensor detects the change in dimensions when a magnetic body is magnetized, which is called magnetostriiction. For instance, iron-cobalt ferrite expands by $\Delta x_{\text{magnet}} = 10^{-6}$ in iron and up to $40 \times 10^{-6}$ in a $-1$ tesla field, permitting detection of -0.2 tesla causes a fractional volume increase of $-10^{-6}$ in iron and up to $4 \times 10^{-6}$ in some iron-nickel (-30% Ni) alloys. A $(370 \text{ nm})^3$ magnetostriiction sensor monitored by stretch sensors with displacement sensitivity $\Delta x_{\text{magnet}} = 0.1$ nm $^{-1}$ atomic radius can in theory detect $B_{\text{magnet}} = 0.1$ microtesla. Magnetostriiction is nonlinear with applied field, and there are significant hysterisis effects which may confuse data interpretation. However, multisensor arrays combining positive- and negative-magnetostriiction materials should enhance signal extraction. Since the change in volume is small, temperature compensation may be required.

Hall effect probe microscopes have demonstrated 10 microtesla sensitivity with $350 \text{ nm}$ spatial resolution, and alternative nanoscale magnetic detectors such as direct-current SQUID (superconducting quantum interference device) magnetometers, strain field devices, and magnetic spin sensors are being developed. In 1999, microSQUID could measure magnetic moments as small as $-10^{-1}$ Bohr magnetons or $-10^{-13}$ J/tesla, roughly $7 \times 10^6$ protons or a 14-nm iron cube. Magnetoeconeophotographs using SQUID magnetometers to measure external electromagnetic fields produced by neural traffic in the brain can register $10^{-6} - 10^{-3}$ microtesla signals. Other biomagnetic measurement techniques have helped to investigate electrophysiological disturbances in heart muscle and in the brain. By 1998, microcoils as small as 100 microns in diameter had been constructed using microcontact printing of 25 micron wide silver wires, producing magnetic flux densities $>0.4$ tesla with a 10 milliamp current, switchable in $<10^{-3}$ sec.

### 4.7.3 Optical Sensing

In biosystems and nanosystems alike, the detection of an optical photon is a high-energy event. A single photon of red light at $\lambda = 700$ nm carries an energy of $280 \text{ zJ}$, a blue photon at $400$ nm conveys $500 \text{ zJ}$, and an ultraviolet (UV) photon at $200$ nm transmits $1000 \text{ zJ}$ to its receptor. These energies, ranging from $65-234$ KJ, exceed thermal noise by a wide margin, hence are all easily detectable reliably by suitable nanosensors. Note that optical photon detection does not require a nanosensor the size of one wavelength or larger—for example, 1-nm chlorophyll molecules have no difficulty absorbing 660-nm photons.

Biological systems provide numerous examples of single-photon transducers, including photopigment molecules such as chlorophyll and rhodopsin which mediate electron transfers in $\pm 1$ picosecond by undergoing conformational changes, triggering a series of biochemical events. An estimated 100-500 cell membrane channels are affected by the detection of a single optical photon. The first documented case of a physiological role for chlorophyll in animals has been found in the fish *Malacosteus niger*, which uses a chlorophyll derivative to see far-infrared light. Mammalian cells also act as indirect UV sensors (e.g., see Section 10.4.2.3), as for instance when energy absorption induces perturbations of membrane structure or a conformational change in membrane proteins, activating the JNK cascade below 300 pJ/micron and centrosomes may serve as cellular directional IR sensors in single fibroblast cells. Integration times for photon sensors are briefly treated in Section 9.4.4.

Artificial “bioelectronic” cyanine-quinone chromophore molecules -1 nm in size can detect optical photons of wavelengths $\lambda = 580-630$ nm, and can trigger and reset in 1-3 picosec (governed by the tunneling times associated with reformation of the zwitterionic states of the inputs), in theory permitting $-100$ GHz modulations to be received. Exciton-generating poly(p-phenylene vinylene) polymer switches have the potential to achieve $-1$THz switching speed, a birefringent Kerr medium also switches in a few picoseconds. Chalcogenide glass has viscosity $>10^{13}$ kg/m/sec in darkness, but $<5 \times 10^{11}$ kg/m/sec under illumination. Proposals for specific implementations of photoergic transducers (Section 6.3.6) and optical cables (Sections 6.4.3.2 and 7.2.5.2) are elsewhere.

#### 4.7.4 Particulate and High-Energy Radiation

High energy photons including x-rays ($\lambda < 2$ nm, $-10^3 \text{ zJ}$) and gamma rays ($\lambda < 0.002$ micron, $-0.1$ picojoule), and ejecta from the decay of radioactive nuclei such as $\alpha$-particles (helium nuclei, $-1$ picojoule) and $\beta$-particles (free electrons, $0.01-0.1$ picojoule) may be extremely destructive to nanomechanical systems (Chapter 13) and are difficult to reliably measure quantitatively. For example, a charged particle with 0.01 pJ of kinetic energy (K.E.) traversing a deaccelerative electrostatic field of $10^5$ volts/m requires K.E./q - 60 microns to slow to a halt. However, particle-induced visual sensations have been observed by dark-adapted astronauts in space, and Ruderfer has suggested that even solar neutrinos may be detectable by the human brain (though most physicists would find this implausible). The nucleoelectric transducer described in Section 6.3.7.1 could be adapted to serve as a directional $\alpha$ particle sensor.

#### 4.8 Cellular Bioscanning

The goal of cellular bioscanning is the noninvasive and non-destructive in vivo examination of interior biological structures. One of the most common nanomedical sensor tasks is the scanning of cellular and subcellular structures. Such tasks may include localization and examination of cytoplasmic and nuclear membranes, as well as the identification and diagnostic measurement of cellular contents including organelles and other natural molecular devices, cytoskeletal structures, biochemical composition and the kinetics of the cytoplasm.

The precision and speed of medical nanodevices is so great that they can provide a surfeit of detailed diagnostic information well beyond that which is normally needed in classical medicine for a complete analysis of somatic status. Except in the most subtle cases a malfunction in any one component of the cellular machinery (Section 8.5) normally provokes a cascade of pathological observables in many other subsystems. Detection of any one of these cascade observables, if sufficiently unique and well-defined, may provide adequate diagnostic information (see Chapter 18) to plan a proper reparative procedure.

DNA is one of the few cellular components that is regularly inspected and repaired. Most homeostatic systems adopt a more simple philosophy of periodic replacement of components regardless of functionality. Nanomedicine allows the philosophy of inspection
and repair to be extended to all cellular components (Chapters 13 and 21). The following discussion briefly describes a few of the sensory techniques that may be useful in achieving this objective.

4.8.1 Cellular Topographies

Tactile topographic scanning provides the most direct means for examining cellular structures in vivo.\(^484-486,2723,2727\) For example, ex vivo live cell scanning in air by commercially available atomic force microscopy (AFM; Section 2.3.3) using a 20-40 nm radius tip allows nondestructive feature resolutions of ~50 nm across the top 10 nm of the cell. Investigators have used the scanning tip to punch a hole in the cell, then pull back and scan the breach, observing the membrane heal itself via self-assembly in real time. Up to -KHz scanning frequencies are possible, with up to 1024 data points per scan line. The nanomechanical manipulator arm described in Section 9.3.1.4 achieves comparable tip velocities and positional accuracies.

Assuming a scan rate of ~10\(^6\) pixels/sec, a micron-scale nanodevice, once securely anchored (Section 9.4.3.3) to the surface of a (~20 micron\(^3\) human cell, could employ a tactile scanning probe to image the 0.1% of plasma membrane lying within its (1.4 micron\(^3\)) vicinity in ~2 sec to ~1 nm\(^2\) resolution (~1 mm/sec tip velocity), or ~50 sec to ~0.2 nm (e.g., atomic) resolution (~0.2 mm/sec tip velocity). Inside the cell, and again post-anchoring, an entire 6 micron\(^2\) mitochondrial surface could be imaged at atomic resolution in ~100 sec; the surface of a 100-nm length of 25-nm diameter microtubule could be atomically resolved in 0.2 sec—though of course in a living cell these structures may be changing dynamically during the scanning process. From Eqn. 5.5, continuous power dissipation of a (0.1 micron\(^3\)) scan head moving at 1 mm/sec through water is ~0.002 pW (~kT/pixel at ~10\(^6\) pixels/sec), though of course the energy cost of sensing, recording, and processing each pixel must be at least ~10 kT/pixel (ignoring the possibility of pre-computational image compression), so the total scanner power draw could be as high as 0.02-0.1 pW in continuous operation. Special scanning tips and techniques should allow topographic, roughness, elastic, adhesive, chemical, electrostatic (charge density), conductance, capacitance, magnetic, or thermal surface properties to be measured.

For larger (~0.1-1 micron) cellular components, identification and preliminary diagnosis of improper structure may be possible using measured surface characteristics which may be matched to entries in an extensive onboard library, perhaps combined with dynamic monitoring of anomalies in cross-membrane molecular traffic using chemical nanosensors. Most membranes are self-sealing, so it should also be possible to gently insert a telescoping member into the target organelle (Section 9.4.5), which member then slowly retracts and extrudes smaller probes with sensory tips in a fixed pattern and step size, allowing the acquisition of detailed internal structural and compositional information.

Small (~1-10 nm) protein-based components of the cell such as enzymes, MHC carriers, and ribosomes are self-assembling or require only modest assistance (e.g., molecular chaperones) to self-assemble. Reversible mechanical denaturation of these smaller proteins, using diamondoid probe structure to displace water molecules (thus reducing hydrophobic forces) and by using specialized handling tools akin to functionalized AFM tips and molecular clamps, may be followed by precise nondestructive amino acid sequencing (Chapter 20) to allow identification and diagnostic compositional analysis. Afterward, the protein molecule may be refolded back into its original minimum-energy conformation, possibly with the assistance of chaperone-like structures in some cases.\(^3045\)

4.8.2 Transcellular Acoustic Microscopy

Acoustic microscopy is another noninvasive scanning technique that can provide nanomedically useful spatial resolutions. Frequencies are in the GHz range, far exceeding the relaxation time of the proteoplasm. A cryogenic acoustic microscope operated at 8 GHz has demonstrated 20 nm lateral resolution in liquid helium.\(^482\) By 1998, the best resolution achieved with water as the coupling fluid has been 240 nm at a frequency of 4.4 GHz.\(^488\) Operated in water at 310 K, a nanomechanical 1.5 GHz acoustic microscope would achieve a far-field minimum lateral resolution \(\lambda/2 ~ 500 \text{ nm}, ~10^5 \text{ voxels per human cell, sufficient to locate and count all major organelles and a few intermediate-scale structures. Scanning acoustic microscopes (SAM) operated at 2 GHz in reflection mode (which allows detection of interference effects) achieve 30-50 nm resolution in the direction of the acoustical axis, and the Subtraction SAM approach reveals topographical deviations of 7.5 nm at 1 GHz.}\(^483\) It has been proposed that picosecond ultrasounds could be used to obtain an image of the cytoskeleton with detail comparable to that of conventional x-ray images of a human skeleton.\(^162\)

Power requirements are a significant constraint on acoustic reflection microscopy (echolocation). In the simplest case, consider an acoustic emitter of radius \(r_E\) which converts an input power of \(P_{in}\) into acoustic power of intensity \(I_E\leq I_{max} = 1000 \text{ watts/m}^2\) (Section 6.4.1) with efficiency \(\epsilon\%,\) conservatively taken here to be 0.50 (50%). The emitter produces a train of omnidirectional pressure pulses of amplitude \(A_p = (2 \rho v_{sound} f)^{1/2} (\text{N/m}^2)\) (Eqn. 4.53) and frequency \(v\) which travels a distance \(X_{path}\) to a target of radius \(r_T\). As the signal reaches the target, its amplitude has been reduced by \((r_E / X_{path})\) due to the \(1/r^2\) dependence of intensity in spherical waves (Section 4.9.1.5), and by \(e^{-\alpha_{tiss} X_{path}}\) by attenuation (Eqn. 4.52) with \(\alpha_{tiss} = 8.3 \times 10^{-6}\) sec/m for soft tissue.

Upon reaching the target, a fraction \(f_{reflect}\) of the signal is reflected as a point source echo; if the target surface has an acoustic impedance similar to liver tissue and the medium is similar to water, then from Eqn. 4.54 and Table 4.3, \(f_{reflect} \approx 0.05 \text{ (5\%). The echo then travels a distance } X_{path} \text{ back to a receiver (which may or may not be located near the emitter), the amplitude again losing } (r_T / X_{path}) \text{ by geometric and } e^{-\alpha_{tiss} X_{path}} \text{ by attenuation. The echo is finally detected by the receiver which can measure a pulse of minimum pressure amplitude } \Delta P_{min} \approx 10^{-8} \text{ atm (Section 4.5.1).}

Combining these relations gives the following result:

\[
P_{in} \geq \frac{\pi \Delta P_{min}^2 X_{path}^4}{2 \rho v_{sound} c \epsilon f_{reflect} \frac{r_T}{r_E} \left(1 - e^{-\alpha_{tiss} X_{path}}\right)} \quad \text{(watts)}
\]

[Eqn. 4.48]

To scan the entire interior of a (20 micron\(^3\)) cell, whether from within or without, requires \(X_{path} \approx 20 \text{ microns}. A frequency of } v = 1.5 \text{ GHz allows } v_{sound}/v = 1 \text{ micron spatial resolution. The fastest possible pulse repetition time is } X_{path}/v_{sound} \approx 13 \text{ nanosec. Assuming } r_T = r_T = 0.5 \text{ micron, } P_{in} = 7 \text{ pW or } 5 \text{ watts/m}^2, \text{ well within the safe intensity range for acoustic radiation (Section 6.4.1). At maximum safe transmitter intensity } I_{max} \text{ the longest scannable path length at } 1.5 \text{ GHz is } X_{path} \approx 46 \text{ microns.}

Power constraints are somewhat less severe in the case of acoustic transmission microscopy (acoustic tomography), a technique which requires a minimum of two physically separated components (e.g., a transmitter and a receiver). Consider an acoustic emitter radiating a series of omnidirectional pressure pulses of frequency \(v\) which travel across a tissue cell of width \(X_{path}\) to be detected by a receiver.
on the other side after a signal transit time $t_{cell} \times X_{path} / v_{cell}$, where $v_{cell}$ is the speed of sound in the cytosol. In the simplest case, assume that the cytosol is clear except for one target of interest along the path (e.g., an organelle) in which the speed of sound is $v_{target} \neq v_{cell}$. The minimum detectable difference in travel time between a signal that passes through the target and one that does not is $\Delta t$ ($10^{-9}$ sec for diamondoid systems), so the minimum resolvable target size is $r_{min} = v_{target} - v_{cell} \Delta t / 2 \Delta t (v_{target} - v_{cell})$ when $r_{min} \approx v_{cell} / v$. For soft tissue targets in water at 1.5 GHz, $r_{min} \approx 30$ microns; for tooth enamel in water (see Table 6.7), $r_{min} \approx 1$ micron. However, the use of sampling gates will permit the detection of phase shifts that are much less than the characteristic time constant of the system in incoming waves, hence the minimum thickness of objects in theory detectable by acoustic tomography can be far smaller than the values for $r_{min}$ estimated above.

Assuming acoustic tomographic power requirement $P_{in}$ in a similar manner as for echolocation, except that the transmitted signal rather than the echo is detected and assuming the signals are of approximately equal strength regardless of the path taken, gives:

$$P_{in} \geq \frac{\pi \Delta \rho_{min}^2}{2 \rho v_{cell}} \frac{\nu_4}{\pi^2} \frac{e^{z_{v_{path}}}}{e^{z_{v_{water}}}} \left(\text{watts}\right)$$

[Eqn. 4.49]

Using the same values as the previous example, $P_{in} = 1$ pW for $X_{path} \approx 45$ microns; for $X_{path} \approx 140$ microns, $P_{in} = 1000$ pW, near $I_{max}$.

In either mode of operation, data gathering is accelerated by positioning more than one receiver around the target. Sensitivity is boosted by increasing emitter size, taking multiple measurements, or by illuminating the target with some large external source located elsewhere in the tissue or even outside of the organ. However, cells and body tissues are mesoscopic “junkyards”—highly heterogeneous media which may produce large numbers of nontarget scattering events, thus increasing the difficulty of extracting signal from noise. Additional complications arise due to:

1. scattering on rough surfaces;
2. rapid pressure variations with range in the Fresnel zone or near field of the transmitted signal;
3. cytoplasmic viscosity inhomogeneities due to the asymmetric arrangement of cytoskeletal structure, granules, vacuoles, and the endomembrane system; and
4. variation in cytoplasmic Young’s modulus due to time-varying tensions in semirandomly-distributed cytosolic fibrillar elements (e.g., a relaxed or contracted state) which alter elasticity and thus the local speed of sound.

### 4.8.3 Magnetic Resonance Cytotomography

Electrostatic scanning is largely ineffective because of Debye-Huckel shielding (Section 4.7.1). Magnetic stray field probes allow resolution of 10 nm physical features, but only for materials with substantial magnetic domains—which most biological substances lack. But nuclear magnetic resonance (NMR) imaging may allow cellular tomography by creating 3-D proton (hydrogen atom) density maps. Atomic density maps of other biologically important elements have nonzero nuclear magnetic moments (including D, Li, B, C, N, O, F, Na, Mg, P, S, Cl, K, Ca, Sn, Te, Cu, and Cd) may also be compiled. For example, sodium imaging is already used clinically to assess brain damage in patients with strokes, epilepsy, and tumors.

In a hypothetical NMR cytometric nanoinstrument, a large permanent magnet is positioned near the surface of the cell or organelle to be examined. This creates a large static background magnetic field that polarizes the protons. The spatial gradient of this field establishes a unique resonant frequency, called the Larmor frequency, within each isomagnetic surface throughout the test volume. A second time-varying magnetic driver field (e.g., vibrating permanent magnet or weak rf field) is then scanned through the full range of resonant frequencies, exciting into resonance the protons (nuclear spins) in each isomagnetic surface, in turn. Depending on the sensor implementation chosen, each resonance detected may cause absorption of driver field energy, an increase in measured impedance, or even a return echo of magnetic energy (if the driver field is operated in pulse echo mode) as the excited protons relax to equilibrium in ~1 sec. The large polarizing magnet is then rotated to a new orientation, moving the isomagnetic surfaces to new positions within the test volume, and the scan is repeated. A 3-D map of the spatial proton distribution may be computed after several scan cycles.

Only those regions within a linewidth of resonance $\Delta B$ will generate an appreciable signal, and the expected spatial resolution (say, along the z-axis) $\Delta z = \Delta B / (dB/dz)$, where $dB/dz$ is the spatial flux gradient of $B$. For a polarizing field $B = 1.4$ tesla placed adjacent to a (20 micron)$^3$ cell, the cross-cell gradient $dB/dz = 7 \times 10^8$ tesla/m; placed next to a (1 micron)$^3$ organelle, $dB/dz = 1.4 \times 10^7$ tesla/m. Assuming an effective NMR linewidth $-0.002$ tesla, minimum spatial resolutions are 29 nm and 1.4 nm, respectively.

Unfortunately, the smallest reliably detectable energy in the sensor element is $-kT e^{SNR}$. The energy required to flip a single proton is $E_{flip} = 4 \pi \nu_1 L_{proton}$ where $L_{proton} = 5.28 \times 10^{-15}$ joule-sec (the quantized spin angular momentum of the proton) and the Larmor resonance frequency $\nu_1 = \gamma_{proton} B$; $\gamma_{proton}$ is the gyromagnetic ratio (4.26 $\times 10^7$ Hz/tesla for protons) and $B$ is the polarizing magnetic flux. However, the population difference between spin-up and spin-down nuclei in NMR is very small. For small energy differences, the Boltzmann distribution only allows a fraction $E_{flip} / 2kT$ to be flipped before the upper and lower state populations are made equal and the absorption disappears. In order to flip $N_{flip}$ protons, the sample volume must include at least:

$$N_{min} = 2 e^{SNR} \left(\frac{kT}{4 \pi \nu_1 L_{proton}}\right)^2$$

[Eqn. 4.50]

For $T = 310$ K, $SNR = 2$ and $B = 1.4$ tesla, $\nu_1 = 60$ MHz and $N_{min} \approx 1.7 \times 10^{11}$ protons.

When scanning a cell or organelle, the vast majority of the protons present are in water molecules (Table 3.2). Water at 310 K has a proton density $n_{water} = 6.7 \times 10^{28}$ protons/m$^3$, compared to $n_{fat} = 6.4 \times 10^{28}$ protons/m$^3$ (palmitic acid). Distinguishing fat from water thus requires a minimum sample volume of $N_{min} / (n_{water} - n_{fat})$ = 55 micron$^3$.

Proteins range from $4.4 \times 10^{28}$ protons/m$^3$ (aspartic acid minus one water) to $7.6 \times 10^{28}$ protons/m$^3$ (all leucine minus one water, average $n_{protein} = 5.6 \times 10^{28}$ protons/m$^3$; $n_{carbohydrate} = 5.8 \times 10^{28}$ protons/m$^3$ (glucose minus one water), allowing NMR cellular tomography to resolve minimum feature sizes of 2–4 microns (5–60 micron$^3$). This resolution may include intracellular structures such as the endoplasmic reticulum (rough ER ~1100 micron$^3$, smooth ER ~400 micron$^3$), the Golgi complex (~500 micron$^3$), the nucleus (~270 micron$^3$), possibly the nucleolus (~50 micron$^3$) (Table 8.17), and even large localized ferritin granule concentrations.
In 1998, the minimum sample size for microcoil (~470 micron diameter) NMR scanning was ~10,000 micron or ~10^{16} protons for a 1-minute measurement cycle. Individual T-cell vacuoles ~1 micron in diameter labeled with dextran-coated iron oxide particles have been imaged by MRI.

### 4.8.4 Near-Field Optical Nanoimaging

Electromagnetic waves of optical wavelength \( \lambda \) that interact with an object are diffracted into two components, called "far-field" and "near-field." The propagation of electromagnetic radiation over distances \( z \geq \lambda \) acts as a spatial filter of finite bandwidth, resulting in the familiar diffraction-limited resolution \( \lambda/2 \). Classical optics is concerned with this far-field regime with low spatial frequencies \( < 2/\lambda \), and conventional optical imaging will be difficult at the cellular level in vivo (Section 4.9.4). (Short-wave x-rays will damage living biological cells.) Information about the high spatial frequency components of the diffracted waves is lost in the far-field regime, so information about sub-wavelength features of the object cannot be retrieved in classical microscopy.

However, for propagation over distances \( z < 0 \), far higher spatial frequencies can be detected because their amplitudes are then of the same order as the sample (\( z \sim 0 \)). This second diffraction component is the "near-field" evanescent waves with high spatial frequencies \( > 2/\lambda \). Evanescent waves are confined to subwavelength distances from the object. Thus a localized optical probe, such as a subwavelength aperture in an opaque screen, can be scanned raster fashion in this regime to generate an image with a resolution on the order of the probe size.

The original Near-field Scanning Optical Microscope (NSOM) surpassed the classical diffraction limit by operating an optical probe at close proximity to the object. The NSOM probe uses an aluminum-coated "light funnel" scanned over the sample. Visible light emanates from the narrow end (~20 nm in diameter) of the light funnel and either reflects off the sample or travels through the sample into a detector, producing a visible light image of the sample with ~12 nm resolution at \( \lambda \sim 514.5 \text{ nm} \). Provided the distance between light source and sample is very short, about 5 nm, with signal intensities up to \( 10^{11} \) photons/sec (~50 nanowatts). This represents a resolution of ~\( \lambda/40 \). The near-field acoustic equivalent is found in the medical stethoscope, which exhibits a resolution of ~\( \lambda/100 \).

Applications include dynamical studies at video scanning rates, low-noise high-resolution spectroscopy, and differential absorption measurements. Optical imaging of individual dye molecules has already been demonstrated, with the ability to determine the orientation and depth of each target molecule located within ~30 nm of the scanned surface. Molecules with nanometer-scale packing densities have been resolved to ~0.4-nm diameters using STMs to create photon emission maps (an electric current generates visible light emanating from a point). Laser interferometric NSOMs have produced clear optical images of dispersed oil drops on mica to ~1 micron in diameter labeled with dextran-coated iron oxide particles. The thermal conductivity of water at 310 K is 0.623 watts/m-K and the energy per 500-nm photon is 400 zJ, so for 1 micron^3 of watery tissue the maximum scan rate is ~35 micron/sec/\( \lambda \), if SNR = 2, thus providing a 1-sec volumetric optical scan of an aqueous ~1-micron^3 sample volume to 1 nm^3 resolution requires a ~1 nanowatt scanner running at ~GHz bit rates, raising sample volume temperature by ~0.03 K.

NSOM permits the determination of five of the six degrees of freedom for each molecule, lacking only the optically inactive rotation around the dipole axis. In principle, it should be possible to produce ~1 nm resolution near-field optical scans of in situ protein molecules, since with atomically precise fabrication and single lines of atoms as conductors a minimum light guide (metal-dielectric-metal) is 3 atoms wide. Given a molecular laser and adequate collimation, photons passing through folded proteins will scatter according to the molecular structure. Detection of sufficient photons comprising these scattering patterns should allow the noninvasive determination of protein structure; polarized photons provide information on chirality. Absorption and fluorescent signals will be visible from phenyl rings, tryptophan, and bound cofactors such as ATP and adenine (which is fluorescent). Positions of monoclonal antibodies on virus surfaces are now identifiable experimentally using NSOM. It should be possible to map binding sites on virus and cell surfaces using fluorescently labelled antibodies. Single molecule detection has been proposed as a tool for rapid base-sequencing of DNA (Chapter 20).

Other optically-based cellular imaging techniques must be distinguished. Optical Coherence Tomography (OCT) uses Michelson interferometry to achieve ~10 micron spatial resolutions over tissue depths of 2-3 mm in nontransparent tissue in the near infrared. However, OCT requires numerous physical components not easily implemented on a micron-size detector (e.g., beam splitter, lens-grating pair, galvanometer mirror, optical prism), femtosecond pulse shaping, and illumination power levels of ~10^7 watts/m^2 >> 100 watts/m^2 "safe" continuous limit in tissue (Section 6.4.2). Bioluminescence techniques in which the light source is placed inside the tissue (e.g., a transgenic mouse with a luciferase gene in every cell of its body) has a spatial resolution limited to ~10% of the depth, or ~10 microns resolution at a ~100 micron depth. Coherent anti-Stokes Raman scattering (CARS) already permits organelle imaging in living cells and can in theory create a point-by-point chemical map of the cell using two intersecting lasers.

Three-dimensional observation of microscopic biological non-living structures by means of x-ray holography requires a high degree of spatial coherence and good contrast between target and surroundings. Good contrast may be achieved in the wavelength range between the K absorption edges of carbon (\( \lambda \approx 4.37 \text{ nm} \)) and oxygen (\( \lambda \approx 2.33 \text{ nm} \)). The "window" where carbon-containing biological objects absorb radiation efficiently but water is relatively transparent has a spatial resolution limited to ~10% of the depth, or ~10 microns resolution at a ~100 micron depth. A 50-micron diameter emitter has been tested that uses near-IR femtosecond laser pulses impinging upon a helium gas target to create a well-collimated (<1 milliradian) beam of coherent soft x-rays at a 1 KHz repetition rate producing a brightness of \( 5 \times 10^9 \) photons/mm^2-milliradian^2-sec, a peak x-ray intensity of \( 10^{10} \) watts/m^2 on the propagation axis behind the He target.

#### 4.8.5 Cell Volume Sensing

Many cellular parameters need not be measured directly in order to be detected by a medical nanodevice. Cell volume sensing is a case in point. An intracellular nanodevice can indirectly monitor changes in the volume of the cell in which it resides by one of two methods. First, measurements of the mechanical deformation of the cellular membrane, stretch-activated channels, or cytoskeletal strains...
and structural changes are quite sensitive to alterations in total cell volume. Second, concentration or dilution of the cytoplasmic environment through cell shrinkage or swelling leads to the activation of various volume-regulation responses which may be detected by the nanorobot. Changes in the concentration of soluble cytosolic proteins may nonspecifically affect enzyme activity via "molecular crowding" and (see Section 8.5.3.3). Minor changes in cell volume can cause severalfold changes in ion channel. Cellular signalling entities that have been linked to the transduction and amplification of the primary volume signal include Ca\(^{2+}\) transients, phosphoinositide turnover, eicosanoid metabolism, kinase/phosphatase systems such as JNK and p38, CAMP, and G-proteins. These signal amplification pathways may be monitored using chemical concentration sensors aboard the medical nanodevice, allowing the nanorobot to eavesdrop (Section 7.4.5.2) on the natural sensory channel traffic of the cell.

### 4.8.6 Noninvasive Neuroelectric Monitoring

In many ways the neuron is the most nanomedically important class of cell in the human body. Nanomedical applications regarding neurons and the brain are addressed in Chapter 25. The following is a brief summary of noninvasive (i.e., no axonal membrane penetration) nanotechnological methods for monitoring the electrical traffic of individual neurons. Noninvasive measurement of axonal traffic within nerve bundles will require multiple sensors and greater sensitivity to compensate for shielding by the perineurium, a tight fascicle layer. Trafficking within nerve bundles will require multiple sensors and greater sensitivity to compensate for shielding by the perineurium, a tight fascicle layer. Trafficking within nerve bundles will require multiple sensors and greater sensitivity to compensate for shielding by the perineurium, a tight fascicle layer. Sensing pathways may be monitored using chemical concentration sensors aboard the medical nanodevice, allowing the nanorobot to eavesdrop (Section 7.4.5.2) on the natural sensory channel traffic of the cell.

#### 4.8.6.1 Electric Field Neurosensing

The "typical" neuron shows that neuronal discharge is a 0.1-20 microns. By 1998 silicon-to-neuron extracellular junctions already permitted direct stimulation of individual nerve cells in vitro without killing the cells, and extracellular electrodes were commonly used to detect neuronal electrical activity noninvasively, both in vivo and in vitro. Artificial electric fields may also be employed to trigger or moderate neural signals (Section 7.4.5.6). Cell membrane capacitance is typically -0.1 picofarads/micron\(^2\), and varies with the state of the health of the cell.

#### 4.8.6.2 Magnetic Field Neurosensing

The magnetic flux density caused by a single action potential discharge is ~0.1 microtesla at the axonal surface, which may be detected by a "compass oscillator" type magnetosensor with a -KHz maximum sampling rate (Section 4.7.2).

It might be possible for artificial magnetic fields to directly influence neural transmissions. Even a static 65 millitesla field has been shown to reduce frog skin Na\(^{+}\) transport by 10-30%. Each neuronal discharge develops an electrical energy of ~20 picojoule \((-10^{-14} \text{ J})\), far smaller than the magnetic energy stored in a B = 1.4 tesla field of a permanent micromagnet traversing an L\(^3\) = (20 microcm)\(^3\) volume which from Eqn. 6.9 is B\(^2\) L\(^3\) / 2 \(\mu_0\) ~ 6000 pJ. If properly manipulated, such a field may be sufficient to enhance, modulate, or extinguish a passing neural signal.

#### 4.8.6.3 Neurothermal Sensing

While neurons come in many shapes and sizes (Chapter 25), our "exemplar" -14,000-micron\(^3\) neuron discharging -90 mV into an input impedance of ~500 Kohms produces -0.2 microampere current per pulse and generates a continuous (average) 100-300 pW of waste heat as measured experimentally (Table 6.8). The discharge rate of 5-100 Hz can produce brief surges up to ~2000 pW during a high-frequency train, but the duty cycle of such trains is far less than 100%, reducing time-averaged dissipation to the observed 100-300 pW range.

Single impulses are measured experimentally to produce 2-7 microkelvin temperature spikes in cold or room-temperature mammalian non-myelinated nerve fibers and -23 microkelvins in non-myelinated garfish olfactory nerve fibers at an energy density ranging from 270-1670 joules/m\(^3\)-impulse from 0-20°C. In non-myelinated fibers the initial heat occurs in two temperature-dependent phases: a burst of positive heat, followed by rapid heat reabsorption (called the negative heat). The positive heat derives from the dissipation of free energy stored in the membrane capacitance, and from the decrease in entropy of the membrane dielectric with depolarization. An L = 20-micron neuron in good thermal contact with an aqueous heat sink at 310 K has thermal conductance L K\(_T\) = 10\(^{-5}\) watts/K, so trains of 5-100 Hz impulses lasting 1 second should raise cellular temperature by 10-30 microkelvins; up to ~200 microkelvin thermal spikes from such trains have been observed experimentally. These events are easily detectable by nanoscale thermal sensors capable of -1 microkelvin sensitivity up to -1 KHz (Section 4.6.3). The -microkelvin heat signature of individual impulses or very short pulse trains can probably be temporally resolved because the minimum pulse repetition time is -10 millisec (at 100 Hz) which is much longer than the thermal time constant for an L = 20-micron neuron (thermal conductivity K\(_T\) = 0.528 watts/m-K and heat capacity C\(_V\) = 3.86 \times 10\(^7\) joules/m\(^3\)-K for brain tissue; Table 8.12) which is t\(_{EQ}\) = L\(^2\) C\(_V\) / K\(_T\) ~ 3 millisec.

#### 4.8.6.4 Direct Synaptic Monitoring

The synaptic cleft between the axonal presynaptic terminal and the dendritic postsynaptic membrane is 10-20 nm in most synapses, although in the vertebrate myoneural junction it may be as large as 100 nm. Contact area per bouton is ~1 micron\(^2\), giving a total gap volume of 10\(^{-10}\)-10\(^{-9}\) nm\(^3\). The density of acetylcholine receptors is highest in muscles along the cists and upper thirds of the junctional folds (-10,000/micron\(^2\)), and is lowest in the extrasympathetic regions (-5/micron\(^2\)). (Other neurotransmitters exist; Table 7.2 and Section 7.4.5.6.) Each action potential discharge triggers the release of -10\(^{10}\)-10\(^{11}\) molecules of acetylcholine into the gap volume of an active neuromuscular junction (diffusion time ~1 microsec), raising the ligand from near zero to ~3 \times 10\(^{-4}\) molecules/nm\(^3\) (-0.0005 M) in -1 millisec, followed by near-complete hydrolyzation by acetylcholinesterase during the 1.2 millisecond refractory period. Into the gap volume may easily be inserted a -10\(^{5}\) nm\(^3\) neurotransmitter concentration sensor (Section 4.2.3) able to measure -100 acetylcholine molecules in ~1 millisec (Eqn. 4.5), thus detecting pulses at the fastest discharge rate. A similar device could be used to precisely regulate neurotransmitter concentration at the junction, and hence the neural signal itself, under nanodevice control (Section 7.4.5.6). Simple electrochemical and mecanochemical artificial synapses have been demonstrated.

#### 4.8.6.5 Other Neurosensing Techniques

Many other noninvasive neurosensing techniques are readily conceivable. For instance, counting rotors (Section 3.4.2) or sodium...
magnetic resonance imaging (Section 4.8.3) could detect changes in the ionic composition (e.g., Na⁺, K⁺) of the periaxonal fluid before, during, and after discharge.

### 4.8.7 Cellular RF and Microwave Oscillations

Starting in 1968, H. Frohlich, observing that millivolt electrical potentials maintained across cell membranes ~10 nm thick give rise to huge fields ~10⁷ volts/m (Section 4.7.1) possibly producing an electret state, theorized that membrane molecules must be highly electrically polarized and thus could interact to produce coherent surface acoustic vibrational modes in the 10-100 GHz (microwave) frequency range. The longest wavelength is about twice the membrane thickness. Interestingly, this frequency span is very close to the maximum trigger/reset frequency for bioelectronic molecules (Section 4.7.3). Note that (-100 mV) (1.6 x 10⁻¹⁹ coul) ~ 4 kT, so a membrane molecule with a single charge on either end should be reorientated by a depolarization wave, coupling pressure waves and electrostatic field fluctuations. However, the direct detection of 10-100 GHz millimeter radiation by non-nanotechnological means is experimentally difficult and controversial because the tests must be performed in vivo in close proximity to an actively metabolizing cell. Therefore, this might prove diagnostic of numerous internal states. Such states may in theory be detectable both within and nearby living cells.

Nevertheless, active cells have shown enhanced Raman anti-Stokes scattering, an effect ascribed to the converse of the Frohlich oscillations. In one study, the normalized growth rate of yeast cultures was enhanced or inhibited when irradiated by CW microwave fields of ~30 watts/m² of various frequencies; growth rate data spanning 62 separate runs revealed a repeatable frequency-dependent spectral fine structure with six distinct peaks of width ~10 MHz near 42 GHz. Investigations of related phenomena are ongoing and voluminous; the interested reader should purge Bioelectromagnetics, the archival journal of this field.

From Eqn. 6.32, 100 GHz waves attenuate only ~1% after passing through ~3 microns of soft tissue. A single electron injected into an integral membrane protein could act as an oscillating dipole, making a 300 volt/m signal 10 nm from the protein antenna with an energy transfer of ~0.004 kT per cycle. 1000 oscillating electrons could produce a measurable field. A 20-micron diameter cell modeled as a nonuniform spherical dipole layer with transmembrane dipoles located 10 nm apart and embedded in a dissipative medium could produce 10⁻⁹ volt/m microwave fields 1-10 microns from the cell surface.

Thus, a variety of rf and microwave electromagnetic emanations may in theory be detectable both within and nearby living cells which could prove diagnostic of numerous internal states. Such states may include cytoskeletal dynamics, metabolic rates, plasmatic type excitations due to the collective motion of ions freed in chemical reactions, positional, rotational or conformational changes in biological macromolecules and membranes, internal movements of organelles and nerve traffic conduction, cellular pinocytosis, cellular reproduction events, cell membrane identity (e.g., distinguishing erythrocyte, Gram-positive and Gram-negative cell coat conductivities at 10 KHz), and cell-cell interactions.

### 4.9 Macrosensing

Macrosensing is the detection of global somatic states (inside the human body) and extrasomatic states (sensory data originating outside the human body). While the treatment here is necessarily incomplete, the discussion nevertheless gives a good feel for the kinds of environmental variables that internally-situated nanodevices could sense. Not all capabilities outlined here need be available on every nanobot, since injection of a cocktail of numerous distinct but mutually cooperative machine species allows designers to take full advantage of the benefits of functional specialization. In many cases, a given environmental variable can be measured by several different classes of sensor device. However, since these devices are microscopic it is in theory possible to operationalize almost all of the macrosensing capabilities described below in one patient using just a billion devices (~10 mm³ whole-body deployment density), a total volume of ~1 mm³ of nanorobots or ~0.1% of the typical ~1 cm³ therapeutic dose (Chapter 19).

A general discussion of methods for communicating macrosensory information to the human user is in Section 7.4.6, and is mentioned briefly in Section 4.9.5 below.

#### 4.9.1 Acoustic Macrosensing

##### 4.9.1.1 Cyto-Auscultation

Can sounds generated by a single cell be detected, and thus be useful for diagnosis? Probably not, given that low frequency acoustic radiators are notoriously inefficient (Section 7.2.2.1). For example, mitochondrion organelles of the giant amoeba *Reticulomyxa* are shuttled back and forth by 1-4 cytoplasmic dynein motors while riding on the outside of a bundle of 1-6 microtubules. Each dynein motor generates 2.6 pN of force and drives the mitochondria at up to ~10 micron/sec, developing 0.3-1.0 x 10⁻¹⁶ watts of mechanical power within each 320-nm diameter organelle. Taking each organelle as a cylindrical acoustic radiator with a mechanical input power of P_in ~ 10⁻¹⁶ watts at ν ~ 1000 Hz, the output acoustic pressure at the organelle surface is only ~10⁻⁹ atm (Eqn. 7.6), an acoustic power intensity 1 ~ 10⁻¹⁵ watts/m² (Eqn. 4.53 which is not detectable by micron-sized nanorobots. Nevertheless, cells and intracellular elements are capable of vibrating in a dynamic manner with complex harmonics that can be altered by growth factors and by the process of carcinogenesis, so the possibility cannot be completely ruled out.

##### 4.9.1.2 Blood Pressure and Pulse Detection

Blood pressure ranges from 0.1-0.2 atm in the arteries to as low as 0.005 atm in the veins. The systolic/diastolic differential ranges from 0.05-0.07 atm in the aorta and 0.01-0.02 atm in the pulmonary artery, falling to 0.001-0.003 atm in the microvessels, or 0.003-0.005 atm if the precapillary sphincter is dilated. In venous vessels, pulse fluctuations are 0.002-0.010 atm in the superior vena cava, 0.004-0.006 atm in the subclavian vein, ~0.004 atm in venules generally, and ~0.0005 atm in the brachial vein. There is also a ~0.05 Hz random fluctuation in the microvessels with amplitude on the order of 0.004-0.007 atm. Both blood pressure and pulse rate can be reliably monitored by a medical nanodevice virtually anywhere in the vascular system using a (68 nm)³ pressure sensor with ~0.001 atm sensitivity (Section 4.5.1). (See also Section 8.4.2.)

Pulse propagation through body tissue is somewhat muted due to absorption in compressible fatty membranes, but most cells lie within 1-3 cell-widths of a capillary so the cardiac acoustic signal should still be measurable using more sensitive detectors. The time-averaged interstitial pressure in subcutaneous tissue is 0.001-0.004 atm.

Arterial pulse waves (vascular oscillations) carry subtle messages about the health of internal organs and the arterial tree. The idea of using pulse waves for diagnosis has a long history dating back 2000 years in China. For example, in the *Book on Pulse Waves* by Wang Shu-He (201-285 AD), waves detected by manual probing are classified using such subjective and qualitative descriptors as floating, deep, hidden, rapid, slow, moderate, feeble, replete, full, thready,
turbulent flow in a tube is

\[ P_{\text{turb}} = P_{\text{lam}} Z = 8 \pi \eta_{\text{air}} v^2 L Z \]  

(watts) \hspace{1cm} \text{(Eqn. 4.51)}

where \( P_{\text{lam}} \) is the dissipation for laminar (Poiseuillean) flow in a long circular cylindrical tube of length \( L \), \( v \) is the mean flow velocity, and turbulence factor \( Z = 0.005 (N_R^{3/4} - (2300)^{3/4}) \), a well-known empirical formula. \( \alpha_{\text{air}} \) is the initial wave amplitude in atm, \( \alpha_x \) is the amplitude a \( x \) distance from the source, and \( \alpha \) is the amplitude absorption coefficient. The function \( F \) expresses the frequency dependence of the attenuation. For pure liquids, \( F = F_{\text{liq}} = v^2 \) (Hz); for example, \( \alpha_{\text{liq}} = 2.5 \times 10^{-14} \text{sec}^2/\text{m} \) for water at room temperature. However, for soft tissues, \( F = F_{\text{tiss}} \), a \( v \) (Hz). Values for \( \alpha_{\text{tiss}} \) (in sec/m) are in Table 4.2; the value for diamond was estimated from the acoustic line discussion in Section 7.2.5.3. Assuming \( 10^3 \) kHz bronchial turbulence noise of initial amplitude \( A_0 = 4 \times 10^{-5} \) atm, \( A_x = 3.7 \times 10^{-5} \) atm through 1 meter of typical soft tissue; \( A_x = 3.4 \times 10^{-5} \) atm even if 0.1 meter of bone is interposed in the acoustic path. Either \( A_0 \) is reliably detected from anywhere in the body using a \( >210 \text{ nm} \) pressure sensor (Eqn 4.29).

4.9.1.4 Mechanical Body Noises

Many other mechanical body noises should be globally audible to properly instrumented medical nanodevices. If normal chewing motions (of hard foods) release 1-10 milliwatts in a \( -100 \text{ cm}^3 \) oral volume with a \( 1 \) sec jawstroke, power density is \( 10 \text{ watts/m}^2 \) or \( 10^{-4} \) atm of tooth-crunching noise. A stomach growl registering 45 dB (vs. 30 dB whisper, 60 dB normal conversation) at 2 meters has a source power of 160 milliwatts; released from a \( 10 \text{ cm}^3 \) gastric sphincter noise gives a \( 2 \times 10^{-6} \) atm acoustic wave, detectable throughout the body. Walking and running releases 20-100 joules/footfall for a 70 kg man; assuming the energy is absorbed within a \( 1 \) cm thickness or within \( 1 \) second by the sole of the foot, Eqn. 4.53 implies an upward-moving planar compression wave of 0.4-2.0 atm, easily detectable by acoustically instrumented nanodevices body-wide. (Shoe insoles dissipate energy and alter the shock wave pulse shape.) The hand-clapping generates 0.02-0.2 atm pulses, also easily detectable.

Lesser noises including \( 30 \) millisecond hiccups at (4-60)/min, intestinal and ureteral peristalsis, sloshing of liquid stomach contents, heart murmurs, a tap on the shoulder by a friend, nasal sniffing and swallowing, clicks from picking or drumming fingernails, crepitations, manustrictures and ejaculations, the rustling noise of clothing against the skin, flapping eyelids, anal towelling, bruits (including murmurs and thrills) due to vascular lesions, dermal impact of water while showering, copulatory noises, urethral flow turbulence during urination, transmitted vibrations from musical instruments, creaking joints, and squeaking muscles can be detected locally if not globally. Implantation of significant interconnected in vivo diamondoid structures may produce increased sensitivity to internal noises, due to the extremely low acoustic absorption coefficient of diamond (Table 4.2).

4.9.1.5 Vocalizations

Average source power for conversational speech in air is \( 10 \) microwatts at the vocal cords (60 dB), up to \( 1000 \) microwatts for shouting (90 dB) and as little as 0.1 microwatts (30 dB) for whispering. Vocal cord surface area \( -1 \text{ cm}^2 \), giving an acoustic intensity \( I = 0.001-10 \text{ watts/m}^2 \). (Using the decibel notation, DB \( = 10 \log_{10} (I/I_0) \), where \( I_0 = 5 \times 10^{-13} \text{ watts/m}^2 \) in air, \( I_0 = 1 \times 10^{-16} \text{ watts/m}^2 \) in water.) In a planar traveling wave, pressure amplitude \( A_p (\text{N/m}^2) \) is related to power intensity \( I \) by

\[ A_p = \left( 2 \rho v_{\text{sound}} I \right)^{1/2} \]  

(N/m^2) \hspace{1cm} \text{(Eqn. 4.53)}

For water at 310 K, \( \rho = 993.4 \text{ kg/m}^3 \) and \( v_{\text{sound}} = 1500 \text{ m/sec} \), therefore \( A_p = 0.0005\text{-0.05 atm} \) for speech, detectable by nanodevices throughout the body due to minimal attenuation at audible frequencies (Section 4.9.1.3). Other easily detectable vocalizations include whistling, humming, coughing, sneezing, rales, wheezing,
Table 4.3 Acoustic Impedance for Specular Reflection by Acoustic Waves Crossing a Material Interface

<table>
<thead>
<tr>
<th>Body Tissue</th>
<th>Impedance (kg/m²·sec)</th>
<th>Body Tissue</th>
<th>Impedance (kg/m²·sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air</td>
<td>400</td>
<td>Liver (25°C)</td>
<td>1.65 x 10⁶</td>
</tr>
<tr>
<td>Lung</td>
<td>1.80 x 10⁶</td>
<td>Blood</td>
<td>1.65 x 10⁶</td>
</tr>
<tr>
<td>Fat</td>
<td>1.39 x 10⁶</td>
<td>Nerve (optic)</td>
<td>1.68 x 10⁶</td>
</tr>
<tr>
<td>Aqueous humor</td>
<td>1.51 x 10⁶</td>
<td>Muscle</td>
<td>1.73 x 10⁶</td>
</tr>
<tr>
<td>Water</td>
<td>1.52 x 10⁶</td>
<td>Lens of eye</td>
<td>1.84 x 10⁶</td>
</tr>
<tr>
<td>Brain (25°C)</td>
<td>1.57 x 10⁶</td>
<td>2.8 x 10⁶</td>
<td></td>
</tr>
<tr>
<td>Skin</td>
<td>1.6 x 10⁶</td>
<td>Skull bone</td>
<td>7.80 x 10⁶</td>
</tr>
<tr>
<td>Soft tissue (avg.)</td>
<td>1.63 x 10⁶</td>
<td>Enamel</td>
<td>1.71 x 10⁶</td>
</tr>
<tr>
<td>Kidney</td>
<td>1.63 x 10⁶</td>
<td>Diamond</td>
<td>6.3 x 10⁷</td>
</tr>
</tbody>
</table>

4.9.1.6 Environmental Sources

Can in vivo nanodevices directly detect sounds emanating from the environment outside of the body, such as other people talking in the same room or a door slamming? The waves from an external acoustic source of power $P_R$ watts travel through the air and, upon arriving at the air/skin interface a distance $x_R$ from the source with amplitude $A_{\text{incident}}$, are transmitted through the interface with amplitude $A_{\text{transmit}}$. For a specular reflector—interface dimensions (human body ~ 2 m) > acoustic wavelength (~0.03-3.0 m for typical ultrasonic frequencies). For $Z_{\text{air}} = 400$ kg/m²·sec and assuming $Z_{\text{skin}} = 1.6 x 10^6$ kg/m²·sec from Table 4.3, then $A_{\text{transmit}} = (5 x 10^{-4})$. For an immediately subdermal nanodevice can detect a minimum $A_{\text{transmit}} = 10^{-6}$ atm, then from Eqn. 4.53 and simple geometry:

$$p_R = \frac{2 \pi x_R^2 A_{\text{incident}}}{\rho v_{\text{sound}}} \quad \text{(Eqn. 4.55)}$$

For STP (1 atm, 0°C) air, $\rho = 1.29$ kg/m³ and $v_{\text{sound}} = 331$ m/sec. If the minimum detectable pressure $= 10^{-6}$ atm (Section 4.5.1) - $A_{\text{transmit}}$, then at a distance of $x_R = 2$ meters the acoustic source must have a power of $= 2000$ watts, far exceeding the $= 1$ milliwatt output of a person loudly shouting. To hear normal conversation at $x_R = 2$ m, minimum nanodevice detector sensitivity falls to $7 x 10^{-11}$ atm requiring a subdermal pressure nanosensor -(17 micron)³ in size (Eqn. 4.29), roughly the dimensions of a single human cell; other methods may prove more efficient (Sections 4.9.5 and 7.4.6.3).

Of course, an ex vivo acoustic nanosensor may receive sound that has passed through no interface, hence may detect pressure waves ~3 orders of magnitude lower in amplitude. Assuming $A_{\text{transmit}} = A_{\text{incident}}$, $p_R = 600$ microwatts, so ex vivo nanorobots with a 0.3 micron³ sensor could hear people shouting at $x_R = 2$ meters. To hear talking (~10 microwatt source) requires a 2.4 micron³ ex vivo pressure sensor (limit - $10^{-7}$ atm), from Eqn. 4.29.

Optimally positioned and calibrated nanomedical pressure sensors could directly measure changes in the ambient barometric pressure to within a $10^{-6}$ atm. Normal atmospheric variation due to weather ranges from 0.94-1.05 atm; such slow moving changes are readily monitored. Very near the Earth's surface, the air pressure $P$ at altitude $h$ above sea level is approximated by $P = c_{\text{baro}}^{kh}$ (atm), where $c_{\text{baro}} = 1.16 x 10^{-4}$ m⁻¹ at 20°C; at sea level, a $10^{-6}$ atm change in pressure reflects a change in altitude of only ~1 cm. However, the opening or closing of a door inside a (~5 m)³ room that displaces >125 cm³ of air also causes a minimally detectable >$10^{-6}$ atm pressure pulse. Other sources of environmental pressure variation such as infrasonic (~0.2 Hz) microbaroms from offshore ocean storms, wind entering through open windows, forced-air currents from central heating or A/C systems, or even the movements of nearby people and pets may be detectable and thus may further confuse the measurement, reducing absolute accuracy unless suitable corrections are made.

4.9.2 Proprioceptive Macrosensing

4.9.2.1 Kinesthetic Macrosensing

Using a navigational transponder network (Section 8.3.3), a population of $10^{11}$ nanodevices each spaced an average ~100 microns apart throughout the body tissues can determine relative location to a positional accuracy of ~3 microns and an angular accuracy of ~2 milliradian, with the data updatable once every millisecond. Thus dispersed, the network can continuously monitor and record the relative positions of all limbs with worst-case (cumulative error) ~0.8 mm accuracy over a 2-meter span (Section 8.3.3). These devices can prepare high-resolution dynamic maps of body position, velocity, acceleration and rotation, allowing precise real-time digital kinesthesia during sports or artistic activities such as gymnastics, pole vaulting, ballet; during transportative activities such as driving cars around hairpin turns, roller-coaster rides, military aircraft maneuvers and space launches; during precision tool-using such as needle-threading, antique watch repair, or while using the fingers as measurement calipers; during self-defense activities requiring complex motions such as karate or judo; and during emergency situations such as automobile crashes and tumbling motions during falls from great heights. All such sensory data is readily outmessaged to the patient or user in real time (Section 7.4).

Nanorobots could assist in path integration or the reconstruction of experienced limb trajectories. Given the ability to determine rotation rates and to measure applied forces and torques, the network should also be able to forecast the anticipated future positions of limbs. By comparing these projections to actual results, the network can then infer the viscosity of the medium in which the activity is taking place (e.g., air or water), whether the environment is stable or is translating or rotating in some direction, and whether the human user is physically supported or in free-fall. Internal nanodevices can directly measure if a body is sitting, standing, laying prone, inverted, falling, floating or diving, and then communicate that information directly to the human user (Section 7.4). Detection of patient activity states—e.g., the patient is sitting, standing, or walking—can be used to control nanorobot behaviors (Chapter 12), activate or deactivate outmessaging displays (Section 7.4.6), and so forth. A
simple macroscale wearable tactile “compass belt” that would convey desired geographical directions has already been proposed.2994

4.9.2.2 Orientational Macrosensing
Since nanopendular sensors can reliably determine the direction of the local gravity field vector to within ~2 milliradians in ~10^{-4} sec (Section 4.3.4.2), the navigational network can poll its members and arrive at an accurate consensus on which direction is up. Nanodevices affixed to relatively stable hard body parts will exhibit more consistent orientational readings. Measurement of the gravity vector allows the vertical orientation of the rest of the body to be precisely fixed in space, especially useful for gymnasts, trapeze artists, underwater divers in murky lakes, and vestibular-impaired individuals, to whom this information may be outmessaged directly (Section 7.4).

The gravitational vector may also be indirectly measured, albeit more slowly and less accurately, by monitoring the body’s natural reactions to changes in the axis of gravitational loading. For instance, the variation in head-to-toe hydrostatic pressure for a 1.7 m tall, 70 kg human standing in a 1-g gravity field is ~0.17 atm; the ability to measure a systematic cross-body differential of 10^{-6} atm allows the detection of a ~10^{-5} g change along the lengthwise aspect, or ~10^{-4} g change along the transverse aspect. Thus a continuously updated whole-body barostatic map allows the human body to serve as a three-dimensional gravity/orientation sensor.

4.9.2.3 Body Weight Measurement
A network of nanodevices can inventory the volume and density of each of the body’s ~10^6 (1 mm)^3 voxels using a combination of acoustic ranging, somatic mapping, and flowmetry. Each voxel is identified as fat, muscle tissue, bone mass, interstitial fluid, and so forth. These measurements allow the body’s weight to be precisely computed as volume times density, rather than the usual method of measuring weight.

4.9.2.4 Gravitational Geographic Macrosensing
Medical nanodevices can measure variations in the gravity field to ~10^{-6} g’s for L = 20 micron gravimeters in a measurement time t_{meas} = 2-9 millisec (Section 4.4.2). This implies that in vivo nanodevices can take precise measurements of their latitude and altitude relative to sea level -100 times every second. Gravity increases toward the poles and at lower altitudes. Specifically, using the formula of Cassinis (accounting for rotational and polar flattening terms), g is given approximately by:

\[ g_{\text{meas}} = g_0 \left[ 1 + k_1 \sin^2(\theta_L) - k_2 \sin^2(2 \theta_L) \right] - k_3 h + k_4 h \rho_{\text{earth}} \]  

where \( \theta_L \) = terrestrial latitude (equator = 0°), h = height above sea level in meters, \( g_0 = 9.78039 \text{ m/sec}^2 \) (equatorial sea-level value of g), \( k_1 = 5.2884 \times 10^{-3} \), \( k_2 = 5.9 \times 10^{-6} \), \( k_3 = 3.086 \times 10^{-6} \text{ sec}^{-2} \), \( k_4 = 4.185 \times 10^{-5} \), and \( \rho_{\text{earth}} = 5522 \text{ kg/m}^3 \).

Since sea-level g varies from 9.78039 m/sec^2 at the equator to 9.83217 m/sec^2 at the north pole, a 20-micron gravimeter (Ag = 10^{-6} g) detects a change in position of 1 arcmin of latitude or ~1900 meters north/south along the Earth’s surface. Similarly, since at 45° latitude g varies from 9.806 m/sec^2 at sea level to 9.803 m/sec^2 at 1000 meters altitude, a 20-micron gravimeter detects a change in altitude of ~3.3 meters (e.g., upstairs vs. downstairs in a house). For comparison, in 1998 high-quality commercial gravity gradiometers measured gradients of ~10^{-3} g/meter and allowed the compilation of micro-g (~1 milligal) resolution aerial gravity maps; interferometers measured the gravitational acceleration of atoms to a precision of 10^{-10}.

To achieve such phenomenal positional accuracies, the nanodevice must be able to computationally resolve several complicating factors. First, localized mass concentrations representing nonuniformities in crustal density produce residuals of up to ±0.0006 m/sec^2, which may be removed from the data using a standard map of known terrestrial isostatic variations and anomalies. Indeed, matching observations to such a map could provide useful longitudinal information as well. Another complication is the variation in gravity due to tidal forces amounting to -3 x 10^{-7} g’s, twice daily, which lies at the limits of detectability for a 20-micron gravimeter. Other minor geodesic and terrain-related corrections, too complicated for discussion here, may also need to be applied in certain circumstances. Note that the presence of nearby heavy objects does not influence measurement accuracy: a 100-ton building 10 meters away adds a lateral acceleration of only 7 x 10^{-5} g’s to a human body.

One final complication is that patient movements create kinematic accelerations that must be distinguished from the gravitational accelerations. Gravity readings can be corrected by taking derivatives of the signals from kinesthetic monitoring to give gravity in the reference frame of the patient’s room, and many individual measurements may be averaged to improve accuracy since the gravity vector normally changes only very slowly.

4.9.3 Electric/Magnetic Macrosensing
4.9.3.1 Vascular-Interstitial Closed Electric Circuits
In vivo studies of the electric properties of blood vessels shows that the walls of veins and arteries present a specific electric resistance ~200 times greater than the conductive medium of blood (plasma)—roughly 200 ohm-m vs. 0.7 ohm-m. Thus blood vessels are properly regarded as relatively insulated conducting cables which can electrically connect an injured tissue with surrounding noninjured tissue. Capillaries form an electric junction with the interstitial fluid, so the electric gradient can be canceled by ionic transport. Cell membranes are insulating dielectrics with resistance and capacitance, penetrated by ionic channels or gates for ionic transport. Additionally, in 1941 Szent-Gyorgyi suggested the possibility of semi-conduction in proteins, a theory which has since been elaborated by many investigators.

B. Nordenstrom has proposed that the above schema represents a system for selective electrogenous mass transport of material between blood and tissue which he calls Vascular-Interstitial Closed (electric) Circuits (VICCs)—in effect, an additional electrocirculatory system operating in parallel with the well-known diffusive, osmotic, and hydrodynamic mechanisms of regular bloodstream materials transport. As long as ions can leak through open pores and migrate through ion channels, long distance transport cannot take place and the VICC system remains primarily a local (e.g., within tissues), not global, electrical circuit. Connections also exist with conductive media including cerebrospinal fluid, bile, and urine.

Nanorobots capable of monitoring the status of local VICCs may rapidly and efficiently acquire a wealth of systemic information without the need for direct inspection of the affected tissues. For example, when a working muscle produces catabolic products such as lactic acid, an electrochemical potential gradient develops...
between the muscle and surrounding tissue in a known manner. Injured tissues become polarized in relation to surrounding tissue, as initial catabolic degradation products acidify the tissue. Malignant neoplasms, benign neoplasms, and internally necrotic granulomas all may polarize electrically in relation to surrounding tissue; a vascular thrombus also contains ionized material. Large deflections in the diffusion potential of blood occur if blood is deoxygenated or is infected with Gram-negative bacteria. Blood changes its electric potential from +500 mV to +1000 mV during spontaneous coagulation, and the spontaneous electric potential at the site of a crush injury of the iliac crest in rats oscillated several times between +190 mV and −60 mV over a four day experimental trial.

Nordenstrom proposes that in vivo electrophoresis via the VICC system may also play a role in mediating leukotaxis. When tissue is artificially polarized by electrodes, accumulation of granulocytes with margination and development of pseudopods is the result. In one experiment, exposure of a mesenterial membrane to a 1 microampere 1 volt field for 30 minutes stimulated diapedetic bleeding near the anode. At higher power, the arterio-capillaries contracted, narrowed, and emptied of blood cells, while the veno-capillaries and venules widened and filled with granulocytes. Monitoring such actions of local VCCs experiencing normal voltage fluctuations could alert medical nanorobots to the presence of local injuries, tissue changes, or pathologies that otherwise might go unnoticed for a considerable time.

4.9.3.2 Electric/Magnetic Geographic Macrosensing

Detection of electric fields to −100 volts/m (Section 4.7.1), the normal vertical atmospheric gradient, in theory might allow determination of aboveground altitude to −1 m accuracy if the time between recalibrations is very brief. However, a human body (which is a good conductor), standing on level ground, acquires a slight negative surface charge and becomes “part of the ground” electrically. This distorts upward the equipotentials that usually run parallel to the ground, thoroughly corrupting an altitude measurement. Nearby lightning hits and storms are detectable, which is useful.

However, medical nanodevices with appropriate magnetometers can measure variations in magnetic field to −0.1 microtesla using a (600 nm)² permanent magnet sensor, in a measurement time Σmean = 0.6 millisecond (Section 4.7.2). Geomagnetic field maps of the Earth’s surface are highly anisotropic in both latitudinal and longitudinal directions; knowledge of the field plus the absolute direction of true north (e.g., using a nanogyroscope; Section 4.3.4.1) gives longitude information from the separation of the planetary magnetic and spin poles. This implies that properly equipped in vivo nanodevices can establish their latitude and longitude on Earth’s surface −1000 times every second by this means. In particular, the horizontal component of the geomagnetic field vector ranges from 0–41 microtesla from magnetic pole to magnetic equator; the independent vertical field component ranges from 0–70 microtesla. Thus a 0.1-microtesla sensor should resolve each component to an accuracy of ~1 arcmin or ~20 km in both latitude or longitude. Terrestrial surface magnetic anomalies (e.g., iron deposits) produce local variations of 0.03–30 microtesla and are thus detectable. Additional complications to these measurements include artificial magnetic field sources, the 0.01–0.1 microtesla solar daily variation, occasional magnetic storms producing erratic geomagnetic fluctuations of 0.01–5 microtesla, sudden-commencement ionospheric electrojets up to 0.3 microtesla, and various long-term secular variations in the geomagnetic field. Radio, television, and direct broadcast satellite signals probably are not detectable by individual nanodevices (Section 7.2.3), but such detection may be indirectly achieved in vivo using nanorobot-accessible dedicated macroscale antennas implanted in the body (Section 7.3.4).

4.9.3.3 Piezoelectric Stress Macrosensing

The piezoelectric effect (Sections 6.3.2 and 6.3.5) is “the production of electrical polarization in a material by the application of mechanical stress”. Many materials in the human body are piezoelectric, including tendon and elastin, dentin and bone. This polarization, or surface charge, varies as the physical stress imposed on the material changes over time. Thus, measurement of the piezoelectric effect in bone or tendon can provide information on the physical loads that are being carried by these materials. For example, a shear stress applied along the long axis of a bone alters the polarization voltage that appears on a surface at right angles to the axis. In another experiment, piezoelectric surface charges on a loaded human femur were measured to range from −131 to +207 picocoulombs/cm² (−8 to +13 charges/micron²) from one end of the bone to the other end, depending upon position along the shaft. Real-time monitoring of this data would permit specific inferences as to the amount of load the bone was carrying and from what direction, including bending, shearing, and twisting forces, from which whole-body activity states could subsequently be inferred. Knowledge of these surface charge variations might also be exploited in osteographic (Section 8.2.4) or functional (Section 8.4) navigation.

4.9.4 Optical Macrosensing

In vivo medical nanodevices can gather little useful optical information from the external environment. Here’s why.

In biological soft tissues, scattering dominates absorption except in the pigmented layers of the epidermis and stratum corneum. Thus the propagation of light in tissues may be regarded as occurring in two steps.

In the first step, optical photons of intensity I₀ falling perpendicularly on skin are transmitted according to Beer’s law through tissue to a depth z with a transmitted intensity of

\[ I_z = I_0 (1 - r_p) e^{-\sigma_z z} \]  

[Eqn. 4.57]

where \( r_p \) is a (specular) reflection coefficient for visible light (Fresnel reflection at the air-tissue surface) − 4%–7%, or 0% if the original light source lies within the body; and the transmission coefficient \( \sigma_z = \sigma_a + \sigma_s \) (scattering coefficient − 30,000 m⁻³) for various human soft tissues at optical wavelengths. (Coefficient values for the most heavily pigmented skin layers may be 5–7 times higher.) Typical values for \( \sigma_s \) over a 10,000-100,000 m⁻¹, average −30,000 m⁻¹ for soft tissue, although exceptionally clear tissues with \( \sigma_a = 1000 m^{-1} \) have been reported. Thus the mean free path of an optical photon in human soft tissue is 10–100 microns, average −30 microns (~1.5 tissue cell-widths), up to an extreme maximum of −1 mm for the most transparent tissues known. For \( \sigma_s \) over a 30,000 m⁻¹, at z = 150 microns, \( I_z / I_0 < 0.01 \) and -99% of all photons have been scattered at least once from their initial path.

In the second step, in an unbounded medium the patch of fully scattered photons continues to propagate through the tissue via diffusion until all photons are absorbed. The complicated governing diffusion equation has not yet been been completely solved analytically, but the asymptotic diffusive fluence is given roughly by

\[ I_d = I_0 e^{-\sigma_z z} \]  

[Eqn. 4.58]
where $\sigma_d$ is the diffusion exponent or effective attenuation coefficient (averaging ~900 m$^{-1}$ for typical soft tissues but with an extremely wide range reported over the ultraviolet, visible, and near-infrared wavelengths, from 10-1,000,000 m$^{-1}$.) As a crude approximation, the initial intensity of the fully scattered photon patch $I_i = I_0 (1 - r_{sp}) n_{scat}$, where $n_{scat} = \sigma_i / (\sigma_i + \sigma_d) \approx 0.987$ is the albedo for single particle scattering.

Over the optical band from 400-700 nm, incident intensity $I_0$ is 100-400 watts/m$^2$ when standing in direct sunlight; artificial lighting in homes and offices is typically 0.1-10 watts/m$^2$; moonlight provides only $I_0 \approx 10^{-6}$ watts/m$^2$; and the absolute threshold for human vision is $10^{-8}$ watts/m$^2$.

The most complex and difficult challenge in neural macrosensing (Section 7.4.6.5) may simplify untangling of the signal compression. Rapid visual field identification using artificial neural nets is also a subject of much current research interest. Eyeball rotations (e.g., via resident intradevice nanogyroscopes; Section 4.3.4.1), eyelid position, pupil aperture, and lens accommodation under ciliary muscle control must be monitored to supplement the vision field width, and significant natural data compression techniques which all must be untangled in real time. Developing algorithms capable of interpreting raw optical nerve traffic, 3018-3021 say, to recognize a specific human face or a specific scene in the vision field, would prove a significant research challenge. (Direct monitoring of photoreceptors or retinal membrane potentials and other techniques (Section 7.4.6.5) may simplify untangling of the signal compression.)

Rapid visual field identification using artificial neural nets is also a subject of much current research interest. Eyeball rotations (e.g., via resident intradevice nanogyroscopes; Section 4.3.4.1), eyelid position, pupil aperture, and lens accommodation under ciliary muscle control must be monitored to supplement the vision field analysis.

4.9.6 Other Macrosensing

We have only scratched the surface of the potential for nanorobot macrosensing. Medical nanodevices can quantitatively monitor variables not normally accessible to human consciousness, such as:

1. hormone and neurotransmitter levels;
2. gastroelectric oscillations and skin conductivity;
3. pupil dilations;
4. drug and alcohol (and breakdown product) concentrations in the blood;
5. internal organ damage or malfunction;
6. digital real-time performance data on vital organs or limbs (e.g., continuous kidney volume throughput, pancreatic insulin output, cholesterol metabolism in the liver, or lactic acid production in specific muscles during exercise); and
7. continuous mapping of thermoregulatory isotherms from which the temperature and heat capacity of the external medium (e.g., swimming pool water, cool night air) and the presence of sunlight or shade on the skin (via dermal thermal differentials) may be inferred (Section 8.4.1.3).

Various chemical substances cyclically increase and decrease in serum concentration with time of day, permitting crude temporal macrosensing (Section 10.1.1). Other substances (e.g., aldosterone; Appendix B) significantly vary in serum concentration depending upon whether the patient is supine or standing, thus permitting limited biochemical postural macrosensing.

If nanorobots can leave and re-enter the body (Section 8.6; Chapter 16), then macrosensing may include direct sampling of the external environment upon demand and the possibilities for remote data acquisition become virtually limitless.
5.1 Flexible Form and Function

It has been asserted that nanomechanical systems fundamentally differ from systems of biological molecular machinery in their basic architecture—specifically, that nanomechanical components are supported and constrained by stiff housings, while biological components often can move freely with respect to one another. As regards medical nanodevices, this may be a somewhat artificial distinction. The likelihood that most nanoscale components will be connected in rigid arrays does not imply that the nanomachines themselves must be entirely rigid in shape, nor does it rule out the possibility that some major nanomachine components may be designed to allow periodic reconfiguration and repositioning.

Why might a flexible shape be useful in nanomedicine? While cell membranes are self-sealing, large wall breaches during cell repair operations can be problematic. A flexible shape makes it easier for a cell repair nanodevice to enter the cytoplasm with minimum disruption, for instance by elongating and narrowing so as to present the narrowest possible aspect during plasma membrane and cytoskeletal penetration (Section 9.4.5). When navigating through narrow passages in hard biological substances such as bone or enamel, a nanorobot with a rigid shape is more likely to scrape or jam than is a more flexibly-shaped device. Extensible volumes allow the projection of compliant mechanical pseudopods of various sizes and shapes from the nanomachine surface. Deformable bumpers make it easier to establish and maintain reliable multidevice linkages in large-scale cooperative nanorobotic architectures.

Flexibility expands the options available for nanodevice mobility to include amoeboid and pulsatile peristaltic locomotion (Section 9.4.3.), surface deformation natation (Section 9.4.2.5.1), and circumvascular tissue diving or nanorobot diapedesis (Section 9.4.4). Fluidlike surfaces are ubiquitous among motile microorganisms. Malleable nanodevices situated on a cardiac or arterial luminal surface can adopt minimum fluid drag configurations.

Microhydrodynamically stable is another factor. For example, when placed in a bloodflow of constant speed in a tubular vessel, a rigid sphere, rod, or disk will tumble as it travels, due to the differential axial velocity field. But external fluid stresses distort a deformable object like an emulsion droplet from its original spherical shape into an ellipsoid oriented at a constant angle to the direction of flow. The fluid stresses are transmitted across the droplet interface; the surface and interior fluid circulates about the particle center in a tank tread motion. Shape changes can also be employed for steering and orientational control during active nanorobot swimming (Section 9.4.2.5). Hard-shelled particles without active mobility don’t marginate laterally in blood vessels; deformable surfaces can radially migrate. Flexible surfaces may also be used to minimize the increment to blood viscosity caused by the presence of bloodborne nanorobots (Section 9.4.1.4).

A flexible or “metamorphic” surface is a nanodevice exterior surface comprised of independently controllable elements that can translate or rotate their relative positions, thus enlarging or contracting total surface area of the device, or changing its shape, with or without altering the membership of elements in the surface, with or without altering the enclosed volume of the entire nanomachine, while maintaining continuous structural integrity and nonpermeability of the surface. The range of possible designs is enormous. Metamorphic surfaces may include integument systems with semiflexible components; hinged elements with fixed relative position and area, allowing variable volume; partially mobile surfaces having unit elements of fixed size but variable position relative to their neighbors, allowing control of surface area as well as volume; and even fully metamorphic surfaces with surface elements free to rotate, alter shape or orientation, slide, or even change membership, permitting maximum surface/volume flexibility.

Device shape is driven by task requirements which may or may not demand surface flexibility, as described in Section 5.2. Section 5.3 presents a number of design alternatives for metamorphic surfaces and manipulators, and briefly considers the impact of surface flexibility on internal configurational design. Section 5.4 examines the challenges of metamorphic bumpers which serve as interdevice fasteners and junctions. Discussion of the biocompatibility of nanodevice surfaces is deferred to Chapter 15.

5.2 Optimum Nanorobot Shape

The optimum nanorobot shape varies according to the function the device is designed to perform and the environment in which the device must operate. A consideration of the many different functions that medical nanodevices are asked to perform suggests several general classes of tasks requiring specific group geometries which drive the choice of optimum individual nanodevice geometry for each task class. For comparison, microscopic unicellular bacteria generally take three basic forms: spherical or ellipsoidal (the cocci), cylindrical or rod-shaped (the bacilli), and curved rod, spiral, or comma-shaped (the spirilla), though there is at least one example of a square bacterium.

5.2.1 Free-Floating Solitary Nanodevices

Free-floating simple nanorobots intended solely as materials delivery or storage devices, or used as omnidirectional communications, navigational or control transponders operating independently in vivo, have no preferred orientation. Thus these nanorobots may be spherically symmetric with rigid surfaces. Spherical particles produce the smallest possible increment in blood viscosity as they tumble, in part because spheres offer the smallest possible interaction surface per unit volume of any geometrical shape. Such simple shapes are supported and constrained by stiff housings, while biological components often can move freely with respect to one another. As regards medical nanodevices, this may be a somewhat artificial distinction. The likelihood that most nanoscale components will be connected in rigid arrays does not imply that the nanomachines themselves must be entirely rigid in shape, nor does it rule out the possibility that some major nanomachine components may be designed to allow periodic reconfiguration and repositioning.

Why might a flexible shape be useful in nanomedicine? While cell membranes are self-sealing, large wall breaches during cell repair operations can be problematic. A flexible shape makes it easier for a cell repair nanodevice to enter the cytoplasm with minimum disruption, for instance by elongating and narrowing so as to present the narrowest possible aspect during plasma membrane and cytoskeletal penetration (Section 9.4.5). When navigating through narrow passages in hard biological substances such as bone or enamel, a nanorobot with a rigid shape is more likely to scrape or jam than is a more flexibly-shaped device. Extensible volumes allow the projection of compliant mechanical pseudopods of various sizes and shapes from the nanomachine surface. Deformable bumpers make it easier to establish and maintain reliable multidevice linkages in large-scale cooperative nanorobotic architectures.

Flexibility expands the options available for nanodevice mobility to include amoeboid and pulsatile peristaltic locomotion (Section 9.4.3), surface deformation natation (Section 9.4.2.5.1), and circumvascular tissue diving or nanorobot diapedesis (Section 9.4.4). Fluidlike surfaces are ubiquitous among motile microorganisms. Malleable nanodevices situated on a cardiac or arterial luminal surface can adopt minimum fluid drag configurations.

Microhydrodynamically stable is another factor. For example, when placed in a bloodflow of constant speed in a tubular vessel, a rigid sphere, rod, or disk will tumble as it travels, due to the differential axial velocity field. But external fluid stresses distort a deformable object like an emulsion droplet from its original spherical shape into an ellipsoid oriented at a constant angle to the direction of flow. The fluid stresses are transmitted across the droplet interface; the surface and interior fluid circulates about the particle center in a tank tread motion. Shape changes can also be employed for steering and orientational control during active nanorobot swimming (Section 9.4.2.5). Hard-shelled particles without active mobility don’t marginate laterally in blood vessels; deformable surfaces can radially migrate. Flexible surfaces may also be used to minimize the increment to blood viscosity caused by the presence of bloodborne nanorobots (Section 9.4.1.4).

A flexible or “metamorphic” surface is a nanodevice exterior surface comprised of independently controllable elements that can translate or rotate their relative positions, thus enlarging or contracting total surface area of the device, or changing its shape, with or without altering the membership of elements in the surface, with or without altering the enclosed volume of the entire nanomachine, while maintaining continuous structural integrity and nonpermeability of the surface. The range of possible designs is enormous. Metamorphic surfaces may include integument systems with semiflexible components; hinged elements with fixed relative position and area, allowing variable volume; partially mobile surfaces having unit elements of fixed size but variable position relative to their neighbors, allowing control of surface area as well as volume; and even fully metamorphic surfaces with surface elements free to rotate, alter shape or orientation, slide, or even change membership, permitting maximum surface/volume flexibility.

Device shape is driven by task requirements which may or may not demand surface flexibility, as described in Section 5.2. Section 5.3 presents a number of design alternatives for metamorphic surfaces and manipulators, and briefly considers the impact of surface flexibility on internal configurational design. Section 5.4 examines the challenges of metamorphic bumpers which serve as interdevice fasteners and junctions. Discussion of the biocompatibility of nanodevice surfaces is deferred to Chapter 15.

5.2 Optimum Nanorobot Shape

The optimum nanorobot shape varies according to the function the device is designed to perform and the environment in which the device must operate. A consideration of the many different functions that medical nanodevices are asked to perform suggests several general classes of tasks requiring specific group geometries which drive the choice of optimum individual nanodevice geometry for each task class. For comparison, microscopic unicellular bacteria generally take three basic forms: spherical or ellipsoidal (the cocci), cylindrical or rod-shaped (the bacilli), and curved rod, spiral, or comma-shaped (the spirilla), though there is at least one example of a square bacterium.

5.2.1 Free-Floating Solitary Nanodevices

Free-floating simple nanorobots intended solely as materials delivery or storage devices, or used as omnidirectional communications, navigational or control transponders operating independently in vivo, have no preferred orientation. Thus these nanorobots may be spherically symmetric with rigid surfaces. Spherical particles produce the smallest possible increment in blood viscosity as they tumble, in part because spheres offer the smallest possible interaction surface per unit volume of any geometrical shape. Such simple shapes are supported and constrained by stiff housings, while biological components often can move freely with respect to one another. As regards medical nanodevices, this may be a somewhat artificial distinction. The likelihood that most nanoscale components will be connected in rigid arrays does not imply that the nanomachines themselves must be entirely rigid in shape, nor does it rule out the possibility that some major nanomachine components may be designed to allow periodic reconfiguration and repositioning.

Why might a flexible shape be useful in nanomedicine? While cell membranes are self-sealing, large wall breaches during cell repair operations can be problematic. A flexible shape makes it easier for a cell repair nanodevice to enter the cytoplasm with minimum disruption, for instance by elongating and narrowing so as to present the narrowest possible aspect during plasma membrane and cytoskeletal penetration (Section 9.4.5). When navigating through narrow passages in hard biological substances such as bone or enamel, a nanorobot with a rigid shape is more likely to scrape or jam than is a more flexibly-shaped device. Extensible volumes allow the projection of compliant mechanical pseudopods of various sizes and shapes from the nanomachine surface. Deformable bumpers make it easier to establish and maintain reliable multidevice linkages in large-scale cooperative nanorobotic architectures.

Flexibility expands the options available for nanodevice mobility to include amoeboid and pulsatile peristaltic locomotion (Section 9.4.3), surface deformation natation (Section 9.4.2.5.1), and circumvascular tissue diving or nanorobot diapedesis (Section 9.4.4). Fluidlike surfaces are ubiquitous among motile microorganisms. Malleable nanodevices situated on a cardiac or arterial luminal surface can adopt minimum fluid drag configurations.

Microhydrodynamically stable is another factor. For example, when placed in a bloodflow of constant speed in a tubular vessel, a rigid sphere, rod, or disk will tumble as it travels, due to the differential axial velocity field. But external fluid stresses distort a deformable object like an emulsion droplet from its original spherical shape into an ellipsoid oriented at a constant angle to the direction of flow. The fluid stresses are transmitted across the droplet interface; the surface and interior fluid circulates about the particle center in a tank tread motion. Shape changes can also be employed for steering and orientational control during active nanorobot swimming (Section 9.4.2.5). Hard-shelled particles without active mobility don’t marginate laterally in blood vessels; deformable surfaces can radially migrate. Flexible surfaces may also be used to minimize the increment to blood viscosity caused by the presence of bloodborne nanorobots (Section 9.4.1.4).
nanomachines are the kind most likely to be deployed in the greatest numbers, and their reduced surface area minimizes potential biocompatibility problems. These free-floating nanodevices must have ready access to all tissues via blood vessels. Since they are nonmotile machines, to avoid getting stuck they must not physically extend wider across their longest axis than the width of human capillaries, which average 8 microns in diameter but may be as narrow as 4 microns (Section 8.2.1.2).

Consider a nanodevice of fixed surface area \( A_n \), total volume \( V_n \), and longest transdevice diameter \( L_n \). Then:

A. For a spherical device of radius \( r \), then \( L_n = 2 r \), \( A_n = 4 \pi r^2 \), and \( V_n = \frac{4}{3} \pi r^3 \).

B. For a prolate spheroidal (football-shaped) device of length \( 2a \), width \( 2b \), and eccentricity \( e = (a^2 - b^2)^{1/2} / a \), then \( A_n = 2 \pi b^2 + 2 \pi a b (\arcsin(e) / e) \) and \( V_n = \frac{4}{3} \pi a b^2 \). For an oblate spheroidal device, \( A_n = 2 \pi a^2 + \pi b^2 \ln\left(\frac{1+e}{1-e}\right) / e \) and \( V_n = \frac{4}{3} \pi a^2 b \). In both cases, \( L_n = 2a \) and the maximum enclosed volume per unit area occurs at \( a = b \) (e.g., a sphere).

C. For a right circular disk or cylindrical device of radius \( r \) and height \( h \), then \( L_n = (h^2 + 4r^2)^{1/2} \), \( A_n = 2 \pi r^2 h \), and \( V_n = \pi r^2 h / 3 \). Maximum enclosed volume per unit area occurs at \( h = \frac{3}{2} r \).

D. For a right circular conical device of radius \( r \) and height \( h \), then \( L_n = 2r \) for \( h \leq \frac{3}{2} r \) (but \( L_n = (h^2 + r^2)^{1/2} \) if \( h > \frac{3}{2} r \) ), \( A_n = \pi r^2 + \pi r (h^2 + r^2)^{1/2} \), and \( V_n = \pi r^2 h / 3 \). Maximum enclosed volume per unit area occurs at \( h = \frac{3}{2} r \).

E. For a right triangular prismatic device (Fig. 5.4) with three equal sides of length \( s \) and height \( h \), then \( L_n = (h^2 + s^2)^{1/2} \), \( A_n = 3 h s + 3^{1/2} s^2 / 2 \), and \( V_n = 3^{1/2} s^3 h / 4 \). Maximum enclosed volume per unit area occurs at \( h = s / 2^{1/2} \).

F. For a cubical device with three equal sides of length \( s \), then \( L_n = (3^{1/2}) s \), \( A_n = 6 s^2 \), and \( V_n = s^3 \).

G. For a right square prismatic device with two equal sides of length \( s \) and height \( h \), then \( L_n = (h^2 + 2 s^2)^{1/2} \), \( A_n = 2 s^2 + 4 s h \), and \( V_n = h s^2 \). Maximum enclosed volume per unit area occurs at \( h = s \).

H. For a right hexagonal prismatic device (Fig. 5.4) with six equal sides of length \( s \) and height \( h \), then \( L_n = (h^2 + 4 s^2)^{1/2} \), \( A_n = 6 s (h + 3^{1/2} s / 2) \), and \( V_n = 2^{7/2} s^3 h / 2 \). Maximum enclosed volume per unit area occurs at \( h = s / 2^{1/2} \).

I. For a truncated octahedral device (Fig. 5.5) of edge \( s \), then \( L_n = (10^{1/2}) s \), \( A_n = (6 + 432^{1/2}) s^2 \), and \( V_n = (128^{1/2}) s^3 \).

J. For a rhombic dodecahedron \( (1010) \) of edge \( s \), then \( L_n = (48^{1/2} / 3) s \), \( A_n = (11.3137) s^2 \), and \( V_n = (3.0792) s^3 \).

K. For a nonregular octahedron (Fig. 5.8) of equatorial edge \( s \) and vertex edge \( (3^{1/2} / 2) s \), then \( L_n = (21^{1/2}) s \), \( A_n = (8^{1/2}) s^2 \), and \( V_n = (1/3) s^3 \).

L. For a regular octahedron \( (1010) \) of edge \( s \), then \( L_n = (2^{1/2}) s \), \( A_n = (12^{1/2}) s^2 \), and \( V_n = (2^{1/2} / 3) s^3 \).

Table 5.1, computed using the above relations, confirms that spheres offer the greatest storage volume per unit surface area and thus are the most efficient shape for this application.

### 5.2.2 Actively Swimming Nanodevices

Nanodevices of another class may swim rapidly through the bloodstream. It should be possible to navigate even the fastest-moving arterial flows if desired. This locomotive capability is risky to the patient if employed by large numbers of devices simultaneously (Section 9.4.1), and may not be necessary for many of the applications proposed in this book.

Bloodstream swimming by nanorobots is qualitatively dissimilar to surface swimming or to a submarine moving through the ocean at depth. Unlike these macroscale analogs where inertial forces prevail, in the microscale environment viscous forces dominate. Inertial and gravitational forces are almost irrelevant. Bloodstream swimming by simple reciprocal motions is not possible (Section 9.4.2.5). Rather, it is necessary to use deformation, helical, or other drive systems, all of which involve either a tubular shape to allow propelled fluids to pass through the center of the device, or an axially symmetric form such as a conical-, egg- or teardrop-shaped spiral to minimize viscous drag while “drilling” through the fluid during locomotion. Red blood cells assume a biconcave shape only when in static equilibrium. In a flowing condition, lone erythrocytes deform into the shape of a bullet or slipper in the capillary blood vessels; \( ^{362} \) at lower shear rates in the arteries, multiple cells aggregate into cylindrical rouleaux oriented roughly in the direction of flow (Section 9.4.1.2). Metamorphic surfaces may allow a nanodevice to configure all fluid contact planes to achieve minimum drag for every velocity vector employed and every fluid traversed.

Are conventional streamlined shapes necessary or useful for micron-scale swimmers? In the macroscopic world, a body traveling through water experiences the least resistance to forward motion (drag) if it is rounded in the front and tapers to a rear point in the familiar shape of a tuna or whale—the animal thus shaped meets little drag, ~10 times less than would a sphere or a person of the same size. \( ^{362} \) However, streamlining and special hydrofoil shapes serve mainly

- A. to reduce induced drag, which largely disappears in nonturbulent microscale flows; and
- B. to reduce pressure drag, an inertial force which also becomes relatively unimportant at the microscale.

For example, the ratio of viscous drag to pressure drag may be computed from Eqs. 9.89 and 9.90 as \( F_{\text{viscous}} / F_{\text{inertial}} = (12 \rho \eta / \pi C_D) R_{\text{nano}} / \rho_{\text{fluid}} (1 / R_{\text{nano}} V_{\text{nano}}) \sim 10^{-5} \) (\( R_{\text{nano}} V_{\text{nano}} \)) in 310 K water. Given that the highest likely nanorobot swimming speed in vivo is \( V_{\text{nano}} \sim 1 \text{ cm/sec} \) (Section 9.4.2.6), then \( F_{\text{viscous}} / F_{\text{inertial}} \approx 100 \) for \( R_{\text{nano}} \approx 10 \) micron. Thus at the microscale, viscous forces predominate. Viscous forces are determined by total surface area in contact with fluid, so the lowest drag on a moving mass is produced (all else being equal) by a shape that presents the minimum possible surface area to the fluid, that is, approximating a sphere. The drag even on extremely pointed shapes differs little whether the object is moving forward or sideways. For instance, experiments show that the most extreme needle-shaped bodies fall about half as fast sideways as they do end-on. \( ^{1378} \) Note that natural micron-scale swimmers such as bacteria and small metazoans are typically ovoid or cylindrical, rather than fishlike, in shape.

An examination of 218 genera of free-floating and free-swimming bacteria \( ^{3582} \) revealed that motile genera are less likely to be spherical and have larger axial ratios (typically 3:1) than nonmotile genera. Spherical shapes were found to produce the largest random dispersal by Brownian motion, and oblate spheroids were rare, possibly due to the increased surface area. Prolate spheroids provided a reduced sinking speed; elongation slightly favored swimming speed, but strongly improved the temporal detection of chemical stimulus gradients by any of three different mechanisms—the probable explanation for the popularity of rod-like shapes in motile bacteria (though others \( ^{3615} \) have been proposed).
5.2.3 Intracellular Nanodevices

Nanodevices intended to perform tasks inside human tissues by leaving the bloodstream and entering the cell or nucleus will require a physical configuration consistent with the particular mission and the specialized tools required (Chapter 21). Metamorphic surfaces may assist in cyto devices in achieving nondisruptive membrane penetrations (Section 9.4.5), deployment and manipulation of flexible mechanical pseudopods or motive appendages (Sections 5.3.1 and 9.3.1.6), and reconfiguration of sensors or other subsystems (Section 5.3.5). Cell repair nanorobots could adopt one configuration to seek the target cell, a second to penetrate the membrane, and yet a third while operating within the cytosol.

Basic device shape will not be driven by hydrodynamic considerations because cell repair nanorobots spend the bulk of their operational time on site, not in transit, and need move only relatively slowly when they travel between worksites (Chapter 21). Nor will device shape be governed by tessellation rules (Section 5.2.4), since these machines will normally spend most of their time in solitary activity or at least out of direct physical contact with other nanorobots working nearby (Chapter 21). Thus the strongest purely geometric influence on these most complex of nanodevices may be volume storage efficiency. A spherical shape or near-spherical icosahedral shape (as in many viruses) can contain the largest possible mass of nanocomputers, mass memory, power supplies, repair consumables, specialized tools, communications and navigational equipment, and so forth.

However, cellular repair machines may also need to dock with other nanomachines at irregular intervals to receive supplemental materials, fuel, or information, so it would be convenient to use a shape with a large number of planar faces across which such transfers might readily be effected. A space-filling shape (Section 5.2.5) also allows ready aggregation of operational units in vivo for ad hoc conferencing plus efficient storage of unused units. A truncated octahedron with a total of 14 hexagonal and square faces (Section 5.2.5) is the space-filling polygonal shape with the highest volume/surface ratio, closest to the sphere (Table 5.1).

5.2.4 Tessellating Nanodevice Aggregates: Nanotissues

Some of the most important applications of medical nanodevices involve cooperative activity by millions or billions of nanorobots in close physical proximity with each other. In such cases it may be essential that individual machines fit together snugly in a mosaic pattern, or periodic tessellation, across a surface to achieve an air-tight or watertight seal. The most important additional complication is that biological surfaces are usually in motion. To accommodate such motion, nanorobots may use inflatable or expansible perimeter “bumpers” to achieve a continuously tight fit. Metamorphic surfaces are used to construct these bumpers (Section 5.4).

This Section discusses nanodevice aggregates whose tasks require them to completely occupy stationary or moving surfaces. Section 5.2.5 discusses optimum shapes for aggregated nanodevices whose tasks require them to uniformly fill volumes with nanomachinery.

5.2.4.1 Tiling Nondeforming Surfaces

If the surface to be covered by a nanorobot aggregate has constant area, such as the exterior of a bone, tooth, or fingernail, or the inner surface of a passive sinus or duct, or the skull and meninges, or certain components of the eyeball (cornea, sclera, etc.), then the question is how to completely tile a fixed surface using prismatic units of a particular shape, a familiar problem in spatial geometry.519,520 Orthogonally-arrayed unit circles achieve a packing density of only 78.54%, while hexagonal packing of unit circles achieves 90.70% coverage. However, as is well-known, triangles, squares, or hexagons are the only regular polygonal prismatic solids that can completely tile a plane by themselves, achieving 100% packing density (Fig. 5.1). There are also an infinite number of irregular planar polygonal tessellations using a wide variety of tile shapes (Fig. 5.2), although nanorobots with these shapes would suffer extraordinarily low volume/surface ratios—an extremely inefficient use of space.

Triangles, rectangles, or regular hexagons (e.g., a roll of chicken wire) can be used alone to tile a cylindrical surface. The deformation of nanorobot bumpers into curved wedge segments allows tiling of

---

Table 5.1 Geometrical Factors for Bloodstream-Traversing Nanorobots of Maximum Volume with Largest Transdevice Diameter Ln = 4 μm

<table>
<thead>
<tr>
<th>Nanodevice Geometrical Type</th>
<th>r</th>
<th>h</th>
<th>s</th>
<th>Enclosed Volume</th>
<th>Surface Area</th>
<th>Volume/Surface</th>
<th>Shown in Figure:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spherical/Spheroidal</td>
<td>2.00</td>
<td>---</td>
<td>---</td>
<td>33.51 μm³</td>
<td>50.27 μm²</td>
<td>0.6667 μm</td>
<td></td>
</tr>
<tr>
<td>Truncated octahedron</td>
<td>---</td>
<td>1.26</td>
<td>---</td>
<td>22.90 μm³</td>
<td>42.85 μm²</td>
<td>0.5343 μm</td>
<td>Fig. 5.5</td>
</tr>
<tr>
<td>Disk/Cylindrical</td>
<td>1.63</td>
<td>2.31</td>
<td>---</td>
<td>19.35 μm³</td>
<td>40.46 μm²</td>
<td>0.4783 μm</td>
<td>---</td>
</tr>
<tr>
<td>Rhombic dodecahedron</td>
<td>---</td>
<td>1.73</td>
<td>---</td>
<td>16.00 μm³</td>
<td>33.94 μm²</td>
<td>0.4714 μm</td>
<td>Fig. 5.6</td>
</tr>
<tr>
<td>Hexagonal prism</td>
<td>---</td>
<td>1.63</td>
<td>---</td>
<td>16.00 μm³</td>
<td>36.48 μm²</td>
<td>0.4386 μm</td>
<td>Fig. 5.4</td>
</tr>
<tr>
<td>Cubic/Square prism</td>
<td>---</td>
<td>2.31</td>
<td>---</td>
<td>12.32 μm³</td>
<td>32.00 μm²</td>
<td>0.3850 μm</td>
<td>Fig. 5.4</td>
</tr>
<tr>
<td>Circular conical</td>
<td>2.00</td>
<td>3.46</td>
<td>---</td>
<td>14.51 μm³</td>
<td>37.70 μm²</td>
<td>0.3849 μm</td>
<td>---</td>
</tr>
<tr>
<td>Regular octahedron</td>
<td>---</td>
<td>2.83</td>
<td>---</td>
<td>10.67 μm³</td>
<td>27.71 μm²</td>
<td>0.3849 μm</td>
<td>Fig. 5.9</td>
</tr>
<tr>
<td>Triangular prism</td>
<td>---</td>
<td>2.37</td>
<td>---</td>
<td>10.87 μm³</td>
<td>31.86 μm²</td>
<td>0.3349 μm</td>
<td>Fig. 5.4</td>
</tr>
<tr>
<td>Non-Regular octahedron</td>
<td>---</td>
<td>2.83</td>
<td>---</td>
<td>7.94 μm³</td>
<td>22.63 μm²</td>
<td>0.3333 μm</td>
<td>Fig. 5.8</td>
</tr>
</tbody>
</table>
ellipsoidal or other randomly curved surfaces provided the minimum radius of curvature \( r_{\text{curve}} \) is relatively large (e.g., \( r_{\text{curve}} \gg L_0 \)). For areas containing smaller radii of curvature (\( r_{\text{curve}} > L_0 \)), a combination of three regular polygonal prisms may be used to completely tile any arbitrarily curved surface: hexagonal prisms for flat sections, mixed with pentagonal prisms to add positive curvature (e.g., convexity, like a sphere) and heptagonal prisms to add negative curvature (permitting concave surface deformations). Artificial fullerene and natural radiolarian structures illustrate this approach (Fig. 2.19). Triangular prisms can also be used to tile arbitrarily curved surfaces; for example, some geodesic domes and the polio virus are symmetrical spherical arrangements of alternating triangles. (Many viruses are regular icosahedra.)

5.2.4.2 Tiling Deforming Surfaces

Most physiological surfaces are subject to periodic deformation, or stretching, along one or both dimensions. Consider the problem of tiling a surface that is periodically deforming in one direction only. A good example in human physiology is the large elastic arteries which circumferentially distend and contract as they absorb ~50% of the stroke volume of each ventricular ejection. Specifically, diameter oscillations in the pulmonary artery and aorta typically range from 9%-12%, averaging 11% for men and women under 35 years of age and 6.5% for people over 65. By contrast, vessel lengths display almost no variation at all.

In order to tile such a monoaxially deforming surface, individual nanorobots must have the ability to expand in one dimension only, using their inflatable bumpers. The nanodevice shape which requires the smallest bumpers (i.e., minimum volume) to accommodate a given linear expansion is the most space-usage efficient and is probably preferred. Assuming continuous perimeter bumpers, compare a series of \( N \) square, hexagonal, or triangular prismatic nanorobots with equal sides of length \( s \), height \( h \), and fully contracted bumpers of minimum depth \( b \) (Fig. 5.3). The center-to-center linear distance between adjacent square units is \( L_0 = s + 2b \), between adjacent hexagonal units along the stretch axis, \( L_0 = 3s/2 + 3b/3 \), and between adjacent triangular units along the stretch axis, \( L_0 = s + 3^{1/2}b \).

Now assume the center-to-center distance increases by \( \Delta L \) along the stretch axis. The increase in bumper volume for the squares is \( \Delta V_s = h \times L_0 \times (N-1)/N \times (L_0 + 3^{1/2}b) \) per unit. The useful interior nanorobot volume is \( U_s = h \times s^2 - h \times (L_0 - 2b)^2 \).

For any choice of \( b/L_0 \geq 0 \), then \( R_{sh} > 1 \) for all \( N > 2 \) and \( R_{sh} > 1 \) for all \( N > 0 \). That is, bumpers on square or triangular devices inevitably require a larger percentage change in nanorobot volume to effect a given linear adjustment, hence the hexagonal prism is always the more space-efficient shape for tasks requiring monoaxial deformations. (Squares are more efficient than triangles except for \( 0.43 \leq b/L_0 < 0.52 \).)

Similar results follow from an analysis of tiling a surface that is simultaneously deforming in both dimensions. Most moving biological surfaces stretch along both axes at once, often at different rates and phases. Examples include the skin (which can stretch up to 100% before permanent damage is observed), the abdominal diaphragm, cardiac chamber walls (papillary muscles thicken while shortening 7-20% during an isometric contraction), alimentary surfaces, and various bladders, glands and sacs. Tendons typically...
stretch 5% (up to 10% in the human psoas tendon) while contracting slightly in width. Apparently hexagonal prisms are also the most efficient shape for tessellating nanorobots aggregating on biaxially deforming surfaces; the intuitive rationale is that the most efficient shape is closest to a circle, with minimal perimeter length enclosing maximal area. In nature, epidermal squamous cells are arranged in hexagonal columns and onion skin cells also form neat arrays of hexagons. Hexagonal territorial partitioning is common among surface-feeding bird and fish species.

5.2.5 Space-Filling Nanodevice Aggregates: Nano-Organs

In specialized structures such as synthetic bones or artificial organs constructed as aggregates of millions or billions of nanorobots, it may be necessary for nanodevices to completely occupy specific three-dimensional volumes. This is the familiar problem of space-filling polyhedra in spatial geometry.

The volumetric packing factor of closely-packed spheres of equal radius is only \((3 \, \pi^2 / 64)^{1/2} \approx 0.68017\) by comparison, protein-interior packing densities are ~60%-85%. The simplest way to completely fill a volume (e.g., packing factor = 1) is to use a prism having a horizontal cross-section of a shape that will tile surfaces, as previously described (Section 5.2.4.1). Multiple planar fully-tiled surfaces can then be stacked vertically to fill the volume. Space-filling forms include triangular, square, and hexagonal prisms (Fig. 5.4). Of these, the hexagonal prism has the highest volume/area ratio and appears to be the most efficient prismatic form.

Only five other shapes can uniformly fill volumes by themselves and thus may be candidates for space-filling nanorobots. The most important of these is the truncated octahedron, similar to the tetrakaidecahedron which can fill space all by itself (Fig. 5.5). The truncated octahedron is formed by cutting the 6 corners off a regular octahedron (which has 8 faces, all triangles; Fig. 5.9), making a 14-hedron consisting of 6 small squares and 8 large hexagons, with 24 vertices and 36 edges. A truncated octahedron surrounded by 14 identical polyhedra contacting and matching each of its faces forms a solid unit in space resembling the original unit—that is, very close to spherical. Such an aggregate with many planar facial contacts permits easy docking, fastening, and transmission of forces in all directions. Weyl, following Lord Kelvin, long ago recognized that this 14-hedron has the highest volume/area ratio of all the space-filling polyhedra (Table 5.1) and is a close mathematical model of soap bubble froth.

Four other single-species space-filling polyhedra are known. The first is the garnet-shaped rhombic dodecahedron, which has 12 faces, all of which are rhombuses (Fig. 5.6). The volume/area ratio of this 12-hedron is intermediate between the 14-hedron and the hexagonal prism; the shape is found in back-to-back formations in the planar hexagonal honeycomb cell of the bee. The second is the closely related rhombo-hexagonal dodecahedron (Fig. 5.7), which is also self-packing. The third is the non-regular octahedron (8 faces, all triangles; Fig. 5.8), which results from truncating a cube such that all eight vertices are replaced by triangular faces with the new vertices coinciding with the intersection of the face diagonals of the original cube. Hence the equatorial edges are all the same but the vertex edges are half the length of a circumscribed cube’s interior diagonal, as distinct from the regular octahedron (Fig. 5.9) with all sides equal.
which is not space-filling. The fourth of the rare space-filling solids is the trapezohedra (Fig. 5.10). 518

There are also an infinite number of nonuniform space-filling systems in which two or more kinds of polyhedra are combined. 518 Nanodevices filling space by this means would require at least two distinct nanodevice species of the appropriate shapes. Perhaps the best-known variant is the octet element, a space-filling unit consisting of an ordered combination of regular tetrahedra and regular octahedra. Unique interlocking interfacial or other features can ensure that major subassemblies of the collective structure can only fit together in one correct way as an aid in automatic self-assembly and quality assurance.

Finally, many biological solids are so thoroughly perforated with holes and tubes that they are more properly regarded as porous rather than completely filled solids. These forms are simulated using open packings of polyhedra arrays 518 that may be employed to form near-arbitrary networks 284 of interior voids and channels in artificial organs, as required. Zeolites and other molecular sieves provide excellent models of this configuration. 382,2468,2469

5.3 Metamorphic Surfaces

Metamorphic surfaces are ubiquitous throughout the animal kingdom. The best known examples are the skins of dolphins 2034,2035 and sharks, 2036,2037 which may actively manipulate the contours of their rubbery integuments in real time to reduce or eliminate skin turbulence and thus maintain near-perfect streamlines to maximize swimming speed (though some have questioned this interpretation). 2022,2038

A precise control of the skin may seem foreign to us because, aside from eyebrow-raising, humans have lost volitional local dermal mobility in most areas of the body. Many lower animals can exercise precise control because they possess a well developed layer of cutaneous striped muscle, the paniculus carnosus, of which only vestigial remnants remain in the subdermal tissues of human beings. In man, the skin is more or less firmly attached to relatively inelastic and immobile fascia, which in turn is attached to major musculature, bone, or other underlying structures, particularly on the extremities and anterior chest wall.

An ability to control a nanorobot surface has many interesting applications and implications. This Section describes basic design
considerations, specific configurations, and other useful aspects of nanoscale metamorphic surfaces.

5.3.1 General Design Considerations

Basic design for any metamorphic surface includes a consideration of minimum feature size, linear extensibility, volumetric limits, areal expansion limits, and maximum distension speed. Other design considerations include tensile and compressive strength along all likely stress vectors at both minimum and maximum surface extensions, and the resistance of surfaces to cuts and tears.

5.3.1.1 Dimple Size

The smallest possible feature, or “dimple size,” on a metamorphic surface is defined by the minimum turning radius of the surface. In most designs, this radius is determined by the minimum lateral extent of the component blocks of which the surface is comprised. Considerations of radiation-induced structural failure in diamondoid materials and components suggest a minimum unit size \( L = 10\ nm \) (Chapter 13), roughly the thickness of the cellular lipid bilayer membrane. A single 180° turn requires a minimum of two linear segments; the minimum structural turning radius is \( r_{\min} = L / 2 \) \( \approx 7\ nm \) at a 90° joint angle, giving a minimum metamorphic feature diameter of \( \sim 14\ nm \). In most practical systems, power and control access requirements necessitate turning radii several times larger than this theoretical minimum, perhaps \( \sim 40\-50\ nm \). Such larger structures should still be able to physically manipulate even the theoretically smallest biological cells (which are \( \sim 40\-50\ nm \) in diameter).

5.3.1.2 Keyhole Passage

The minimum aperture through which a longitudinally stretched but laterally compressed nanodevice may pass is determined mainly by the minimum size of its internal incompressible components, not the minimum turning radius of its skin. The smallest working complex nanomachines of the human body, the ribosomes, are \( \sim 25\ nm \) in diameter; the minimum size of a \( \sim 100\ pW \) mechanochemical power supply is \( \sim 50\ nm \) (Section 6.5.3). Add to this the width of the integument component blocks and it is difficult to imagine even the most flexible artificial nanodevice squeezing through a keyhole less than \( 50\ nm \) in diameter—roughly the maximum dimension of the intercellular contact space. A design limit of \( 100\ nm \) would be more conservative.

In the macroscopic world, the octopus can stretch itself quite thin, passing rubberlike through small holes and narrow crevasses less than 10% of its size; arms, eyes, and even head can alter shape and elongate when necessary, with keyhole passage limited by its beak, the only hard part in its body.

5.3.1.3 Extensibility

The most useful property of a flexible metamorphic surface is its extensibility, defined here as the percentage length increase of a fully stretched material relative to its length at zero stretch. Maximum values for natural materials used to construct the human body include 7% for cartilage, 10% for collagen, 30% for muscle, 60%-100% for skin, and up to 170% for arterial wall material if muscles are artificially relaxed. Linear elasticities of \( \sim 1000\% \) are commonplace in artificial gels, rubbers, and other elastic materials, corresponding to \( \sim 1000\)-fold enclosed-volumetric changes. The El Tor strain of the Vibrio cholerae microbe shrinks 300-fold to the size of a large virus when plunged suddenly into cold salt water and remains viable in that state.

If the shape of a metamorphic surface can be controlled to \( r_{\min} \) (Section 5.3.1.1), then nanodevices may extrude these surfaces to define useful working subvolumes such as manipulatory appendages (e.g., prehensile fingers; see below), locomotive appendages (e.g., nanopseudopods; Section 9.3.1.6), large exterior hydrodynamic features (e.g., stabilizing fins), tools (e.g., a screw-shaped prow for easier cell penetration; Section 9.4.5.2), reconfigurable mechanical data arrays (e.g., Braille-like surface texturing), engulf formations (Section 5.3.4), perimeter contact bumpers (Section 5.4), or even entire second skins (see below). In theory, such extrusions can be quite large relative to device size.

For example, consider a nanodevice of volume \( V_n \) with an onboard pocket of volume \( V_f = f V_n \) containing \( N_b = f V_n / L_b^3 \) metamorphic unit surface blocks each \( L_b^3 \) in size. This is sufficient metamorphic material to extend a hollow cylindrical (hemisphere-capped) finger of diameter \( d_f \) out a total length \( l_f = f V_n \pi d_f L_b \).

For \( f = 0.01(1\%) \), \( V_n = 1\ micron^3 \), \( L_b = 10\ nm \), and \( d_f = 100\ nm \) (\( \pi \) \( d_f \) \( L_b \) + 31 blocks per circumference), \( N_b = 10^4 \) blocks in the pocket and \( l_f = 3\ microns \) maximum linear extension. Thus in theory, a finger which can extend three times the body length of the entire nanorobot can be stored in a (215 nm)\(^3\) pocket at the nanodevice surface. Ten such fingers, functionally equivalent to two microsized human hands, would occupy only \( (10/6) \pi (2.5)^3 \approx 0.08(8\%) \) of total nanodevice surface area and 10f - 0.10(10\%) of nanodevice volume.

Note also that a volume \( V_f \) of metamorphic blocks could be used to construct an enlarged second skin surrounding the entire nanodevice, enclosing an expanded volume \( V_e = (V_f/6L_b)^{1/2} \) if the shell is of thickness \( L_b \). For \( V_n = 1\ micron^3 \) and \( L_b = 10\ nm \), nanorobot volume may double \( (V_e/V_n = 2) \) if 9.5% of device volume is metamorphic blocks; nanorobot volume may expand up to tenfold if -28% of its volume is in controllable metamorphic blocks.

There are two limiting cases of extensibility. The first case is isoareal extension, in which total surface area remains constant while volume increases dramatically. An example is the erythrocyte, which under osmotic stress expands from its normal disk shape into a sphere due to the influx of water. Volume rises from 94 microns\(^3\) to a high of 164 microns\(^3\) when spherized, a 74% increase, but surface area rises from 135 microns\(^2\) to just 145 microns\(^2\), a mere 7% increase.

The second limiting case is isovolumic extension, wherein surface area expands at constant volume, a transition of perhaps more relevance to nanodevices containing an irreducible fixed volume of onboard nanomachinery. For instance, if a spherical device of radius \( r \) morphs into a disk of equal volume with height \( h_d \) and radius \( r_d = \alpha r \), surface area increases by a factor of:

\[
e_{\text{area}} = \frac{3\alpha^3 + 4}{6\alpha} - 1
\]

[Eqn. 5.3]

For \( \alpha = 4.7 \), \( h_d - r_s / 17 \) and areal extensibility \( e_{\text{area}} = 10.2 \) (1020%) while enclosed volume remains unchanged.

5.3.1.4 Reactivity

Another important parameter of nanorobot metamorphic surfaces is their reactivity—the speed with which they can execute a full-range structural modification. Energy requirements are key, given a power budget of \( P_m \) (watts) for morphing. In particular, metamorphic block velocity is limited both by the sliding power dissipation due to sliding block surfaces of area \( S \) that are sliding at a velocity \( v_{\text{slide}} \) or

\[
v_{\text{slide}} = \left( \frac{P_m}{k_b S} \right)^{1/2}
\]

[Eqn. 5.4]

where \( k_b = 400\ kg/m^2\cdotsec \) (Section 4.3.4.2), and by the drag power dissipated in driving a circular surface of radius \( r \) through a 310 K
aqueous (plasma, interstitial or cytosolic) environment of viscosity \( \eta = 1.1 \times 10^{-3} \text{ kg/m/s} \) at a velocity \( v_{\text{drag}} \), given by

\[
v_{\text{drag}} = \left( \frac{P_m}{6 \pi \eta r} \right)^{1/2} \quad \text{(Eqn. 5.5)}
\]

Assuming we wish to restrict total morphing energy dissipation to a budget of \( P_m = 0.1 \text{ pW} \) for a 1-micron\(^3\) nanorobot, then for \( S = 6 \text{ micron}^2 \) of sliding surfaces and \( r = 0.5 \text{ micron} \), \( v_{\text{drag}} \approx 0.3 \text{ cm/sec} \) (\( < v_{\text{slide}} = 0.6 \text{ cm/sec vacuo} \)). Using 0.3 cm/sec, the full-range morphing action to, say, double the nanorobot diameter (-0.5 micron of radial motion) requires ~0.2 millisecond for the surface of the entire nanorobot to complete the motion.

The 10-finger manipulator described in Section 5.3.1.3 may also extend at \( v_{\text{drag}} \approx 0.3 \text{ cm/sec} \) while remaining within the 0.1 pW energy budget, although a single individual finger may extend at \( v_{\text{drag}} \approx 1 \text{ cm/sec} \) allowing maximum deployment within the cytosol in ~0.3 millisecond. Thus metamorphic surfaces in vivo can exhibit ~kHz frequency, large-motion oscillations. By comparison, deformed erythrocytes when released from micro-pipette suction recover their normal biconcave shape in ~100 milliseconds. Therefore, the full range motion of the metamorphic gel fibers ~1 micron thick can shrink to 4% of their initial volume in ~100 milliseconds. 537

### 5.3.2 Metamorphic Surface Configurations

At least five classes of metamorphic surface configurations are readily discernible. Undoubtedly, many more may be conceived. The structural models briefly described below are not intended as specific design proposals but rather as geometrical representations of alternative design pathways that could provide the desired capability.

#### 5.3.2.1 Accordion Model

The Accordion Model is characterized by a surface folded in a repeating-W pattern, as in a Japanese fan or butterfly wing pleating; photographic and accordion bellows use what is known in origami\(^*\) as a “basic fold.” Point/line vertices may employ rigid hinges or flexural members. 1251 Fold geometry may be double-triangular, triangular-square, or double-square; may consist of segments of varying lengths; or may consist of a series of hinged blocks (Fig. 5.11). This surface remains flexible even near full distension, provided that obtuse angles may be continuously accessed. The main drawback of this model is its likely propensity to surface fouling in vivo due to the large number of concave pockets formed during flexure.

#### 5.3.2.2 Parasol Model

In the Parasol Model, the metamorphic surface consists of a series of overlapping plates pressed snugly together in the vertical dimension, but free to slide in the horizontal plane (Fig. 5.12). Each plate has at least one orthogonal stabilizing keel or “handle” through which the segment is connected to subsurface control (Section 5.3.3) or rigidification mechanisms. The minimum number of plate planes is two, which allows both one-dimensional and (limited) two-dimensional extensions. In the two-dimensional case, the lower plates are allowed to rotate codirectionally under a torsion-spring restoring force applied in the plane of the lower surface, keeping all edges tight under the upper plates to maintain leakproofness. A modest areal expansibility of \( e_{\text{area}} \approx 0.28 \) (28%) is depicted. The maximum number of plate planes is limited by nanodevice radius (hence maximum stack depth) and operating specifications such as surface rigidity, rugosity, reactivity and controllability. Self-scraping plate produce a fouling-resistant dysopsonic design—the device simply “shrugs” in all directions, neatly guillotining any biological adherents from their attachment points on the diamondoid surface. Vertical spring tensions are adjusted to keep plates pressed tightly together, ensuring water-tightness even while in motion. Corrugation features can be added to the underside of each plate to increase contact area and watertightness through plate tippling in curved configurations. Related crudely analogous structures include systems of fixed scales (e.g., lizard or snake skins) and roof tile shingle patterns on houses.

Consider an annular cylindrical section of diameter D comprised of a two-plane (\( p = 2 \)) parasol surface with square top plates of area \( L^2 \) and rectangular bottom plates of area \( L(L-h) \) where \( h \) is the width of the handle and also the minimum overlap of adjacent plates at full extension. Then \( A_{\text{max}} = N_{\text{plates}} L^2 = \pi D L \) and \( A_{\text{min}} = A_{\text{min}} + (N_{\text{plates}} - 1)(L^2 - 3hL), \) hence area extensibility \( e_{\text{area}} = (A_{\text{max}} - A_{\text{min}})/(A_{\text{max}} - A_{\text{min}}) \). For \( N_{\text{plates}} \approx 1 \) and \( h \ll L \), the theoretical limit for a 2-plane parasol is \( e_{\text{area}} = 1.00(100\%) \).

Extensibility is greatly improved by using additional plate planes. In the compact configuration of Figure 5.13, for \( p \gg 1 \), \( A_{\text{min}} = N_{\text{plates}} L^2, \) \( A_{\text{max}} = N_{\text{plates}} (pL - p^2 h + h) \) and area extensibility \( e_{\text{area}} = \frac{p - 1 - (h/L)(p^2 - 1)}{1}, \) with maximum \( e_{\text{area}} \) occurring at \( p = L/2h \). Hence for \( L = 10 \text{ nm} \) and \( h = 1 \text{ nm} \), \( h/L = 0.1 \) and maximum \( e_{\text{area}} = 1.60(160\%) \) using \( p = 5 \) plate planes; the minimum plausible \( h/L = 0.01 \), which gives maximum extensibility \( e_{\text{area}} = 24.00(2400\%) \) using \( p = 50 \) plate planes, giving a maximum surface rugosity of ~98 nm at full distension if \( h = 2 \text{ nm} \).

If an external mechanical pressure is applied perpendicular to a parasol surface, the degree of deformation will depend strongly on various details of design. Given the extensive cabling and spring-loading of plates a surface stiffness of \( k_s \approx 10 \text{ N/m} \) should be

\[\begin{align*}
\text{Fig. 5.11. Accordion model.}
\end{align*}\]

---

*The ancient practice of origami (the art of folding three-dimensional objects out of paper without cutting or pasting) has systematically explored the geometries of folded flat sheets; the mathematics of origami is well-studied.*\(^{106-111}\)
Fig. 5.12. Schematic of parasol model.
achievable, in which case a point force of 1 nN deforms the parasol surface by ~0.1 nm.

5.3.2.3 Telescoping Model

Metamorphic surfaces using the Telescoping Model (Fig. 5.14) consist of nested, tight-fitting congruent shapes or blocks under tension fitted tightly into sleeves. The surface expands as progressively smaller segments telescope out from the main body. Typical configurations might include nested annular cylinders, sliding "dominoes" (tethered blocks), the "trombone" (sleeved blocks), or even a multisegment approach. In nature, centipedes, worms, and wasps with extended thoraxes use tubular segments to achieve flexible surfaces.

Given the many sliding interfaces and higher pressures involved, telescoping surfaces should be carefully designed to preclude possible surface oxidation. At high enough pressures sliding diamondoid plates may strip each other's passivating hydrogens, allowing newly exposed carbon atoms in the contacting surfaces to weld the plates tight; unpassivated graphene sheets (Section 5.3.2.4) may be optimal for this configuration.

5.3.2.4 Flexible Fabric Model

Two properties of metamorphic surfaces are most useful: flexibility and extensibility. The graphene sheet (of which fullerene tubes and ellipsoids are made; Section 2.3.2) is probably the ideal example of a flexible but nonextensible isoreal nanofabric that could be used to wrap a nanorobot exterior with any continuous shape having turning radii as small as 1.1 nm. The graphene sheet is a one-atom thick purely carbon surface (no hydrogen passivation yet chemically inert) arranged in a mostly hexagonal array of atoms with an occasional pentagon or heptagon for curvature (Section 5.2.4.1). This sheet has the highest tensile strength of any known 2-dimensional network, about 50 times the strength of high-carbon steel and 1.5 times the strength of crystalline diamond (Table 9.3). It has a higher packing density (atoms/nm²) than any other network made of any other atoms in the periodic table—even a 2-D slice of diamond. The graphene sheet is effectively impermeable under normal chemical conditions—naked carbon atoms or small carbon cluster radicals will not readily bond with it, and even helium atoms up to 5 eV (~40,000 K) just bounce off. The material is probably extremely bioinert (Chapter 15). Graphene is also an excellent conductor of heat (as good as diamond) and electricity (better than copper if appropriately doped) along the plane of the sheet, but is a poor cross-plane conductor.

The ultimate flexible and extensible volume in common experience is the rubber balloon, which readily distends its volume by several orders of magnitude with an areal extensibility of -15,000%. Rubber is a linear polymer of isoprene (C₅H₈)n with individual molecules of n = 1000-5000 heavily cross-linked during vulcanization with 1-2% sulfur, somewhat resembling a network of coiled springs that is easily stretched. Although isoprene surfaces are porous and bioactive, hence inappropriate for nanorobot exteriors, Drexler has proposed an analogous coiled pleat configuration (Fig. 5.15) that could probably be implemented using graphene sheets embedded with reciprocal non-zero-curvature elements to provide the needed countertension. The triple-pleat configuration shown has areal extensibility $e_{\text{area}} \approx 8.00(800\%)$. As with the Accordion Model, a drawback of this design is a propensity for surface fouling in vivo due to the presence of numerous trapping pockets.

Other more dynamic configurations of flexible fabrics may be imagined. Graphene sheets may be unfurled from interior storage, like a spring-roller window shade, a rolled-up sleeping bag, or a serpentin or pito (coiled "party blowouts"). For example, R.C. Merkle notes that the minimum gas pressure differential needed to uncoil and inflate a collapsed, tightly-coiled graphene tube of radius R is given by:

$$p_{\text{unflate}} \approx 2 E_c / R$$  \hspace{1cm} [Eqn. 5.6]

where $E_c$ is the energy of two graphite sheets held together by the van der Waals forces between them. Merkle estimates $E_c \approx 0.25$ J/m² from computer simulations, in line with Kelly's experimental value²²⁸⁰ of $E_c = 0.234$ J/m². Thus a coiled tube with $R = 100$ nm...
requires at least $p_{\text{surface}} \geq 50 \text{ atm}$ to inflate; larger diameter tubes require less pressure differential to uncoil.

Fabrics may be comprised of tightly-woven 1.1-nm fullerene tubes with the length of each such “thread” individually controlled by winding or unwinding from numerous internal spools. In biology, various species of worms achieve linear extensibilities of up to 800% by changing the relative pitch angles of helically wound inextensible collagen fibers. A dynamically reconfigurable support trusswork sheathed in flexible surface materials also permits ready shape-changing; the octet truss geometry is one of the strongest known [P. Salsbury, personal communication, 1997].

5.3.2.5 Block Exchange Model

The most difficult surface to construct is also the most volume-efficient and indefinitely extensible configuration—a surface consisting entirely of replaceable blocks or prisms of fixed geometry. In this model, surface blocks connected by snap-together interfaces are disconnected, rotated, exchanged, moved, and reconnected by subsurface block-moving mechanisms which may include manipulator arms (Section 9.3.1), conveyor systems (Section 3.4.3), and the like. Cytoskeletal components (e.g., subsurface actin microfilaments; Section 8.5.3.11) are similarly lengthened or shortened by adding or subtracting monomeric units during movement and various intracellular transport processes. Blocks combining at least one pentagonal, hexagonal, and heptagonal face can be removed, rotated to expose a new face, then refastened in a new relationship to their neighbors to alter surface concavity or convexity at will. Design constraints include a requirement that each edge of the surface faces must be of the same length, and that blocks should have angles $\leq 90^\circ$ on all edges of the face that will form the surface, so that the face can be flush with adjacent faces [C. Phoenix, personal communication, 1999]. Blocks containing different “letters” from a pattern “alphabet” on various faces (e.g., different Braille-like dots on each surface, like spots on dice cubes) can also be used to post tactile-readable information mechanically on the nanorobot surface, or individual rods in rod arrays can be selectively extended or retracted from the surface to create tactile-readable displays similar to refreshable Braille displays used by the deafblind (Section 7.4.6.1).

Block-built tubular members or protrusions of any other shape may be extruded, lengthened, rotated, or reabsorbed, as required. Internal block exchange allows periodic antifouling maintenance to be performed, eliminating biomatter contamination of the external nanorobot surface. Careful design should permit a waterproof boundary to be maintained throughout the reconfiguration process (Fig. 5.16). The extensibility and reactivity of efficient block exchange models has already been explored in Section 5.3.1.

5.3.3 Metamorphic Power and Control

The power required to control a given section of metamorphic surface may be provided by two simple counterbalancing mechanisms. First, the space between an interior hard shell and the distensible outer surface is pressurized using a working fluid such as gas or water. Second, the pressurization force is resisted by a gridwork of independently controllable nanoscale cables each attached to specific sets of metamorphic surface units (e.g., Fig. 5.12). Applying tension to a cable causes the units to which it is attached to selectively retract; releasing tension allows those same units to distend. This system is crudely analogous to the system of circular and longitudinal muscle fibers found in the intestines, and to the hydrostatic skeleton found in small marine organisms such as earthworms, nematodes, and to a lesser extent in echinoderms, some mollusks, caterpillars and spiders. A counterbalanced surface reverts to its compact shape in the event of a physical breach with escape of working fluid, an important fail-safe design element.

Fully inflating one of the 3-micron long metamorphic fingers described in Section 5.3.1.3 requires filling its 0.015 micron$^3$ interior volume with ~700,000 molecules of N$_2$ gas at a 2-atm working pressure, accomplished in 70 millisec within a 0.1 pW power budget using a bank of 10 sorting rotors (Section 3.4.2). This gas may be stored onboard in a (41 nm)$^3$ (~0.00007 micron$^3$) pressure vessel at 1000 atm (Table 10.2). If energy efficiency is the primary concern, it should be possible to recover most of the compressive sorting power using a generator subsystem to charge an energy storage buffer as the gas passes from the high to the low pressure regime. If distension speed is the primary design goal, gas may be rapidly vented from
the pressure vessel into the finger volume in milliseconds, bypassing the generator subsystem.

Of course, such rapid venting causes the working gas to cool. Ignoring the normally relatively small amount of work done on the external medium, the temperature change during free expansion of a van der Waals gas (Section 10.3.2) is simply:

$$\Delta T_{\text{expand}} = -\left(\frac{A_{vdW} \mu_{gas}}{C_{V_{gas}}}ight)\left(V_{\text{final}}^{-1} - V_{\text{init}}^{-1}\right)$$  \{Eqn. 5.7\}

where $A_{vdW}$ = van der Waals gas constant for intermolecular attraction, $\mu_{gas}$ is the moles of gas present, $C_{V_{gas}}$ is volumetric heat capacity of the gas, and $V_{\text{init}}$ and $V_{\text{final}}$ are the starting and ending gas volumes, respectively. Taking $A_{vdW} = 1.390 \times 10^{-6}$ m$^6$-atm/mole$^2$ (Table 10.1) and $C_{V_{gas}} = 20.8$ J/mole-K for N$_2$ gas, and from the above example $\mu_{gas} = 1.2 \times 10^{-18}$ moles ($\sim$700,000 molecules N$_2$), $V_{\text{init}} = 6.9 \times 10^{-23}$ m$^3$, and $V_{\text{final}} = 1.5 \times 10^{-20}$ m$^3$, then $\Delta T_{\text{expand}} = -120$ K. However, this temperature change is rapidly conducted throughout a micron-scale diamondoid nanorobot structure in $\sim10^{-5}$ sec (Section 4.6.1), since diamond is $\sim3200$ times more thermally conductive than water. The maximum instantaneous temperature decline in a $V_{\text{nanorobot}} \sim 1$ micron$^3$ block of diamond with $C_{V_{\text{diamond}}} = 1.82 \times 10^6$ joules/m$^3$-K is only $\Delta T_{\text{cool}} = \Delta T_{\text{expand}} \frac{C_{V_{\text{diamond}}}}{\mu_{\text{gas}} / V_{\text{nanorobot}}} = -120$ K. However, this temperature change is rapidly conducted throughout a micron-scale diamondoid nanorobot structure in $\sim10^{-5}$ sec (Section 4.6.1), since diamond is $\sim3200$ times more thermally conductive than water.

Controlled power transmission may also be achieved by dividing the extensible surface into compartments, then pressurizing or decompressing adjacent compartments to produce differential movement, or triggering contractile gels in various compartments; by manipulating tension to tendon bundles affixed to the end of a lengthy protuberance to control tip position (Section 9.3.1.2); by mechanically rotating an internal ribcage consisting of noncoplanar ovate sections; by acoustically triggering a progression of locks and ratchets; or by using telescoping screw drives (Section 9.3.1.4) or electrostatic drives (Section 6.3.5) to extend or retract specific surface segments. In biology, butterflies exemplify the first of these methods. Upon emerging from its pupal skin, the insect pulls its abdominal segments inward to raise its blood pressure, which inflates the veins of the wings and expands their membranes.

The dexterity of a controlled protuberance will depend upon the fineness of the control mechanism, the stiffness of the metamorphic configuration, the degree and speed of extension, the applied tip load, special jointing and numerous other factors; -0.1-1 nm positioning accuracy and application of nanonewton forces at the tip should be feasible. Pressure-driven ratchets may transduce acoustic power/control pulses transmitted internally to sensors or end effectors located at the tip.

### 5.3.4 Engulf Formations

Flexible nanomachines may adopt special physical configurations optimally suited for the defense of the human body. Cytocidal mechanisms are discussed in Section 10.4, but consider here a metamorphic spherical nanorobot of radius $r_s$ which reshapes itself into a thin disk of equal volume, thickness $h_d$ and radius $r_d = (4 r_s^2 / 3 h_d)^{1/2}$. This accomplished, the nanodevice next folds itself into a hollow ball, enclosing a spherical cavity of radius $r = r_d / 2$. Areal extensibility $e_{area}$ required to complete the first transformation, the principal source of surface stretching, is given by Eqn. 5.3.

Thus a spherical nanorobot 4 microns in diameter ($r_s = 2$ microns), after first flattening itself into a pancake that is $h_d = 0.1$ micron thick with $e_{area} \sim 12.00(1200\%)$, can then curl into a hollow ball and completely surround a spherical biomass ~8 microns in diameter. This is large enough to engulf virus particles, most known bacteria, most in vivo nanodevices, and most of the formed elements in human blood including platelets, erythrocytes, and lymphocytes.

### 5.3.5 Reconfiguring Surface-Penetrating Elements

The ability to reshape a nanodevice surface implies that some configurational flexibility of external components is required. It must be possible to add, remove, or reposition stable nanodevice elements whose function requires penetrating the metamorphic surface, including sensors (Chapter 4), manipulators (Section 9.3), sorting rotor banks (Section 3.4.2), antigen semaphores (Section 5.3.6), and bulk material intake and outlet ports. A schematic of the process for removing such an element from the surface while maintaining watertightness is shown in Figure 5.17; to install an element, the process is reversed.

Some limited configurational flexibility of internal nanorobot components may also be useful but should be restricted to nonessential or insensitive systems to avoid unnecessarily multiplying system complexity. Any movable trusswork should rotate around a configurationally stable core. Storage areas, nanofactory work volumes, pressure vessels, fluid/tool transfer pathways, and voids are the easiest internal elements to compress or reshape during a nanodevice reconfiguration event. Internal transfer stream mechanisms like conveyor belts may be designed modularly to permit limited modifications of device functionality. A few special purpose internal subsystems may tolerate positional or rotational uncertainty. For example, a clocking subsystem emitting a synchronizing periodic omnidirectional acoustic click may lie in any orientation within the

![Fig. 5.17. Schematic representation of a watertight block surface reconfiguration: removing an embedded sensor element (block motive and control components omitted).](image-url)
nanorobot; acoustic transit time across a 1-micron diamondoid nanorobot is $\Delta t_{\text{min}} = 10^{-10}$ sec $< \Delta t_{\text{min}} = 10^{-9}$ sec (Section 10.1), thus the clock may be positionally insensitive as well.

5.3.6 Presentation Semaphores

It will often be necessary to modify nanorobot surface biochemical characteristics, as for example to present self-antigen to ensure biocompatibility (Chapter 15), to present targeted non-self-antigens to facilitate cytocide (Section 9.4.7) or elimination from the body (Chapter 16), or to create a traveling solvation wave (e.g., hydrophilic, lipophilic) on the nanorobot exterior surface to facilitate cytopenetration (Section 9.4.5.3). At concentrations of 1-10 nanomolar, there are $10^4$-$10^5$ MHC Class I plasma membrane molecules per typical tissue cell, or $10$ MHC proteins/micron$^2$ at the cell surface (Section 8.5.2.1).

The MHC Class I molecule (Section 8.5.2.1) consists of a $45,000$ dalton glycosylated polypeptide chain crudely shaped like a “hand” which is noncovalently grasping a $12,000$ dalton nonglycosylated peptide microglobulin (Fig. 8.33). Eliminating anchor and attachment components which do not participate in recognition leaves an active antigen “semaphore” component measuring roughly $3$ nm $x$ $6$ nm $x$ $8$ nm that must protrude from the cell surface.

To present this semaphore to the external environment at a nanorobot surface, a device roughly analogous to the manipulator arm described in Section 9.3.1.4 may be used. The manipulator arm includes a $7$-nm diameter internal tool transport channel which may be useful in some special applications. Pure diamondoid in the flawless crystalline form will be clear and largely colorless; impurities commonly found in commercial and investment gemstones that impart bluish (~0.1% B substituted for C), yellowish (~0.1% N substituted for C), or other (e.g., blood-red, kelly green) hues are unlikely in nanomanufactured materials, as are graphicitic deposits that produce black streaking or onyxlike coloration. Perception of chromaticity is limited by such factors as the minimum spatial resolution of the human eye (~0.5 arcmin or ~30 microns at closest focus of 10 cm), the maximum wavelength of visible light (~0.7 micron), and the degree of physical proximity of individual colored nanodevices when aggregated.

The simplest method for imparting a specific surface color is to add a thick coating of engineered corundum. Pure corundum (aluminum oxide) is colorless and has a hardness and chemical inertness only slightly inferior to diamond. Sapphire is the best-known crystalline form. Sapphire can be manufactured in a full spectrum of blues, pinks, yellows, teals, lavenders, greens, whites, and all intermediate hues. More colors with greater intensity exist with sapphires than with any other gemstone. The broad color palette is achieved by replacing aluminum atoms with ~0.1% iron atoms and ~0.01% titanium atoms. Rubies, also corundum crystals, achieve their vivid reds by replacing a few aluminum atoms with atoms of chromium. The biocompatibility of sapphire and ruby coatings is largely unexplored. If this should present a problem, a thin overcoating of transparent diamond preserves the color effects while presenting a potentially bioactive interface (Chapter 15) to the external environment. Given the differences in crystal structures and atomic lattice sizes between diamond and sapphire, there may be some sacrifices in materials properties at such interfaces; D.W. Brenner [personal communication, 1999] is

![Diamondoid Structure](image)

Fig. 5.18. Schematic of presentation semaphore mechanism.
unaware of any computational modeling studies having been performed on the atomically-bonded diamond-corundum interface as of 1998.

In larger devices or in macroscopic aggregates, color may be created and dynamically manipulated by embedding micron-scale light-emitting solid-state lasers in the surface. For a surface to appear a certain color, embedded monochromatic optical lasers must emit light of sufficient brightness to compete with other illuminated surfaces that may be present in the visual field. The required intensity is probably ~1 watt/m² (Section 4.9.4) or ~1 pW/micron²—a feasible energy emission budget for micron-scale nanorobots. For example, Shen’ 532 has constructed 0.86-micron thick three-color electrically-tunable organic light-emitting devices which generate ~0.5 pW/micron² with ~1% energy efficiency; red-only LEDs have achieved efficiencies up to 16%,1054 and larger diode lasers have achieved up to 60%-66% efficiencies.3144,3145 Organic light-emitting devices (OLEDs) using small organic molecules can have high brightness (2-4 pW/micron²), half-lifetimes of ~4000 hours, and can be made with a wide range of emission colors in a ~300 nm thick sandwich.3188 White-light organic electroluminescent devices ~100 nm thick have produced ~3 pW/micron² at 15 volts;1048 0.1-1% energy efficiency is typical.1049 (Commercially available LEDs like Hewlett-Packard’s gallium arsenide VCSEL blue-light laser chip have active layers ~10 microns thick and ~11% energy efficiency. LED chips typically have lifetimes ~10⁸ sec, and other microcavity lasers are well-known.3255

If surface color is a serious design objective, arrays of frequency-selective absorbers or emitters such as rhodopsin, fluorescein, carotenoids, luciferins, or engineered porphyrins can be placed below a thin transparent diamondoid window. Active manipulation of covershades to block or expose these windows could permit rapid modulation of chromatic surface characteristics (e.g., polychromatic sparkling). Diverse surface pattern morphologies are theoretically available, dynamic texture arrays can produce marked changes in color and other optical properties.1041 Coherent modulation of the optical characteristics of aggregated nanorobots over large areas allows the creation and control of optical patterns visible to the human eye (e.g., epidermal displays; Section 7.4.6.7) with ~micron resolution at frequencies up to ~10 KHz consistent with a conservative ~1 cm/sec covershade speed. The visibility of subdermal chromomorphic devices is greatly reduced by photon scattering and absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4).

If surface color is a serious design objective, arrays of frequency-selective absorbers or emitters such as rhodopsin, fluorescein, carotenoids, luciferins, or engineered porphyrins can be placed below a thin transparent diamondoid window. Active manipulation of covershades to block or expose these windows could permit rapid modulation of chromatic surface characteristics (e.g., polychromatic sparkling). Diverse surface pattern morphologies are theoretically available, dynamic texture arrays can produce marked changes in color and other optical properties.1041 Coherent modulation of the optical characteristics of aggregated nanorobots over large areas allows the creation and control of optical patterns visible to the human eye (e.g., epidermal displays; Section 7.4.6.7) with ~micron resolution at frequencies up to ~10 KHz consistent with a conservative ~1 cm/sec covershade speed. The visibility of subdermal chromomorphic devices is greatly reduced by photon scattering and absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4). Pure suspensions of spherical nanorobots in clear fluid will partake of the absorption processes at greater depths in tissue (Section 4.9.4).

5.4 Metamorphic Bumpers

Nanorobots may use circumferential metamorphic bumpers to achieve a continuously tight fit to neighboring devices while laboring cooperatively in nanotissues overlaying moving biological surfaces (e.g., see Chapter 22). Bumpers are crudely analogous to the cell-cell adherens junction found in human epithelial sheet cells.531 In this junction, a continuous adhesion belt (the zonula adherens or belt desmosome) surrounds each of the interacting cells in the sheet, resulting in a narrow adhesive zone mediated by Ca²⁺-dependent transmembrane linker glycoproteins. Within each cell, a contractile bundle of actin filaments runs adjacent to the adhesion belt and parallel to the cellular membrane, serving as control cables for this biological junctional “bumper” system. Oriented contraction of the actin filaments causes specific localized movement of the sheet, as for example the rolling of epithelial tissue into a closed tube. Nanorobots linked by narrow equatorial-band or wide whole-wall bumpers into nanotissues may perform similar feats, under continuous computer control.

The other principal class of anchoring junction is the spot desmosome, a buttonlike point of intercellular contact that rivets together certain types of epithelial tissue cells. Each desmosome is a spoon-shaped molecule about 30 nm long, spaced at ~8 nm intervals across the cell surface, with the straight tail of each spoon firmly embedded in a subsurface cytoskeletal plaque and the head of the spoon protruding ~10 nm outside of the cellular lipid bilayer membrane.512,531 Matching spot desmosome heads on neighboring cells adhere, gluing the cells together while maintaining a 10-20 nm intercellular gap.

Tighter junctions (~2 nm) called occluding junctions form nearly watertight seals in certain critical areas such as between the cells lining the digestive system. Blood platelets are also intensely reactive cells that respond to a variety of stimuli to undergo shape change, adhesion, primary and secondary aggregation, and a process known as viscous metamorphosis in which membrane fusion occurs between adjacent platelets with the loss of membrane integrity.530

5.4.1 Nanojunction Mechanisms

Neighboring nanorobots must be able to reliably identify the presence of others and then join together tightly into cohesive nanotissues, forming planar or three-dimensional arrays. In nanodevices, a broad array of fastening technologies from both the biological and the engineering worlds is available for design inspiration.

For simple adhesion, micromechanical velcro,532 sticky tethers,535 magnetic latches (F_min ~ 1600 nN for A_magnet = 1 micron², B = B_magnet = 1.4 tesla; Eqn. 4.45), or other mechanical adhesives offer the simplest means of attachment. Interlocking mechanisms are common in engineered materials (e.g., plug and socket, hook and eye, distension bulbs or dilators)—many of which are simple derivations of two simple machines (the plane and lever) commonly employed in mechanical engineering. Hepatic cells comprising liver walls are connected by “pegs” that fit into depressions in neighboring cells in a crude analogy to snap fasteners (Section 8.2.5, and see below); similar knobbed adhesive interfaces and docking envelopes have been proposed for nanodevices. Complementary knobs of various shapes may also be used on the surfaces of different nanorobot species, or on different faces of the same species, to provide simple mechanical recognition or orientational control. Judicial selection of device shape which mandates specific mateable junction geometries can force a large nanorobot aggregate into desired global configurations (e.g., structured nano-organisms).

Pressed tightly together, two complementary diamondoid surfaces create a strong van der Waals adhesive interface with tensile strength ~3 x 10⁹ N/m²; interposing scattered atomic-scale bumps ~0.2 nm in diameter reduces this adhesion to ~3 x 10⁸ N/m²,30 or ~3000 atm (Section 9.2.1). Providing each metamorphic nanorobot species with complementary Braille-like adhesion plates may provide either a simple means of universal attachment or a unique recognition and interlock pattern of great complexity. Adherent plates are readily detached by inserting corrugated wedge mechanisms or by deforming the adhesive pattern. For comparison, adhesion between biological cells (as determined by measurements of the force required for detachment of ~20-micron tissue cells by micromanipulation, shear, or jets of fluid) is typically ~1-100 nN/micron².1454,1455,1458,1459
Reversible adhesion may also be achieved using surface-mounted matched nanomechanical manipulator arms (Section 9.3.1) which can grasp or release their counterpart in an adjacent cell, under computer control.

5.4.2 Transbumper Communication

Besides occluding and anchor junctions, mutually attached biological cells use gap junctions or "communicating junctions" to exchange inorganic ions and other small water-soluble molecules like sugar and ATP. Connexons are hexagonal transmembrane proteins with tiny pores; when connexons in the membranes of two cells are aligned, they form a continuous aqueous channel linking the two cell interiors that can open and close in a way similar to the acetylcholine receptor channel described in Section 3.3.3. When the calcium level of a cell rises—often a signal that the cell is sick or compromised—connexons close and quarantine the unhealthy neighbor. A force of 6-10 pN is required to physically pry apart each connexin-32 hepatic cell gap junction unit, each unit a combined pair of hexagonal cylinders ~7.5 nm in height and 7 nm in diameter with a 2-nm pore through the center. Another communicative adhesive mechanism is provided by bacterial sex pili, filamentous surface adhesive organelles that are long thin hollow protein tubes (typically ~300 per cell) with sticky receptors on their ends that bind to recipient cell walls, allowing DNA to pass from cell to cell. Integrins (Section 8.5.2.2) mediate cell binding to extracellular matrix and anchorage-dependence mechanical signals.

Similar mechanisms may be employed in metamorphic bumper walls to establish stable physical conduits among neighboring nanorobots in an array. Such conduits may transmit information, materials, nanoscale components, power (acoustic or hydraulic), or even structural forces (including tension, compression or torsion) to provide stability or movement of the entire nanotissue assemblage. Hinges and related mechanisms allow contact to be maintained as the metamorphic surface expands, contracts, twists or translates.

5.4.3 Bumper Mechanics

Methods of inflation and deflation of metamorphic volumes have been described in Section 5.3.3. Interbumper connection forces may be $10^7$-10^9 N/m^2, easily exceeding potentially disruptive forces typically encountered inside the human body. Consistent with available power resources and likely operating environments, metamorphic bumpers may be driven at frequencies of 10-1000 Hz at full-range amplitude, or faster in cases of limited-range amplitude oscillations. This response rate matches or exceeds the speeds of all important mechanical tissue movements found in the human body including respiration (0.1-1 Hz), pulse (1-2 Hz), muscle contractions during violent exercise (1-10 Hz) or tetanic contractions (15-60 Hz), and even nerve cell discharge events (5-100 Hz). The fastest known animal muscle speeds are found in the ~90 Hz rattleshaker of the western diamondback rattlesnake; the ~90 Hz hummingbird wingbeat; the ~200 Hz toadfish swim bladder; the synchronous muscle contractions of the katydid, Neoconocephalus robustus, at 212 Hz, and the cicada, Chlorocysta viridis, at 224 Hz, and finally the record 1046-2200 Hz wingbeat of the tiny midge Forcipomyia.
6.1 Nanodevice Energy Resources

Device energetics may represent the most serious limitation in nanorobot design. Almost all medical nanodevices will be actively powered. Mechanical motions, pumping, chemical transformations and the like all require the expenditure of energy, measured in joules. Even a drug molecule interaction with a biological receptor site reduces free energy by ~50 kT (Section 3.5.2) or ~210 zJ (1 zeptojoule (zJ) = 10^{-21} joule) at 310 K. Heat dissipation is also a major consideration in nanomachine design, particularly when large numbers of nanomachines are deployed. Power, of course, is the rate of energy consumption or production, measured in joules/sec or watts.

Energy, like money, has both a storage and a transactional character. Section 6.2 reviews the various forms of stored energy that may be accessible to working nanodevices in vivo. Section 6.3 describes how one form of energy can be converted into another form, while Section 6.4 discusses how energy may be transmitted from one place to another—both representing the transactional aspect of energy use. Section 6.5 closes the Chapter with an enumeration of issues and techniques useful in assessing energy requirements and performance restrictions in medical nanodevice design.

6.2 Energy Storage

For medical nanorobots, onboard volume is a precious and limited commodity. Viscous forces dominate inertial and gravitational forces, so mass is almost irrelevant. Hence energy stored per unit volume (joules/m^3) is an appropriate figure of merit for nanoscale energy storage devices.

Energy storage devices may be required to maintain temporary power during subsystem failures, metamorphic transitions, or during temporary unavailability of environmental energy resources; to provide supplementary supplies during brief periods of overcapacity consumption; or to buffer normal energy usage among subsystems generating large fluctuations. Stored energy may also be used to power short-lived medical nanodevices on missions of limited duration. For example, a 1 micron^3 storage device with storage density of 2 kT/nm^3 (~10^7 joules/m^3) contains sufficient energy to power a 10 picowatt (pW) nanorobot for ~1 second. Chemical storage devices (providing up to 10^{11} joules/m^3; Section 6.2.3) may extend power short-lived medical nanorobots on missions of limited duration. For example, a 1 micron^3 storage device with storage density of 2 kT/nm^3 (~10^7 joules/m^3) contains sufficient energy to power a 10 picowatt (pW) nanorobot for ~1 second. Chemical storage devices (providing up to 10^{11} joules/m^3; Section 6.2.3) may extend this duration to 10^4 sec (~3 hours).

6.2.1 Gravitational Energy Storage

Not surprisingly, the gravitational field offers one of the weakest forms of energy storage available to nanodevices. The energy density attainable using a storage device of density \( \rho \) and size \( L \) (height) in a gravity field \( g \) is

\[
E_{\text{storage}} = \rho g L \quad \text{(joules/m}^3) \quad \text{(Eqn. 6.1)}
\]

For \( \rho = 2000 \text{ kg/m}^3 \), \( g = 9.81 \text{ m/sec}^2 \) and \( L = 1 \text{ micron} \), \( E_{\text{storage}} = 2 \times 10^{-2} \text{ joules/m}^3 \).

6.2.2 Mechanical Energy Storage

6.2.2.1 Pendulums and Springs

Energy may also be stored in mechanical systems. A gravitational pendulum of cord length \( r \) with a bob of density \( \rho \) and characteristic diameter \( L \) that is tangentially displaced a distance \( \Delta x \) has potential energy \( -\rho L^3 \Delta x \) stored in a volume of order \( -r L \Delta x \) for small \( \Delta x \), hence energy density is approximately:

\[
E_{\text{storage}} = \rho g L^2 / r \quad \text{(joules/m}^3) \quad \text{(Eqn. 6.2)}
\]

Taking \( \rho = 2000 \text{ kg/m}^3 \), \( g = 9.81 \text{ m/sec}^2 \) and \( L = r = 1 \text{ micron} \), then \( E_{\text{storage}} = 2 \times 10^2 \text{ joules/m}^3 \), the same as Eqn. 6.1, as expected.

Stretched springs provide significantly greater energy storage capacity. For example, a diamondoid spring of size \(-L\) and stretching stiffness \( k \), has harmonic potential \((1/2) k x^2\) for a displacement \( x \) and volume \( L^3 \), with energy density:

\[
E_{\text{storage}} = k x^2 / 2 L^2 = \left( \frac{E}{L} \right) \left( \frac{x}{L} \right)^2 \quad \text{(Eqn. 6.3)}
\]

where \((x/L) = \text{strain}\) and \(E = \text{Young's modulus}\). Conservatively taking strain = 5% and \(E = 1.05 \times 10^{12} \text{ N/m}^2 \) for diamond, then \( E_{\text{storage}} = 1.3 \times 10^8 \text{ joules/m}^3 \). However, strains may be applied in three dimensions as well as in tension, shear, or torsion, so total energy storage may be somewhat higher. The fracture surface energy of the weakest \([111] \) diamond plane \( E_f = 5.3 \text{ joules/m}^2 \), and the distance between \([111] \) planes is \( L_{\text{plane}} = 0.24 \text{ nm} \) given that there are \(-1.8 \times 10^{10} \text{ bonds/m}^2 \), so the theoretical maximum mechanical energy storage density in a diamond block is \( E_{\text{storage}} = E_f / L_{\text{plane}} = 2 \times 10^{10} \text{ joules/m}^3 \).

6.2.2.2 Flywheels

The maximum energy that can be stored in an axially spinning flywheel is \( E_{\text{flywheel}} = (1/2) I_{\text{fly}} \omega_{\text{max}}^2 \), where \( \omega_{\text{max}} \) is the maximum angular velocity (bursting speed) given by Eqn. 4.17 and \( I_{\text{fly}} = (1/2) \pi r^4 h \) \( \text{m}^2 \) is the rotational inertia of a disk of radius \( r \), height \( h \), and mass \( m = \pi r^2 h \rho \). Dividing by disk volume, all geometric variables drop out giving maximum \( E_{\text{storage}} = (1/2) \omega_{\text{max}}^2 \sigma_w \) for flywheels of any size. Adopting a fairly aggressive working stress \( \sigma_w = 10^{11} \text{ N/m}^2 \), \( E_{\text{storage}} = 5 \times 10^{10} \text{ joules/m}^3 \), an excellent power density. In 1998, commercially available electric micromotors already achieved energy storage densities of \(-10^5 \text{ to } 10^6 \text{ joules/m}^3 \).

More intuitively, J. Sidles points out that the energy that can be stored in a compact flywheel is of the order of either:
1. the chemical binding energy per atom, times the number of atoms in the flywheel (same as chemical energy storage), or
2. the yield stress of the material comprising the flywheel, times the volume.

Both methods give estimates that agree to within an order of magnitude. Thus flywheels offer little advantage over chemical fuels in terms of stored energy density (Section 6.2.3), regardless of their size or material composition.

If high-velocity flywheels must ride on bearings of large stiffness, this may cause sufficient frictional drag to render flywheel energy storage impractical even for vacuum-isolated systems. Consider a cylindrical sleeve bearing of radius $r_{bear}$ and length $l_{bear}$ axially supporting the flywheel described in the previous paragraph with axial bearing stiffness of $k_b = 1000 \text{ N/m}$ and bearing surface velocity $v_{bear} = v_{fly} \left(\frac{r_{bear}}{r}\right)$ and flywheel rim velocity $v_{fly} = \left(4 \frac{E_{storage}}{\rho}\right)^{1/2}$. According to Drexler,\(^{10}\) the drag power for a nanoscale bearing is dominated by band-stiffness scattering as:

$$P_{\text{drag}} = \left(1.42 \times 10^{-31}\right) \frac{k_b^{1/2}}{l_{bear}^{1/2}} \rho v_{fly}^2 \quad \text{(watts)} \quad \text{[Eqn. 6.4]}$$

The energy initially stored in the flywheel is:

$$E_{fly} = \left(\frac{\pi}{4}\right) r^2 h \rho \frac{v_{fly}^2}{2} \quad \text{[Eqn. 6.5]}$$

so the time required for the flywheel to lose half of its energy due to drag friction, its “energy half-life” or $\tau_{1/2}$, is:

$$\tau_{1/2} = \frac{E_{fly}}{P_{\text{drag}}} \ln(2) \quad \text{[Eqn. 6.6]}$$

Thus a flywheel of radius $r = 200 \text{ nm}$ and thickness $h = 20 \text{ nm}$ ($m = 9 \times 10^{-18} \text{ kg}$ for diamond) supported by a bearing of radius $r_{bear} = h/2 = 10 \text{ nm}$ and length $l_{bear} = h = 20 \text{ nm}$ with maximum flywheel energy density $E_{storage} = 5 \times 10^{10} \text{ joules/m}^3$ has $v_{fly} = 2400 \text{ m/sec}$, $v_{bear} = 120 \text{ m/sec}$, $E_{fly} = 13 \text{ picojoules (pJ)}$, $P_{\text{drag}} = 25 \text{ pW}$, so $\tau_{1/2} = 0.35 \text{ sec}$ and the flywheel loses 99% of its energy in just 2.3 sec. A much larger ~1 micron\(^3\) flywheel of radius $r = 500 \text{ nm}$, thickness $h = 1500 \text{ nm}$ and $v_{bear} = 50 \text{ nm}$ loses 99% of its energy in just 140 sec.

On the other hand, it may be possible to substantially improve the energy storage time if a bearing with much lower drag is available. R. Merkle points out that a spinning, perfectly symmetric diamondoid disk made of isotopically pure carbon and supported by two coaxial carbyne rods on either side should exhibit negligible bearing losses and no vibration. The carbyne rod is one dimensional, and the electron cloud along the rod is rotationally symmetric (for single and triple bonds), so the disk cannot readily interact with its rod supports. The major source of energy loss is off-axis rotation during spin-up which could induce vibration in the carbyne support, so spinning the disk up to speed may require larger bearings which could be disengaged after spin-up.

A lateral velocity sufficient to destroy the C-C carbyne bonds ($E_{\text{thermal}} = 550 \text{ zJ/bond}$) is $v_{\text{destroy}} = \left(2 E_{\text{lateral}} / m\right)^{1/2} = 0.4 \text{ m/sec}$ for $r = 200 \text{ nm}$ and $h = 20 \text{ nm}$—much larger than the flywheel thermal velocity $v_{\text{thermal}} = 0.04 \text{ m/sec}$ (Eqn. 3.3). Since $F \sim 30 \text{ N}$ of force will break a C-C bond (Section 4.4.1), a lateral acceleration sufficient to destroy the support rods requires $F/m \sim 0.3 \times 10^9 \text{ g's}$, far in excess of the ~0.4 g's typically anticipated for 1-micron spherical nanorobots in vivo (Section 4.3.3.2). Thermal noise will cause the flywheel housing and supports to jiggle, but the phonons will be transmitted along a one-dimensional pathway (the carbyne rod) which cannot couple to the disk rotation. Energy losses due to rotational resistance (as the nanorobot constantly rotates to new spatial orientations) can be made negligible using a nanogyroscopic gimballed housing (Section 4.3.4.1). Safety issues must also be addressed, such as providing energy-absorbing device housings that are highly explosion-resistant.

### 6.2.2.3 Pressurized Fluids

Compressed fluids can store mechanical energy limited only by the tensile strength and aspect ratio of their container (Section 10.3.1) and the rupture strength of their valving system. A “conservative” working stress for diamondoid pressure containers is $\sigma_w = 10^{10} \text{ joules/m}^3$ (~100,000 atm). Gases or liquids, either of which are compressible, may be used as the working fluid depending on energy density and design buoyancy requirements (Section 10.3.6). For example, cycling between 1-1000 atm of pressure at 310 K the density of compressed water varies from 993.4-1038.0 kg/m\(^3\) while the density of compressed oxygen varies from 1.26-670 kg/m\(^3\) (Table 10.2). The energy stored in a compressed fluid is $\sim \sqrt{\text{Joules/Vol}}$; thus at the ~1000 atm cycle limit, the compressed water stores ~2 x 10\(^6\) J/m\(^3\) and the compressed gas stores ~1 x 10\(^6\) J/m\(^3\). (Materials are thermodynamically forbidden to have negative volume compressibilities. However, lanthanum niobate and a few other crystals exhibit negative linear compressibilities; such materials may be used to fabricate porous solids that either expand in all directions when hydrostatically compressed with a penetrating fluid or behave as if they are incompressible.\(^{109}\))

For safety reasons, fluid pressures in excess of 1000 atm ($10^8 \text{ joules/m}^3$) should rarely be used in nanomedical systems (Chapter 17). Any mechanical energy storage system with readily accessible catastrophic energy release modes that is operated near maximum capacity (e.g., fluids >10\(^3\) atm) in vivo risks causing significant damage to nearby tissue cells in the event of device malfunction or rupture.

### 6.2.3 Chemical Energy Storage

Chemical energy storage offers an ideal combination of high storage density, abundant physiological resources, and superior safety in the event a device is physically compromised. As shown in Table 6.1 (values computed from stoichiometric reaction formulas), chemical energy storage density (fuel only\(^{*}\)) ranges from $10^8 \text{ joules/m}^3$ to $10^{11} \text{ joules/m}^3$, as compared to $10^6$-$10^9 \text{ joules/m}^3$ for mechanical media and $10^8$-$10^9 \text{ joules/m}^3$ for electric/magnetic storage (Section 6.2.4). Electrochemical batteries typically achieve $10^6$-$10^9 \text{ joules/m}^3$ (e.g., 5-micron thick rechargeable lithium microbatteries have energy density ~2 x 10\(^7\) joules/m\(^3\)).\(^{588}\) and tin-lithium material can reach ~8 x 10\(^9\) joules/m\(^3\)).\(^{715}\) As of 1998, prototype ~4000 nm\(^3\) Cu/Ag nanobatteries had demonstrated ~7 x 10\(^7\) joules/m\(^3\) or ~2 x 10\(^8\) watts/m\(^3\) power density.\(^{589}\)

Table 6.1 offers many interesting revelations:

1. Assuming “safe” ~1,000 atm storage, a pressurized hydrogen/oxygen mixture has the greatest energy per unit mass (excluding the mass of the containment vessel) but has almost the lowest energy per unit volume, hence is too inefficient for

\(^{*}\)Not all fuels require oxidizer, e.g., ATP. For fuels that require oxidizer, taking oxidizer volume into account (e.g., O\(_2\) at 1000 atm) reduces overall energy storage density to ~7% of fuel-only density for diamond, ~25% for most organics, and ~65% for hydrogen. Also, note that $E_{avail} = E_{storage} - E_{activation}$ where $E_{avail}$ is available energy and $E_{activation}$ is the net unrecoverable activation energy required to release the stored chemical energy.)
Table 6.1 Energy Storage Density for Various Chemical Fuels

<table>
<thead>
<tr>
<th>Energy Storage Fuel</th>
<th>Storage Density (joules/m³)</th>
<th>(joules/kg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP 1.4 x 10⁸</td>
<td>1.0 x 10⁵</td>
<td></td>
</tr>
<tr>
<td>H_2 @ 10⁵ atm (g) 4.9 x 10⁹</td>
<td>1.2 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Nitroglycerine 1.0 x 10¹⁰</td>
<td>6.3 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Glycine (amino acid) 1.0 x 10¹⁰</td>
<td>6.5 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Wood 1.1 x 10¹⁰</td>
<td>1.5 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Urea 1.4 x 10⁰</td>
<td>1.1 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Methanol 1.8 x 10⁰</td>
<td>2.2 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Vegetable protein 2.3 x 10⁰</td>
<td>1.7 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Acetone 2.4 x 10⁰</td>
<td>3.1 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Glucose 2.4 x 10⁰</td>
<td>1.6 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Glycogen (starch) 2.5 x 10⁰</td>
<td>1.8 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Animal protein 2.5 x 10⁰</td>
<td>1.8 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Carbohydrate 2.6 x 10⁰</td>
<td>1.7 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Gasoline 2.8 x 10⁰</td>
<td>4.4 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Butane 3.0 x 10⁰</td>
<td>4.9 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Fat 3.3 x 10⁰</td>
<td>3.9 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Palmitic acid (lipid) 3.3 x 10⁰</td>
<td>3.9 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Palmitin (lipid) 3.4 x 10⁰</td>
<td>4.0 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Leucine (amino acid) 3.5 x 10⁰</td>
<td>2.7 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>Cholesterol (lipid) 4.2 x 10⁰</td>
<td>3.9 x 10⁷</td>
<td></td>
</tr>
<tr>
<td>H_2 @ 10⁵ atm 7.2 x 10⁰</td>
<td>1.2 x 10⁸</td>
<td></td>
</tr>
<tr>
<td>Diamond 1.2 x 10¹¹</td>
<td>3.3 x 10⁸</td>
<td></td>
</tr>
</tbody>
</table>

free-swimming medical nanorobots where interior volume is a scarce commodity.

2. The classic explosive, nitroglycerine (included solely for comparison), ranks poorly on either measure of energy density.

3. ATP, sometimes proposed as an alternate energy source for in cyto operations, has the lowest energy density on the list, though its use cannot be excluded in the earliest-generations of medical nanodevices for metering out small energy packets of known magnitude.

4. Lipids and fats generally store more energy per unit volume than carbohydrates, proteins, or even hydrogen.

5. Cholesterol, twice as plentiful as serum glucose on a gm/cm³ basis (Appendix B), is in theory the most favored biochemical energy storage molecule with nearly double the storage density of glucose.

6. Diamond has the highest known oxidative chemical storage density, possibly surpassed only by fullerene materials, probably because it also has the highest atomic density per unit volume.

7. Hydrogen stored in solid form (H_2 solidifies at ~57,000 atm at room temperature, making 600 kg/m³ crystals), yields an extraordinarily high chemical energy density, second only to diamond.

Combining the values for molecular energy density given in Table 6.1 with the bloodstream concentrations given in Appendix B reveals that the chemical energy content of blood plasma (~ 3 liters in adult males) includes -3 x 10⁶ joules/m³ for all lipids (~10⁷ joules/m³ from total cholesterol alone), -2 x 10⁶ joules/m³ for serum glucose, -1 x 10⁶ joules/m³ for each of the common amino acids, and even -4 x 10⁶ joules/m³ for urea. Cytoplasmic chemical energy content is of similar magnitude, with the addition of -6 x 10⁶ joules/m³ for ATP. (See Table 6.4 for related data.)

The question naturally arises whether some biocompatible artificial energy molecule could be added to the human bloodstream to provide a supplementary chemical energy source for a working in vivo nanorobot population. To this end, an injection of ~0.7 cm³ of diamond colloid (the most energy-dense chemical fuel known—it has been used as rocket fuel) encapsulated in trillions of suitable nanoscale biocompatible carrier devices provides an energy resource equal in size to the entire serum glucose supply, a negligible -0.01% addition to whole blood volume. Ten trillion 0.1-micron³ passive carriers would have a mean separation of ~0.1 microns in the blood.

6.2.4 Electric and Magnetic Energy Storage

The energy density in a static electric field of strength E traversing a material of dielectric constant \( \varepsilon \), is given by:

\[
E_{\text{storage}} = \frac{1}{2} \varepsilon_0 \varepsilon R \]

where \( \varepsilon_0 = 8.85 \times 10^{-12} \text{ farad/m} \) (permittivity constant) and dielectric constant \( \varepsilon = 5.7 \) for diamond. Electrostatic motors (Section 6.3.5) in nanomechanical systems may exhibit an electric field strength of ~0.2 x 10⁸ volts/m. However, the maximum field that may be employed in an electrostatic energy storage device is limited by the dielectric strength or breakdown voltage \( E = 2 x 10⁹ \text{ volts/m} \) for diamond (about the highest known for any material), giving a maximum electric storage density of 1.0 x 10⁷ joules/m³.

What about magnetic storage density? Since isolated magnetic poles (analogous to the electron) are not known to exist, magnetic field energy can be stored only in an array of aligned atomic dipoles. The energy density of the static magnetic field of a permanent magnet comprised of atoms with dipole moment \( \mu \) and number density \( N \), producing a flux density \( B \) at 100% saturation is given by:

\[
E_{\text{storage}} = \mu_B N \]

For iron atoms with bulk density 7860 kg/m³, then \( N = 8.5 \times 10^{28} \text{ atoms/m}³ \) and \( \mu_B = 1.8 \times 10^{-23} \text{ ampere-m}² \), giving \( E_{\text{storage}} = 2.1 \times 10⁶ \text{ joules/m}³ \).

Only a negligible amount of magnetic energy is stored in a magnetic field created by a permanent current loop in a nanoscale ring of superconducting material. For a wire loop of radius \( R \) carrying current \( I \), the peak magnetic flux density is \( B = \mu_0 I / 2 R \) at the center of the loop, so the peak energy density is given by:

\[
E_{\text{storage}} = \frac{\mu_0 I^2}{8 R}
\]

Aluminum conductors in integrated circuits are limited to \( I_d = 3 \times 10^{10} \text{ ampere-m}² \) due to electromigration; thin-film high-temperature superconductors have achieved \( I_d > 3 \times 10^{10} \text{ ampere-m}² \). Taking \( I_d = 10^{10} \text{ ampere-m}² \) (for \( d_{\text{wire}} = 0.1 \text{ micron} \)), and \( \mu_0 = 1.26 \times 10⁶ \text{ henry/m} \), then \( E_{\text{storage}} = 6.3 \times 10⁶ \text{ joules/m}³ \).

Electromagnetic waveguides, radiator cavities and fiberoptic closed loops are too lossy or of inappropriate scale to permit direct
nanodevice photonic energy storage. Energy stored in excited or partially ionized molecular or atomic states, coherent (lasing) and fluorescing media, and enzymatic activated complexes (e.g., at peak activation energy) generally also lack sufficient duration or stability to be useful, although the metastable excited electronic 2 3S state of solid He at 19.8 eV has a 2.3-hour lifetime and thus a theoretical storage density of 5 x 10^{11} joules/m^3 for 100% electronically excited solid helium.661

6.2.5 Nuclear Energy Storage

By nanotechnological standards, the energy stored in atomic nuclei is huge. For instance, the energy density of an un-ionized radioactive atom of U^{235} is 1.5 x 10^{18} joules/m^3, counting the kinetic energy of all fissionable decay products in the total. Hydrogen that undergoes fusion into helium actually provides a much poorer volumetric energy storage density than for fission, ~4.4 x 10^{16} joules/m^3 (again assuming storage of fuel as molecules), largely due to the comparatively high atomic number density of fissionable heavy metals. The highest practical energy density would be achieved by storing some theorized form of matter-friend stabilized antimatter perhaps converted to a two-phase hypergolic (self-igniting) fuel,656 up to a maximum of ~2 x 10^{14} joules/m^3 (fuel only) for platinum/antiplatinum annihilation. The difficulty, of course, is accessing this potential resource in a controlled and well-shielded fashion (Section 6.3.7).

6.3 Power Conversion

Almost all energy available to biological processes on Earth originates in the consumption of the highest naturally-available energy density resource—nuclear fuels in the Sun. After transfer across space via high-energy photons, this energy is absorbed via photosynthesis by plant life on Earth and converted into chemical energy stores of moderate energy density. These chemical energy stores are then consumed by animal life and converted into mechanical or electric energy stores of lower energy density, or are completely degraded to heat.

Medical nanodevices join this energy economy by consuming onboard energy stores (Section 6.2) or by absorbing fresh energy resources from the environment, converting these resources into other forms to accomplish useful work and possibly recovering for reuse a portion of this energy via reversible or regenerative processes, then finally releasing heat as the ultimate outcome of irreversible processes. Clearly the key to medical nanorobot power supply is the efficient conversion of energy from one form to another.

The energy conversion matrix in Table 6.2 provides a convenient conceptual framework with which to organize and guide our discussion. Representative technologies are listed for each cell in the matrix. An exhaustive discussion of every matrix element is beyond the scope of this book. Instead, a selection of the most important categories are illustrated with one or two specific examples. Serial chains of multiple conversion processes may in some cases provide increased efficiency over competing pathways, conserve volume or mass, provide faster conversion (e.g., permit higher power density), allow partial energy regeneration, or serve other specific design objectives.

A note on nomenclature: In this book, energy conversion processes are named using the source energy first, followed by the resultant energy form. Thus a device which converts chemical energy into acoustic energy employs a “chemoacoustic” process, and so forth.

6.3.1 Thermal Energy Conversion Processes

The second law of thermodynamics says that it is impossible to convert heat into useful work if the heat reservoir and the device are both at the same temperature, as demonstrated by Feynman’s classical example of the Brownian motor using an isothermal ratchet and pawl machine,611 although nonequilibrium fluctuations, whether generated by macroscale electric fields or chemical reactions far from equilibrium, can drive a Brownian motor.696 It has also been suggested that reversible-energy-fluctuation converters can obtain useful electrical work from thermal Nyquist noise, up to power densities of 10^{15}-10^{16} watts/m^3 at ~300 K.1606,1607

Of course, a reversible Carnot-cycle heat engine can extract useful work from even a small temperature differential with a Carnot efficiency of e% = ΔT / T. For example, a nanorobot circulating with the blood between core and peripheral tissues may experience a temperature variation up to several kelvins during each vascular circuit of duration t_{circ} ~ 60 sec (Section 8.4.1). From this small temperature differential an ideal biothermal thermomechanical engine may extract a maximum power:

\[
P_n = \frac{V_n C_V \Delta T e\%}{t_{circ}} = \frac{V_n C_V (T_2 - T_1)^2}{T_2 t_{circ}} \quad \text{[Eqn. 6.10]}
\]

where nanorobot thermal storage volume is V_n = 1 micron^3, heat capacity C_V = 4.19 x 10^5 joules/m^3-K for a device filled with water, T_2 = 310 K at the human body core and T_1 = 307 K at the periphery. The thermal store, vacuum-isolated to prevent heat loss (see below), is equilibrated in the hotter core environment to T_2, which heat is then stored until the device reaches the cooler peripheral environment at T_1. From Eqn. 6.10, this temperature differential yields at most P_n ~ 0.002 pW with efficiency e% = (T_2 - T_1) / T_2 ~ 0.01(1%) and a peak (accessible) energy density of P_n \cdot t_{circ} / V_n ~ 10^5 joules/m^3. The change in temperature can be made to cause gas in a three-dimensional coiled piston to slowly expand or contract, driving a rod back and forth thus providing a cyclical linear mechanical output, a Stirling engine configuration. The gas expansion is isobaric and reversible because thermal equilibrium time t_{EQ} ~ V_n C_V / h K_t = 10^5 sec for a conduction layer of thickness h ~ 0.5 micron and thermal conductivity K_t = 0.623 watt/m-K for water at 310 K, so t_{EQ} << t_{circ}.

(Exploiting the diurnal variation in mean body temperature, typically ranging from 309.3 K in early morning to 310.4 K in the evening, produces at most ~2 x 10^7 pW of power.) A nanorobot resting on the epidermal surface may exploit the temperature differential between skin and air, up to 8-13 K (Section 8.4.1.1) giving a maximum Carnot efficiency e% ~ 0.04 (4%); for classical radiative transfer (see below), the nanorobot develops a net power through an L = 10 micron^2 epidermal contact surface of P_n - \sigma (T_2^4 - T_1^4) L^2 e\% = 0.04 pW.

Nakajima541 has built and operated a 50 mm^3 Stirling engine working at 10 Hz between 273-373 K producing 10^2 watts (power density 2 x 10^5 watts/m^3), and has demonstrated the theoretical engineering feasibility of microscale Stirling engines. In 1993 Jeff Snieckoski of Sandia National Laboratories constructed a 50-micron steam engine on a silicon chip producing forces ~100 times higher than those of electrostatic motors of similar size.3486 (The steam was produced electrically.) Computer simulations of a molecular-scale steam engine have been performed by Donald W. Noid at Oak Ridge National Laboratory.3488 A conservative and practical upper limit to nanorobot Carnot efficiency is probably ~50% (T_2 = 620 K).
A heat engine may exploit the temperature difference between the largely isothermal human body acting as a sink and a hot, high-capacity source of stored heat energy. Because the rate of conductive heat loss is scale-dependent, such exploitation is not feasible in nanodevices relying on stored heat sources unless a vacuum isolation suspension is employed (Section 6.3.4.4). As a simple demonstration, consider a vacuum-isolated spherical thermos bottle filled with a hot working fluid of heat capacity $C_V$ at initial temperature $T_2$. Conduction and convection are eliminated; heat loss in vacuo occurs only by radiative transfer. In the classical macroscopic formulation, radiated power $P_r = \frac{4}{15} \pi \sigma e (T_2^4 - T_1^4)$ (watts), where $\sigma = 5.67 \times 10^{-8}$ watts/m$^2$-K$^4$ (Stefan-Boltzmann constant). The thermal energy contained in the hot material is $H_r = \frac{4}{3} \pi r^3 C_V (T_2^2 - T_1^2)$ (joules); hence the time required for half of the energy to radiate away is:

$$\tau_{1/2} = \frac{H_r}{P_r} = \frac{8}{15} \pi \sigma e (T_2^2 - T_1^2)$$  \hspace{1cm} \{\text{Eqn. 6.11}\}

For $r = 1$ micron, $C_V = 4.2 \times 10^6$ joules/m$^3$-K for water, $e = 0.02$ for polished silver, $T_1 = 310$ K inside the human body, and $T_2 = 350$ K up to 647 K (~critical temperature of water at 218 atm pressure), $\tau_{1/2} \approx 6$ sec (at $T_2 \approx 350$ K) starting from an initial (accessible) energy density of $\approx 10^8$ joules/m$^3$; $H_r = 1.5$ nanojoule for a 1-micron core at 647 K. Almost the entire thermal energy store (~99%) leaks away in just 40 sec (at $T_2 = 350$ K). Smaller thermos bottles leak even faster, due to the $r^{-1}$ dependence of $\tau_{1/2}$.

Radiators lying within <1 micron of a lower-temperature material surface exhibit near-field anomalous radiative transfer (Section 6.3.4.4 (E)) and thus exhibit different cooling characteristics than Eqn. 6.11 predicts. Taking $P_r = P_{\text{n}}$ from Eqn. 6.21 for spherical surfaces $<200$ nm apart, then $\tau_{1/2} \approx 0.01 (h^2 c^2 / k^3) (r C_V / \sigma_{\text{cond}} T_2^2)$ (seconds); for $T_2 = 647$ K and $r = 1$ micron, $\tau_{1/2} \approx 10$ sec using a germanium shell but $\approx 10^{-3}$ sec using a boron shell.

Other thermomechanical transducers include sandwich cantilevers (Section 4.6.3) made of composite materials with high coefficients of linear expansion (e.g., heated metal bimorphs$^{357}$). Nitinol or other temperature-sensitive shape-memory alloys,$^{548}$ thermally-driven phase-change microactuators,$^{545}$ thermally-powered contraction turbines,$^{357}$ and thermally-driven contractile proteins.$^{1261}$ Thermochemical transducers that make use of thermal energy stored as a
phase change of a refrigerant can display energy densities of $-10^8$ joules/m$^3$,\textsuperscript{1097} and a thermoacoustic Stirling engine with no moving parts has been demonstrated.\textsuperscript{3267}

A thermoelectric transducer may be constructed from a crystal with piezoelectric properties. When such a crystal is heated or cooled, charges are produced on its surfaces (called pyro-electricity\textsuperscript{553} or heat electricity)\textsuperscript{3088} setting up mechanical strains in the crystal that produce the same electrical effect as the application of external forces in piezo-electricity.\textsuperscript{551} Both bone and tendon exhibit the pyro-electric effect;\textsuperscript{3088} all pyroelectric materials are piezoelectric, though the converse is not true.\textsuperscript{3089} Thermocouples are another example of direct thermoelectric energy transduction, and thermophotovoltaic generators are well-known.\textsuperscript{1983}

A high emissivity blackbody radiator provides thermooptical transduction at temperatures >650 K, increasing in efficiency up to -6000 K.

### 6.3.2 Mechanical Energy Conversion Processes

It has been suggested\textsuperscript{520} that “simple physical shaking” may be sufficient to provide power to nanorobots, much like a self-winding watch that employs mechanical rectification. Material moves through human lymphatic vessels (Section 8.2.1.3) in this manner: The lymphatics have no direct musculature (or cilia) of their own, but depend upon external physiological motion sources such as blood vessel contractions and skeletal movements to induce periodic pressures on the lymphatic vessels, moving material along via one-way valving. Experiments using respiratory chambers found that spontaneous physical activity or “fidgeting” consumed an average of 348 Kcal/day (range 100-700 Kcal/day) or ~17 watts of power\textsuperscript{2935} during normal respiration can obtain $P_n \sim 3 \times 10^{-6}$ pW assuming $X_{move} = 2$ cm, $t_{move} = 3$ sec. In normal walking with $X_{move} = 0.6$ m and $t_{move} = 1$ sec, $P_n = 0.06$ pW; for the most violent handwaving exercise with $X_{move} = 0.3$ m and $t_{move} = 0.2$ sec (5 Hz), $P_n = 2$ pW. All of these figures optimistically assume 100% mechanical efficiency.

Nanorobots resident in the chest walls experience cyclical displacements of $X_{move} = 2.5$ cm and $t_{move} = 3.3$ sec (0.7 Hz) for normal shallow breathing at 10/min, up to $X_{move} = 5$ cm and $t_{move} = 1$ sec (1 Hz) for deep breathing during heavy exertion (Section 8.2.2), a chest velocity range of $v_{move} = 2$-10 cm/sec, giving $P_n = 3-400 \times 10^6$ pW (Eqn. 6.12) for a pendulum transducer. Another alternative in the chest is a simple spring-loaded stretch transducer. If the diaphragm and chest wall muscles cost $P_{chest} = 1-25$ watts to operate depending upon exertion level, then the applied force $F_{chest} / t_{move} = 50-250$ N; an $X = 5$ cm displacement for a 90-cm circumference chest cavity gives a $\delta = 5\%$ distension. Thus a two-phase L = 100 nm stretch transducer produces $P_n = 2 \delta L e\% F_{chest} / t_{move} = 150-2500$ pW, conservatively taking mechanical efficiency $e\% = 0.001$ (0.1%) to account for poor coupling and a highly nonspecific tissue stress tensor. Similar power levels are available in contractile cardiac tissues, near joints that flex during normal arm or leg motions, and in the pedal dermis during walking.

Another example of mechanomechanical transduction is a simple gear train, which transmits mechanical rotational power from one location to another. In one example given by Drexler,\textsuperscript{10} a 17 nm$^3$ steric gear pair transmits 1 nanowatt of mechanical power, giving a power density of $6 \times 10^{10}$ watts/m$^3$ with a mechanical efficiency of 99.997%. Complex hectomicron-scale gear trains have been fabricated.\textsuperscript{558} Sandia’s Microelectronics Development Laboratory mass-produces 100-micron motors and gears.\textsuperscript{1259} Properly designed molecular bearings will have lifetimes that are not limited by wear but only by the static lifetime of the bearing (e.g., due to radiation damage).

In 1998, there was at least one unconfirmed experimental report of direct mecanochemical energy transduction,\textsuperscript{2924} in which the mechanical energy of stirring of water over a catalyst bed of powdered cuprous oxide was alleged to produce H$_2$ and O$_2$ gas effluent.

Mechanoacoustic transduction may be achieved by operating the pressure-driven actuators described in Section 6.3.3 in reverse, using an oscillating mechanical energy input such as a reciprocating rod to drive the piston. See Section 7.2.2.1 for a more complete treatment of the vibrating piston acoustic radiator.

Mechanoelectric transduction is commonplace in atomic force microscope (AFM) sensors using piezoresistive cantilevers that produce a varying electrical potential in response to changing mechanical loads. Typical coupling constants (e.g., mechanoelectric efficiency) are 11% for polyvinylidene fluoride (PVDF), and 35-59% for lead zirconate titanate (PZT).\textsuperscript{3995} Transduction of mechanical to electrical signals can also be achieved by mechanically modulating a tunneling contact junction as a mechanically controlled electrical switch, as occurs at the tip of a scanning tunneling microscope (STM) in response to varying physical loads. Fullerene molecules also change their electrical resistance in response to mechanically-applied loads, providing mecanoelectrical switching (Section 10.2.2). Alternatively, operating the nanoscale electrostatic DC motor described in Section 6.3.5 in reverse converts it into a mecanoelectric generator, transforming rotational mechanical energy into electrical current at high efficiency. Cochlear outer hair cells are capable of both mecanoelectrical and electromechanical transduction.\textsuperscript{3997}

Cardiac mecanoelectric transducers have also been proposed or tested.\textsuperscript{593,722,723,3513-3517} In 1966, Ko\textsuperscript{533} fabricated a mechanical converter that converted heart movement into the vibration of a...
piezoelectric rod to generate AC electric power that was rectified and used to power a pacemaker; such devices were installed on dogs' hearts and generated 30 microwatts for several months. Electropaced skeletal muscles (latissimus dorsi) have been used to compress an implanted plastic balloon counterpulsation device to produce pulmonary artery diastolic pressure augmentation. There are proposals to use piezoelectric bimorphs to transform the mechanical energy of diaphragm motions during breathing, and of the circumferential movements of the aorta and other elastic arteries during each pulse cycle into electrical energy, but transducer fatigue and biocompatibility issues were problematical. Piezoelectric fields developed in moving bones and collagenous tissues possibly could be tapped. Indeed, many biological materials have been found to be piezoelectric, including tendon, elastin, silk, dentin, ivory, wood, aorta, trachea, intestine, and even nucleic acids. A gait-powered battery-charging system developed by Freeplay Power Group in 1999. Keyboard typing and used to power a pacemaker; such devices were installed on dogs’ hearts and generated 30 microwatts for several months. Electropaced skeletal muscles (latissimus dorsi) have been used to compress an implanted plastic balloon counterpulsation device to produce pulmonary artery diastolic pressure augmentation. There are proposals to use piezoelectric bimorphs to transform the mechanical energy of diaphragm motions during breathing, and of the circumferential movements of the aorta and other elastic arteries during each pulse cycle into electrical energy, but transducer fatigue and biocompatibility issues were problematical. Piezoelectric fields developed in moving bones and collagenous tissues possibly could be tapped. Indeed, many biological materials have been found to be piezoelectric, including tendon, elastin, silk, dentin, ivory, wood, aorta, trachea, intestine, and even nucleic acids. A gait-powered battery-charging system developed by Freeplay Power Group in 1999. Keyboard typing produces 0.1-30 milliwatts/finger (Section 7.4.2.1).

Piezoluminescent crystals provide mechanooptical power transduction, and micromachined optical shutters have demonstrated mechanical switching of optical signals.

6.3.3 Acoustic Energy Conversion Processes

Most clearly useful for medical nanorobots are acoustomechanical transducers (Section 4.5.1) that can directly apply motive power to internal nanomechanical manipulator and computational systems. Drexler has described a simple transducer that can function as a pressure-driven actuator, which is abstractly modeled as a constant-force spring that resists the motion of a piston moving between two limit stops (Fig. 6.1). Such a device which experiences a cyclical volumetric change of \( \Delta V \) in response to above-ambient pressure pulses of amplitude \( \Delta P \) at a frequency \( v_p \), converts acoustic pulse energy into the mechanical energy of piston motion with almost 100% thermodynamic efficiency when the expansion is isothermal and reversible (e.g., when thermal equilibration time \( t_{EQ} \) << \( v_p^{-1} \), which should always hold for \( v_p \) ≈ 1 GHz).

Recognition of four additional restrictions on the net mechanical power \( (P_n) \) available from a piston-type transducer provides a very conservative power estimate:

\[
P_n = v_p (\Delta P \Delta V - E_{friction} - E_{inertia} - E_{drag} - E_{heat}) \quad [Eqn. 6.13]
\]

The product of \( \Delta P \), the change in pressure during a stroke, and \( \Delta V \), the change in piston volume, is the change in Gibbs free energy per cycle. The piston energy lost to friction may be estimated from Eqn. 6.4 using the observation that interface velocity \( v = 2 v_p L \) for a piston of dimension -L³, thus \( E_{friction} \approx 10^{-25} v_p^3 L^6 \); for \( E_{friction} \leq kT \), L < 0.05 micron for \( v_p = 1 \) GHz, L < 5000 microns for \( v_p = 1 \) MHz. The energy lost in overcoming piston inertia - 1/2 \( m \nu^2 \) where \( m = \rho_p L^3 \), so \( E_{inertia} \approx 2 \rho_p v_p^2 L^5 \) for a piston of density \( \rho_p \); for \( E_{inertia} \leq kT \), L < 4 nm for \( v_p = 1 \) GHz, L < 60 nm for \( v_p = 1 \) MHz. Fluid viscous drag occurs when the piston must push itself through a viscous external aqueous medium in which the pressure pulses are traveling. However, if \( L \ll \lambda = \nu \lambda_{sound}/v_p \) then the range of piston motion is far smaller than one wavelength and \( E_{drag} = 0 \) for submicron transducers: at \( v_p = 1 \) GHz, \( \lambda = 1.5 \) microns in water at 310 K. Finally, \( E_{heat} \approx kT \) because useful work may not be extracted from an isothermal medium. \( L_{min} \) is the size of the receiver piston for which \( P_n \ll 0 \).

Of course, this is a very conservative approach because much depends upon how predictable the oscillatory source is. J. Soreff notes that in the case of externally supplied, constant frequency, constant amplitude pressure waves, both \( E_{inertia} \) and \( E_{heat} \) are negligible because the capture of energy from the piston can be phase locked to the incoming waves, and the piston spring constant can be tailored to put the kinetic energy of the piston into a potential energy store at the end of its travel. Phase locking still requires an energy expenditure of \( kT \) for sensing, but this sensing could be spread over many cycles, given a source with a long coherence time.

Table 6.3, with values computed using Eqn. 6.13 and acoustic pressure estimates from Section 4.9.1, shows that a modest amount of acoustic energy is available from normal physiological processes within the human body, which are assumed to generate planar traveling waves. Path length is not important here because of the weak attenuation at such low frequencies (Section 4.9.1.3). For a single (-1 micron)³ transducer, usable received acoustic power up to \( P_n \approx 0.004 \) pW is regularly available from continuous sources. Intermittent or sporadic usable power sources up to \( P_n \approx 5 \) pW are also available. In many cases these sources would be in phase with the biothermal transducer described in Section 6.3.1. Large epidermally-placed nanorobots may convert the force of wind resistance into usable energy. In theory, nanorobots may receive \( P_n = (1/2) \epsilon \rho_{air} A v_{run}^3 \approx 1 \) pW for efficiency \( \epsilon = 0.30\% \), STP air density \( \rho_{air} = 1.29 \) kg/m³, and transducer area \( A = 1 \) micron² for a person walking or running at \( v_{run} = 1.8 \) m/sec (-4 mph). However, as a practical matter much less power is available to micron-size nanorobots because there is a near-skin boundary layer of relatively still air.

External sources of acoustic energy can also be employed (Section 6.4.1). Power pulse cycling depends on sensor design. For example, to avoid damaging the aforementioned acoustic sensor devices with large acoustic pulses intended to provide power, power pulses should be shaped with a slow initial amplitude rise sufficient to gently “peg” all sensor pistons to their distal stops, followed by a sharp upramp to peak amplitude, then a shallow decline sufficient to gently return sensor pistons to their proximal stops, followed by a sharp decline.
in pulse amplitude, thus safely completing the power cycle. An acoustic-powered legged microrobot ~1 mm in size with resonant actuators has been fabricated and tested.3525

The Space Thermoacoustic Refrigerator (STAR) developed by Steven L. Garrett for NASA in the early 1990s uses 160 dB sound waves in a confined space to create a spatial thermal gradient in a working fluid, an example of acoustothermal energy conversion.3446 Sonoluminescence caused by focused high-frequency sound waves in water provides an example of acoustophotonic power transduction.543716 Acoustochemical processes are well-known in the field of sonochemistry,625,1084,1085,1523 including ultrasonic depolymerization. Acoustomechanical fluid-driven microturbines ~250 microns in diameter have been fabricated and operated at ~1 KHz frequency with lifetimes of ~10⁸ revolutions.1218

### 6.3.4 Chemical Energy Conversion Processes

The following Section describes possible energy sources for chemically-powered foraging nanorobots. Chemically-powered tethered or stored-energy nanorobots may have different constraints (see Sections 6.4.3.5, 6.5.3, and 6.5.4).

#### 6.3.4.1 Human Chemical Energy Resources

The first step in evaluating chemical power transduction alternatives for medical nanorobots is to assess the chemical energy resources readily available within the human body. A complete inventory is beyond the scope of this book, but a few broad conclusions may be drawn.

Table 6.4 summarizes major representative chemical energy resources available within typical individual human tissue cells, within the human blood volume, and within an entire 70 kg adult human body. Burning a protein or carbohydrate in oxygen releases ~4.1 Kcal/gm (17 x 10⁶ joules/kg); burning lipids in oxygen releases ~9.3 Kcal/gm (39 x 10⁶ joules/kg). In theory, proteins represent the most plentiful energy resource in blood and cells, but extensively accessing this source may require disassembly of essential cellular structures, considerable preprocessing (e.g., denaturation and lysis), and disposal of nitrogen-containing waste products.* Free amino acids and short peptides are too dilute to be of much value. ATP, usually available only inside cells with a high turnover rate, is also of very limited utility (though it is fairly stable—one can buy bottles of powdered ATP and store it for years in the freezer). Fats potentially offer the most plentiful natural energy source in the human body, providing over 80% of body heat in the absence of carbohydrates. Unfortunately, high energy density serum lipids such as cholesterol are not freely available but instead are present only in three bound forms that ensure solubility:

1. chylomicrons and other plasma lipoprotein carriers (protein-coated lipid droplets <0.5 microns in diameter), to transport lipids throughout the body from the intestine after absorption of dietary fat, or from the liver after lipid synthesis to storage in adipose tissue or for utilization elsewhere in the body;3525
2. fatty acids bound to serum albumin, to transport fat from adipose tissue to elsewhere in the body; and
3. ketone bodies (acetoacetate and beta-hydroxybutyrate) to transport lipids processed by or synthesized in the liver.

Anhydrous lipid droplets 0.2-5 microns in diameter are often found in the cytosol. Excepting operations inside adipocytes (wherein triglyceride droplets occupy nearly the entire cytosol), fat utilization by nanorobots would probably require some physical nanopipetting or other preprocessing to dislodge the lipid from its carrier, and different energy extraction pathways may be required depending on whether the fatty molecules thus liberated are saturated or unsaturated, etc.

Carbohydrates are soluble in water and thus exist in free form throughout the body. Large individual molecules of glycogen (10-40 nm in diameter) float freely as granules in the cytosol, usually coated with a ~5 nm thick monolayer of digestive enzymes (Section

---

*The human body secretes urea because it lacks the necessary enzymes to further process the nitrogen. Nanorobots could make use of bacterial enzymatic or other chemical processes to obtain energy from nitrogenous wastes—for example, by exploiting the ammonia oxidation reaction: 4NH₃ (g) + 3O₂ (g) → 2N₂ (g) + 6H₂O (g) + 2105 zJ.
Table 6.4  Estimated Chemical Energy Resources in the Human Body

<table>
<thead>
<tr>
<th>Chemical Fuel</th>
<th>Energy Available in 70 kg Human Body (joules)</th>
<th>Energy Available in 5.4 Liters of Whole Blood (joules)</th>
<th>Energy Available in Typical 20-µm Cell (joules)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP</td>
<td>2.2-4.1 x 10^6</td>
<td>1.7-3.1 x 10^4</td>
<td>0.6-1.0 x 10^9</td>
</tr>
<tr>
<td>Glycine</td>
<td>3.5-4.8 x 10^6</td>
<td>2.8-6.0 x 10^4</td>
<td>0.9-1.2 x 10^9</td>
</tr>
<tr>
<td>Leucine</td>
<td>1.2-1.7 x 10^5</td>
<td>2.0-2.9 x 10^4</td>
<td>3.0-4.3 x 10^4</td>
</tr>
<tr>
<td>Lactate</td>
<td>0.6-2.6 x 10^5</td>
<td>0.4-1.6 x 10^4</td>
<td>0.5-2.4 x 10^4</td>
</tr>
<tr>
<td>Glucose</td>
<td>0.5-1.1 x 10^6</td>
<td>3.6-7.7 x 10^4</td>
<td>0.5-1.1 x 10^7</td>
</tr>
<tr>
<td>Urea</td>
<td>0.7-1.4 x 10^6</td>
<td>1.2-2.4 x 10^4</td>
<td>1.8-3.5 x 10^4</td>
</tr>
<tr>
<td>Glucose, free</td>
<td>3.3-4.9 x 10^6</td>
<td>5.6-8.2 x 10^4</td>
<td>0.8-1.2 x 10^7</td>
</tr>
<tr>
<td>Glycogen (starch), granules</td>
<td>4.2-6.3 x 10^6</td>
<td>0.1-1.6 x 10^4</td>
<td>1.3-8.0 x 10^7</td>
</tr>
<tr>
<td>Cholesterol (lipid)</td>
<td>1.4-2.8 x 10^7</td>
<td>2.4-4.7 x 10^7</td>
<td>3.6-7.0 x 10^7</td>
</tr>
<tr>
<td>Total carbohydrate</td>
<td>3.3-5.4 x 10^7</td>
<td>0.6-1.1 x 10^4</td>
<td>0.8-9.0 x 10^-6</td>
</tr>
<tr>
<td>Total protein</td>
<td>1.3-2.5 x 10^8</td>
<td>1.9-2.0 x 10^7</td>
<td>1.4-2.9 x 10^-5</td>
</tr>
<tr>
<td>Total fat (lipid)</td>
<td>1.6-9.8 x 10^8</td>
<td>0.9-1.3 x 10^7</td>
<td>6.2-9.4 x 10^-6</td>
</tr>
</tbody>
</table>

8.5.3.7). Glucose is the preferred human body fuel—a nearly constant bloodstream inventory of ~5 grams is maintained homeostatically from a buffer supply of ~350 grams of glycogen stored mainly in the liver and muscles. Glucose transporters maintain cytosolic concentrations at comparable levels to the blood. Nerve cells can only metabolize glucose and ketones; sperm cells are surrounded by a fructose bath, from which they absorb chemical energy to power flagellar engines during their short-lived locomotion (Section 6.3.4.2).

Because of its abundance, high energy density, and relative ease of use (e.g., high solubility), glucose is probably the ideal fuel for most in vivo nanomedical applications. A micron³ onboard store of glucose fuel powers a 10 pW nanorobot for 2400 sec of operation without further refueling. The equilibrium free glucose mass in a cytosolic space has sufficient fluid volume to supply adjacent cells at energy conversion efficiency (see below), and that the local extracellular space has sufficient fluid volume to supply adjacent cells at peak rates, then the maximum theoretical total continuous power draw in cyto is 70,000-300,000 pW for foraging glucose-consuming nanodevices with unlimited access to oxygen (e.g., onboard storage). If oxygen supplies were restricted to that which could passively diffuse through cell walls, the comparable in cyto diffusion-limited total oxygen current would be I_{O_2} = 2 x 10^{13} O_2 molecules/sec, a ~4000 pW power limit. (For comparison, the basal power consumption of a typical 20-micron human cell is ~30 pW; Table 6.8.)

The following Sections include brief technical sketches of several possible approaches to chemooxygenic transduction. The methods outlined are rather crude. An ideal design might involve complex nanomachinery that has yet to be designed, which may perhaps more closely resemble in efficiency and function the enzymatic and “proton pump” biological nanomachines that are found in nature, or the natural chemophotonic transducers employed in bioluminescence.3564

6.3.4.2 Biological Chemomechanical Power Conversion

Many examples of direct chemomechanical power transduction are found in nature. Perhaps the most familiar is the mechanism of muscle contraction mediated by the sliding of interdigitating myosin and actin filaments—the actomyosin molecular motor.1246 Myosin is constructed of domains joined by hinges and is powered by ATP. The globular head domain of the myosin motor, called myosin subfragment-1 or “S1”, is a -10 nm wide at the head and ~20 nm long, containing ~50,000 atoms (~500 nm³). When the ATP molecule attaches to an open ATPase pocket on S1, this action causes a conformational change in the S1 protein, which releases its hold on the associated actin filament. The pocket can then close further, clamping the terminal phosphate from the ATP (making ADP) and powering yet another conformational change which increases affinity for actin. The phosphate is released as S1 again binds to actin, pulling the pocket open wider and triggering a ~2 millisecond power stroke as the spent ADP molecule is ejected.576 During each 0.05-sec cycle, the myosin motor advances 10 nm along the actin filament (~5 micron/sec velocity) pulling with a force of ~5 pN, a mechanical energy output of ~50 zJ, or ~10^{-6} pW (power density ~2 x 10^6 watts/m³). The energy released in converting one ATP molecule to ADP is ~83 zJ, so the motor is ~60% efficient. A typical macroscale muscle may contain ~10^{11} individual actomyosin motors. The net energy efficiency of a muscle is at most 25%, doing the greatest work when it shortens by only ~10% of its length.2035 In 1998, the myosin superfamly of protein motors contained 15 different classes, and ~30 different myosin species were known.2279

Other molecular motors similarly transduce chemical energy into mechanical motion to drive the beating of cilia (dynein; Section 9.3.1.1) and to drive the transport of vesicles and organelles along tracks within cells (kinesin; Section 9.4.6). In each case, the binding of ATP (or GTP) induces conformational transitions in the ~12 nm
motor proteins, which are reversed by the hydrolysis of bound ATP and release of ADP and phosphate, advancing the motor protein along the tubulin track and pulling with a force of ~2 pN. Another example is a component of the mitochondrial enzyme ATP synthase: F1-ATPase, a ~1 nm rotor spinning in a ~12 nm barrel, the smallest rotary motor known, producing forces >100 pN and a rotary torque of ~40-80 pN-nm under high load. A different type of chemomechanical motor protein—an enzyme called lambda exonuclease found in a bacteriophage virus—chews up one of DNA’s double helix strands, leaving behind a single strand; the motor pulls >5 pN while moving along the strand, and is powered directly by the energy liberated from DNA’s broken bonds.

In human mitochondria, the combined TCA (tricarboxylic acid) cycle (aka. Krebs cycle) and cytochrome chain oxidative-phosphorylation convert 39 molecules of ADP to higher-energy ATP during the net metabolism of one glucose molecule, a chemomechanical energy conversion efficiency of ~65%. Combining this metabolic process with the myosin motor described earlier would yield a glucose-powered chemomechanical transducer that is ~39% efficient.

While all known eukaryotic motors are powered by ATP or GTP, neither is required for bacterial flagellar rotation. Rather, the flagellar ionic motor is a rotary device, driven by the pH differential between the cytosolic and extracellular sides of the device. This ~5 x 10^6 atom, 50,000 nm^3 motor consists of a shaft attached to a rotor that spins up to 300 Hz (15 Hz at high load) inside a fixed stator ~30 nm in diameter attached to the peptidoglycan layer, driven by the flow of ~1000 protons (H+ ions) per revolution. Stator and rotor surfaces lie ~0.5 nm apart.

Figure 6.2 illustrates how this proton-gradient driven machine works. The rotor has a set of ~100 proton-accepting sites around its periphery which interact with 8 channel complexes embedded in the stator. Each channel complex has two half-channels, one accessible from the (relatively alkaline) cytosol, the other from the (relatively acidic) extracellular space. A proton is transferred from a half-channel to an acceptor site on the rotor, but cannot be donated to the other half-channel unless the rotor rotates. Referring to the Figure to trace one cycle, in the geometry shown in (A), site 2 is empty and site 3 is filled, so the complex can move right but not left. If random motion carries the complex one step to the right (B), the proton on site 3 can move into the cytosol and a new proton can move into the adjacent empty site 4 (C). The rotor then rotates counterclockwise a distance equal to the spacing between acceptor sites, to relieve the spring tension (D). Reversing the geometry (which takes ~1 millisecond, controlled by a protein switch) allows the motor to run backwards; this occurs naturally every few seconds in the absence of any environmental stimulus. The motor develops ~10^-4 pW with a power density of ~2 x 10^6 watts/m^3. Motor efficiency is <5% at low load, 50%-99%+ at high load.

Flagellar ionic motors that run on sodium ions rather than protons have been found in some bacteria; no doubt other alternatives could be employed in chemomechanically powered medical nanorobots. Drexler estimates from first principles that chemically-driven engines can be designed to operate at >99% efficiency at a power density of ~10^9 watts/m^3.

Other natural chemomechanical engines involve fluid movements, including evaporative-driven ascent of sap in plants, hydrophilic- and hygroscopic-driven expansion due to water absorption, and osmotic pumps in cells, leaves, and plant roots.

6.3.4.3 Artificial Chemomechanical Power Conversion

Chemomechanical gel actuators have been reviewed by Tatara, who describes a variety of polyelectrolyte gels and ion-exchange resins used to make centimeter-scale “mechanochemical pistons” and 100-micron polyelectrolyte fibers used to make chemomechanical “finger” actuators; unfortunately, these electrically-stimulated transducers may fatigue after only ~500 cycles. More interesting is the Sussmann-Katchalsky chemomechanical turbine (Fig. 6.3)—a continuous collagen fiber contracts as it enters a 10M LiBr solution, producing a contractile force that is converted into a torque to rotate the shaft of the output pulley. The device has been patented,
built, and operated with an energy conversion efficiency of ~40%. While it would be difficult to scale down this turbine to submicron dimensions, it nonetheless constitutes another useful proof of principle that direct chemomechanical power transduction may need only surprisingly simple mechanisms.

Astumian has examined the use of nonequilibrium chemical drivers for micron and submicron Brownian motors. For instance, when ATP (negatively charged in solution) binds to a protein, it changes the protein’s net charge. Thus two proteins that react with ATP at different rates would feel different electrostatic forces; this force differential can drive a motor. Astumian estimates that even a crudely designed chemically-driven Brownian motor could move in 10-nm steps at ~3 microns/sec, developing ~0.5 pN of force, ~3 x 10^6 pW of power, and a power density of ~10^6 watts/m^3. Many types of cellular protrusions such as filopodia, lamellipodia, and acrosomal extension do not involve molecular motors but rather the transduction of chemical bond energy into mechanical energy using Brownian ratchets to rectify random thermal motions by expending chemical energy in actin or tubulin polymerization.

6.3.4.4 Glucose Engines

There are many pathways for glucose oxidation in the human body, some which directly involve molecular oxygen. For instance, the glucose-6-phosphate dehydrogenase and the 6-phosphogluconic dehydrogenase enzymes both employ molecular oxygen directly to oxidize a glucose derivative. Although evolution has largely replaced these primitive pathways and a few can even be found in mammalian tissues. For example, enzymes in the liver use the reactions:

\[
\text{Gluconic acid} + H_2O_2 \rightarrow \text{Glucose oxidase} + 131 \text{ zJ} \quad \text{[Eqn. 6.14]}
\]

with

\[
H_2O_2 \rightarrow \text{catalase} \rightarrow H_2O + \frac{1}{2}O_2 + 166 \text{ zJ} \quad \text{[Eqn. 6.15]}
\]

One very crude approach for medical nanorobots is to use mecha- nochemistry techniques to force an exothermic chemical reaction, making a change in volume or generating heat which is subsequently exploited to produce mechanical motion. Given that glucose may be the preferred chemical fuel for nanomedical systems (Section 6.3.4.1), we concentrate our attention on the high-energy exothermic oxyglucose combustion reaction

\[
C_{6}H_{12}O_{6} + 6O_2(g) \rightarrow 6CO_2(g) + 6H_2O(g) + F_{\text{gly}} (= 4,765 \text{ zJ}) \quad \text{[Eqn. 6.16]}
\]

The kinetic model for the glucose oxidation reaction has not been extensively studied at pressures >30 atm, so it has not yet been proven that an oxyglucose mixture can be ignited at high pressure near human body temperature at 310 K (37°C). Such a model could have a complex structure, such as the well-studied stoichiometric H_2/O_2 mixture which shows a minimum ignition temperature of only ~380 K (107°C) at ~2500 atm.

Pure sugar in air is not ignited by open flame (e.g., bunsen burner, matchstick), but readily burns if a catalyst such as ash is added. Exposed to moist air below 323 K (50°C), each molecule of anhydrous glucose hygroscopically absorbs a single molecule of noncovalently-bound water, yielding the monohydrate. Heated above 323 K, anhydrous glucose “caramelize”s when it reaches ~433 K (160°C). Caramelization is not oxidation, but rather is an endothermic decomposition process involving successive dehydration, condensation, and polymerization reactions, that includes the making and breaking of covalent bonds, resulting in brown melamines and possibly some carbonization. At 433 K the total translational kinetic energy of the 7 oxyglucose reactant molecules would be 7(3/2)(kT) ~ 63 zJ.

What is the minimum ignition pressure (P_{ignition}) of glucose? Pressures of 5,000-15,000 atm deactivate antibodies, enzymes, and proteins (Table 10.3), and ignition temperatures normally fall with rising pressure in combustion reactions. In food science, the well-known Maillard reaction (“browning”) between amino acids from protein (mainly lysine) and reducing sugars (glucose and lactose) involves a potpourri of Amadori rearrangements, Schiff’s base formations, and Strecker degradations, and occurs during heating (cooking) and pressurization (e.g., the glucose-lysine system at 6000 atm). The rate increasing 2-3 times with each 10 K temperature rise. Pressure-induced peroxidation of lipids has been shown with browning greatly accelerated by pressurized oxygen. It is also well-known that the hydrocarbons found in oils and greases, and many other organic compounds, will “ignite almost spontaneously” in oxygen stored in compressed tanks at ~150 atm. Upon contact with liquid oxygen (LOX), paper, textiles, asphalt, tar, kerosene, wood, stainless steel, teflon, and silicones are combustion hazards, and pulverized organic materials such as sawdust, polystyrene and charcoal, and powdered magnesium, can spontaneously ignite or explode when saturated with LOX these materials have exploded after an impact as slight as a footstep. The density of oxygen molecules trapped in a mechanochemical binding site at 1000 atm is 12.6 molecules/nm^3, not much below the ~21.5 molecules/nm^3 density in LOX (Table 10.2).

Lacking firm experimental data, we provisionally approximate P_{ignition} as the minimum ~63 zJ activation energy (required to induce covalent bond modification during caramelization), divided by the total reactant volume at 310 K computed by summing the volumes of one glucose (1.91 x 10^{-28} m^3/molecule) plus 6 molecular oxygens (8.85 x 10^{-29} m^3/molecule at 860 atm using van der Waals’ equation; Section 10.3.2). The required energy density is 8.7 x 10^7 joules/m^3, which can be supplied to the reactants by applying a
mechanical pressure of $P_{\text{ignition}} \approx 860$ atm. Operating the glucose engine at $P_{\text{ignition}} \approx 1000$ atm, possibly including durable catalysts to accelerate the reaction, thus would ensure complete and reliable combustion. (Note that biology already achieves full stepwise glucose oxidation at 310 K, constituting a general proof of principle.) J. Soreff suggests several interesting design alternatives:

1. Convert glucose to a more stable fuel such as methane (which will not caramelize), possibly by biochemical means, incurring some loss in efficiency;

2. precompress the fuel/oxidizer mixture in a single tank to subignition pressures, then trigger the reaction purely thermally or possibly by adding free radical initiators such as benzoyl peroxide to the tank; or

3. nitrate the glucose hydroxyls, producing combustion and evolution of N$_2$ which must then be recycled back to HNO$_3$ via nitrogen fixation (Chapter 19) to nitrate the next batch of fuel.

Once oxyglucose combustion has taken place, the energy thus released may be converted to mechanical energy by various means. In the ideal case of an isothermal expansion at constant (P = 1000 atm) pressure, $P \Delta V = E_{\text{ghu}}$ would imply a volume change $\Delta V \approx 47$ nm$^3$, producing a ~5 nm power stroke on a piston of area ~10 nm$^2$ and a ~1 nN applied force. Vacuum isolation of the combustion chamber assembly (see below) helps prevent rapid energy dissipation via thermal conduction to the external medium. Burning 10$^9$ sec$^{-1}$ glucose molecules in a ~500,000 nm$^3$ combustion/engine device (which includes a generous volumetric allowance for support structure) generates ~5 pW with a theoretical power density of ~10$^{10}$ watts/m$^3$. Note that while the oxyglucose reaction products occupy 0.047 nm$^3$ more volume than the reactants, relying upon this volume change only allows a highly inefficient (~0.1%) extraction of ~5 zJ per glucose molecule at 1000 atm operating pressure.

A complete design for a working nanoscale glucose engine is beyond the scope of this book. However, as a concept demonstration without regard to maximizing efficiency, one simplistic implementation might employ a combustion chamber at the center of a vacuum-isolated “thermos bottle.” A series of combustion events maintains the isolated heat source at a high operating temperature, thus providing a sizable temperature differential against the external medium. This differential can subsequently be used to drive a microscale Stirling engine to produce a mechanical power output,541 which requires at least five critical design elements:

A. Combustion Mechanism—Figure 6.4A shows a heat-generating mechanism which employs two opposed sorting rotors. These rotors contain binding sites which must be exceptionally robust and damage-resistant. Each rotor is a cylinder 40 nm tall with ten staggered rings of opposed binding sites (Fig. 6.4B). The left rotor has binding sites which accept one glucose molecule at a time from the aqueous fuel solution, simultaneously applying sufficient force to strip off the associated water of hydration. The right rotor has binding sites each of which accepts oxygen molecules, six at a time. During operation, these “crushing rotors” compress paired reactant packets into the activation volume required for the oxyglucose combustion reaction to occur. This reaction volume may change shape during the compression cycle, perhaps employing a judicious series of insertions and retractions of binding pocket rods in order to accommodate intermediate reaction products and multistep reaction events. The end result is to force a complete combustion at the rate of 3 x 10$^6$ glucose molecules per second (~7 pW). Effluent molecules are ejected into the outer jacket, transferring their heat to the thermally-conductive structure in <10$^{-9}$ sec (Section 4.6.1). The glucose solvation water discharge pathway is not shown in the Figure.

B. Fuel Tanks—The combustion mechanism has fuel tanks sized to stoichiometric requirements from which combustion reactants are drawn during each 100 millisec discharge cycle. Tanks are refilled once during each discharge cycle with a saturated 70% glucose (aqueous) solution and O$_2$ at >1000 atm pressure. During refueling, the thermos structure equilibrates to 310 K in ~10$^{-9}$ sec by thermal conduction through the refilling mechanism, which physically plugs into the refueling sockets. Refilling requires <10$^{-7}$ sec. The refilling mechanism then disconnects and withdraws at least 100 nm, after which the combustion mechanism raises thermos temperature back to the 600 K operating temperature in ~30 millisec. Thermos temperature then remains at 600 K throughout the remainder of the 100 millisec discharge cycle. Presumably this cycle is fast enough to avoid glucose fuel caramelization at 433 K inside the tank; if not, a faster cycle or a lower operating temperature may be used, forcing a somewhat lower efficiency on the entire design. (The boiling point of the water solvent exceeds 433 K at >6.4 atm (Section 10.3.2); at 600 K, well below water’s critical temperature $T_{\text{crit}} = 647.3$ K (Section 10.3.2), water remains liquid at pressures $\geq 140$ atm.)

C. Rotor Drive Source—Stacked above the 40-nm tall fuel/rotor complex is a second O$_2$ storage tank (~15 nm tall) at >2000 atm pressure (Figure 6.4B). This gas passes through a simple ducted turbine mechanism which is coaxial with the two rotors, driving rotor rotation. Spent drive gas passes into the main O$_2$ storage tank and
Basic Capabilities • Power

is later used as combustion reactant. Drive gas is recharged, and effluent gases are drained off, each process taking ~10^-7 sec, once every 100 millisec discharge cycle.

D. Electrodynamic Suspension—It is crucial to avoid any thermally conductive physical pathways until the cycling Stirling engine (or other intentional energy transfer mechanism) calls for heat. Even the best thermal insulators such as wood, plastic and glass have conductivities $K_t = 0.01-1$ watts/m-K (Table 8.12, Appendix A), so a 10-nm cube of such material straddling a $\Delta T = 290$ K temperature differential conducts away the entire energy store of $E_{\text{store}} \sim 1$ pJ in ~0.1-10 microsec. Using $N_{\text{rod}} = 6$ carbyne rod supports, each of cross-sectional area $A_{\text{rod}} \sim (0.2 \text{ nm})^2$, length $L_{\text{rod}} = 100$ nm, and $K_t = 2000$ W/m-K for diamond, the thermal leakaway time is:

$$t_{\text{E leak}} = \frac{E_{\text{store}} L_{\text{rod}}}{N_{\text{rod}} A_{\text{rod}} K_t \Delta T} = 0.7 \text{ microsec} \quad \{\text{Eqn. 6.17}\}$$

still much too fast. (J. Soreff notes that glasslike disorder in molecular chains may be obtained by placing $^{12}$C and $^{13}$C atoms randomly throughout the carbyne structure, rendering local most of the high spatial and temporal frequency modes and possibly significantly reducing thermal conduction. For somewhat weaker strands, N/CH or O/NH/CH$_2$ units might be intermixed with similar effect.) Fortunately, vacuum non-contact isolation of the entire combustion assembly also is possible via electric levitation. Earnshaw’s theorem$^{2172}$ states that static levitation against gravity is not possible in classical physics using any combination of fixed magnets or electric charges. However, stable levitation may be achieved by using quantum effects (Section 9.2), diamagnetic suspension materials,$^{2173,2174}$ rotation of the suspended object as in the levitron,$^{2175}$ oscillating magnetic fields,$^{2176}$ or dynamic feedback control mechanisms. Three-dimensional electro-dynamic confinement of charged micron-sized aluminum particles was demonstrated experimentally in the late 1950s.$^{654}$ From the formulas given, a 100-nm particle with 10 surface charges (charge/mass ~ 1) can be stably confined to a 200-nm volume in a circular Lissajous pattern using a ~1 volt hyperbolic quadrupole AC field at a ~10 MHz resonance frequency plus a ~10 mV DC field to neutralize gravity. A stiffer confinement field on micron-scale particles has been produced inside a spherical void electrodynamic levitator constructed as two joined hemispheres,$^{657}$ and more complex configurations have been designed.$^{658,659}$ Electric suspension bearings for micromotors using a resonant circuit driven by radio frequency AC voltage can achieve stable levitation of charged objects without requiring feedback signals or sensors,$^{656,657}$ and stable passive trapping may also be achieved using cylindrically symmetric dielectrophoretic levitation electrode structures.$^{660}$ Translational and rotational motions may be imposed on the suspended object by applying appropriate DC or oscillating fields.

Levitation is particularly effective at the smallest scales. Electrically levitating an object of thickness $t$, density $\rho$, and dielectric strength $\kappa \varepsilon_0$, against a gravitational acceleration $g$ requires an electric field$^{656}$

$$E_{\text{lev}} = \left( \frac{8 \rho g t}{\kappa \varepsilon_0} \right)^{1/2} \quad \{\text{Eqn. 6.18}\}$$

If $r = 1000$ kg/m$^3$, $g = 9.81$ m/sec$^2$, $t = 100$ nm, $\kappa = 2.1$ (e.g., Teflon), $\varepsilon_0 = 8.85 \times 10^{-12}$ farad/m, then $E_{\text{lev}} = 20,000$ volts/m or 20 mV across a 1 micron gap (~typical in biological cells). Two unit charges separated by 100 nm in vacuo feel a mutual repulsion of ~0.1 pN, sufficient to resist an external acceleration of ~$10^4$ g’s on a (100 nm)$^3$ block of density $\rho = 1000$ kg/m$^3$; $10^4$ g’s is the largest natural acceleration imposed on objects of this size inside the human body (Section 4.3.3.2) and is near the maximum safe acceleration for human cells (Chapter 11).

One possible stable open-loop levitation configuration is shown schematically in Figure 6.4C, wherein the combustion assembly is affixed to a simple permanently-polarized dipole electret carrier which is itself electrodynamically levitated by same-sign control electrets positioned at either end. Thermal electrets (electric field applied across heated material$^{2177,2178}$) may be made of wax, Teflon, or Mylar; several polymer electrets have extrapolated lifetimes of several thousand years at room temperature.$^{727}$ Homocharge electrets (electron beam-embedded charges$^{2177}$) may have longer lifetimes at higher temperatures, or other thermostable electret materials with designed nanoscale charge patterns may be found.$^{3099}$

E. Thermal Insulation. The combustion assembly shown in Figures 6.4A and 6.4B measures (110 nm)$^2 \times 90$ nm. For an isolated hot
surface in proximity to a cooler surface, both surfaces of area A and separated by vacuum, the rate of heat transfer is given by:

\[ P_{\text{classical}} = A e_i \sigma \left( T_2^4 - T_1^4 \right) \]  

(Eqn. 6.19)

with variables as defined for \( P_i \) in Eqn. 6.11; the peak vacuum emission wavelength is given by Wien’s relation as:

\[ \lambda_{\text{max}} = \frac{0.2 \ h \ c}{k T_2} \]  

(Eqn. 6.20)

which, for \( h = 6.63 \times 10^{-34} \) joule-sec, \( c = 3 \times 10^8 \) m/sec, \( k = 0.0138 \) xJ/K, and \( T_2 = 310 \) K, gives \( \lambda_{\text{max}} = 9.3 \) microns; at \( T_2 = 540 \) K, \( \lambda_{\text{max}} = 5.3 \) microns.

At planar separations \( d \geq \lambda_{\text{max}} \), radiative power is independent of separation distance. However, at smaller separations \( d < \lambda_{\text{max}} \), radiative heat transfer may be greatly enhanced due to near-field coupling of nonradiative electromagnetic modes in the surfaces.\(^{62,63}\)

In particular, surfaces at temperatures between 300-600 K transfer heat approximately as \(-\sigma_{\text{cond}} d^{-4}\) for 0.2 microns \( \leq d \leq 2 \) microns, where \( \sigma_{\text{cond}} \) is electrical conductivity. Surfaces with 10 nm \( \leq d \leq 200 \) nm again exchange energy independent of \( d \), but according to:

\[ P_{\text{anomalous}} = 0.574 A \left( \frac{4 \pi^2}{h^2 c^2} \right) \sigma_{\text{cond}} T_2^4 (T_2 - T_1) \quad \text{(watts)} \]  

(Eqn. 6.21)

The basic energy exchange relation for \( d < 200 \) nm is simply \( P_{\text{glu}} = P_{\text{anomalous}} + P_{\text{ext}} \), where the combustion power \( P_{\text{glu}} = n_{\text{glu}} E_{\text{glu}} / t_{\text{glu}} \), \( n_{\text{glu}} \) = number of glucose molecules oxidized per cycle (e.g., \( 3 \times 10^7 \)), \( t_{\text{glu}} \) = combustion cycle time (e.g., 100 milliseconds), and \( P_{\text{ext}} \) is the working power extracted by the Stirling engine. Assuming \( d = 100 \) nm, a highly conductive silver or aluminum surface with \( \sigma_{\text{cond}} \sim 4 \times 10^{18} \text{ sec}^{-1} \) results in a negligible temperature differential \((T_2 - T_1) \sim 0.0001 \) K. However, a germanium surface (melting point 1221 K) with \( \sigma_{\text{cond}} \sim 10^{15} \text{ sec}^{-1} \) could allow a useful operating temperature \( T_2 > 600 \) K, giving a maximum Carnot efficiency \((T_2 - T_1) / T_2 = 0.48\%\) for gas internal combustion engines). With device volume ~0.002 micron\(^4\), power density would be ~10\(^9\) watts/m\(^3\). A boron surface (m.p. 2573 K) with \( \sigma_{\text{cond}} \sim 10^7 \text{ sec}^{-1} \) might provide even better vacuum insulation properties. More research is required on the mechanisms of anomalous radiative heat transfer in order to fully assess the feasibility of this approach.

6.3.4.5 ChemoElectric Cells

The human body provides a renewable source of chemoelectric energy that may be tapped by various means. As long ago as 1959, Pinneo and Kesselman reported powering an FM transmitter by simply inserting two steel electrodes into the brain of a cat, generating 0.5 microamperes at 40 millivolts. In another experiment by Reynolds in 1963, electrodes inserted subcutaneously and abdominally into an anesthetized rat produced 400 millivolts and ~10 microamperes, powering an oscillator circuit for 8 hours with no sign of decline. Since then, three classes of bioelectric energy have been investigated for medical use: those in which both anode and cathode are consumed and go into solution, and those in which the cathode is inert. In the first group, zinc (anode) and silver chloride (cathode) electrodes have been implanted in human test subjects for up to two years. The Zn goes into solution as Zn\(^+\); the AgCl cathode is converted into Ag and the Cl\(^-\) ion goes into solution. This chemoelastic source generates steady currents of ~0.2 microamperes/micron\(^2\) at -1 volt or ~10 pW/micron\(^2\) of electrode surface at 70%-90% efficiency, an energy density of 10\(^7\) watts/m\(^3\) assuming 20-nm thick electrodes. Similar output is obtained from galvanic pairs in the second group, of which the best studied is a zinc anode coupled with a palladium or platinum-black cathode; molecular oxygen combines with water at the cathode producing OH\(^-\) and H\(^+\) ions. Unfortunately, galvanic sources provoke a significant host reaction, including formation of layers of necrotic debris, free neutrophils, granulation tissue and complete fibrous connective tissue encapsulation of long-term implants.\(^{590-592}\)

Biogalvanic energy sources\(^{590,3527-3530}\) exploit the electrochemical potential between metallic electrodes (the fuel) in an electrolyte solution. Two broad classes of galvanic pairs which generate electrical energy have been investigated for medical use: those in which both anode and cathode are consumed and go into solution, and those in which the cathode is inert. In the first group, zinc (anode) and silver chloride (cathode) electrodes have been implanted in human test subjects for up to two years. The Zn goes into solution as Zn\(^+\); the AgCl cathode is converted into Ag and the Cl\(^-\) ion goes into solution. This chemoelastic source generates steady currents of ~0.2 microamperes/micron\(^2\) at -1 volt or ~10 pW/micron\(^2\) of electrode surface at 70%-90% efficiency, an energy density of 10\(^7\) watts/m\(^3\) assuming 20-nm thick electrodes. Similar output is obtained from galvanic pairs in the second group, of which the best studied is a zinc anode coupled with a palladium or platinum-black cathode; molecular oxygen combines with water at the cathode producing OH\(^-\) and H\(^+\) ions. Unfortunately, galvanic sources provoke a significant host reaction, including formation of layers of necrotic debris, free neutrophils, granulation tissue and complete fibrous connective tissue encapsulation of long-term implants.\(^{590-592}\)

The biofuel cell\(^{590,3526,3527}\) relies upon redox reactions in which neither anode nor cathode is consumed but merely act as catalysts, potentially a great advantage in nanomedical applications. The biofuel cell that has received the most attention for nanomedicine is the oxyglucose cell, which could rely upon an electrochemical process in which glucose is oxidized at the anode and molecular oxygen is reduced at the cathode according to the following equations:

\[ \text{C}_6\text{H}_12\text{O}_6 + 6\text{H}_2\text{O} \rightarrow 6\text{CO}_2 + 24\text{H}^+ + 24\text{e}^- \]  

(Eqn. 6.22)

\[ 6\text{O}_2 + 12\text{H}_2\text{O} + 24\text{e}^- \rightarrow 24\text{OH}^- \]  

(Eqn. 6.23)

Decades-old experiments using platinum black electrodes\(^{594}\) produced ~0.6 picoamps per micron\(^2\) of electrode surface at ~0.3 volts, with power levels of 0.2 pW/micron\(^2\) in vitro (declining over time due to the formation and absorption of gluconic acid at the anode), but only 0.004 pW/micron\(^2\) in vivo (when implanted in rat or rabbit test animals) in part due to poisoning of the catalytic anode action by proteins. Complete electrochemical oxidation of glucose has not yet been demonstrated experimentally.\(^{591-593}\)

Controlled-permeability membranes may eliminate these problems in nanomedical applications, but even using the aforementioned electrodes might allow a continuous power density of 10\(^7\) watts/m\(^3\) for a 10 pW oxyglucose biofuel cell using 50 micron\(^2\) of electrodes 20-nm thick. Since gluconic acid can also be oxidized at potentials

*Interestingly, when the rat was finally administered an instantly lethal intracardial injection of Nembutal the voltage required 75 minutes, after death, to fall to zero.*
which oxidize glucose, gluconate may not be the only byproduct, thus appropriate means of disposing of any potentially undesirable byproducts must be included in a practical design. The addition of specific catalysts (e.g., ruthenium) to the electrodes could be helpful.

Ethanol fuel cells, though inefficient, are already in wide commercial use. For example, the Lion alcohol sensor employs catalytic platinum electrodes (supported on a 1-cm² PVC matrix disk assembly impregnated with acidic electrolyte) to produce electron flow by oxidizing ethanol to acetic acid (producing two free electrons per alcohol molecule) at the anode and reducing atmospheric oxygen at the cathode. Normal alcohol working concentrations are 5-900 parts per million typically producing up to -10 microamps at -5 millivolts across a 390 ohm load, or ~50 nanowatts (power density at most ~0.1 watt/m³). The Lion fuel cell can also produce energy from all primary and secondary aliphatic alcohols but not from aldehydes, ketones, ether, esters, hydrocarbons or carboxylic acids. Other manufacturers of electro-chemical fuel cells for alcohol sensing include PAS Systems, Guth Laboratories, and Intoximeter. Direct methanol fuel cells also were under development in 1998.

The following analysis was prompted by R. Merkle’s suggestion that an oxyglucose biofuel cell using nanoscopic membranes might achieve significantly higher power densities than indicated by past experiments (Fig. 6.5). Consider a proton (H⁺) exchange nanomembrane consisting of a 1-nm thick diamondoid sheet containing a number of very narrow pores each lined with atoms of oxygen, fluorine or nitrogen, creating negatively charged channels with high proton affinity. A channel narrow enough to admit only protons but nothing else (~0.1 nm, excluding even helium atoms) and surrounded by a -3 atom thick support wall makes a ~1 nm³ pore structure ~1 nm wide.

The power generated by each pore due to proton flow is \( P_{pore} = q N V_p \) taking \( q = 1.6 \times 10^{-19} \) joule/proton-volt, \( N \) is the proton flow rate in protons/sec, and \( V_p \) is the electrical potential through which the charges fall. Experimental observations suggest \( V_p \) may reach 0.75 volt at zero load, but averages 0.3-0.6 volts at moderate loads; for this analysis, \( V_p \) will be taken as ~0.5 volt.1017 As for \( N \), the nicotinic acetylcholine receptor channel (Section 3.3.3) has a 0.65-2.2 nm inside diameter and allows 2.5 x 10⁷ Na⁺ ions/sec to pass while the channel is open, comparable to the transport rate of artificial transmembrane peptide nanotube ion channels.1177 Taking this flow rate as representative for \( N \), then \( P_{pore} \approx 2 \) pW/pore.

Redox fuel cells using transition metal catalysts typically achieve \( n_{cat} \approx 10 \) catalytic events/sec/catalytic atom.1017 Oxidation of each glucose molecule produces 24 protons, so the required catalytic rate is \( N/24 \approx 10^9 \) glucose molecules/sec per pore which in turn requires \( N / (24 n_{cat}) \approx 10^6 \) metal catalyst atoms per pore. Taking Pt atoms as representative, \( 10^7 \) Pt catalyst atoms have a volume of ~1500 nm³. Including the 1-nm³ pore structure and doubling the volume to allow for catalyst dispersal, fluid access and structural overhead gives a 3000 nm³ single-pore catalytic unit generating 2 pW, or a power density of ~7 x 10¹¹ watts/m³. (Assumed catalyst atom usage is ~10¹⁷ atoms/watt; in 1999, the best Pt-catalyzed proton exchange membranes required ~10¹⁸ atoms/watt.)3264 The current flow of 4 picoamps through each (0.1 nm)³ pore gives a current density of \( 4 \times 10^8 \) amps/m², well below the ~10¹⁰ amp/m² maximum current density in bulk aluminum (Section 6.4.3.1). Since ~10⁶ glucose molecules/sec are consumed per pore and each glucose molecule represents 4765 ßJ of free energy, then total energy available is ~4.8 pW/pore. Thus it appears the device may be ~2 pW / 4.8 pW ~ 40% efficient, though this figure is highly voltage-dependent.

J. Soreff notes that the oxyglucose biofuel cell can in theory operate near the thermodynamic limit but may require the design of a complex sequence of catalytic reactions, not unlike microbial metabolism.2427 The glucose/gluconic acid source may have lower efficiency due to incomplete oxidation but could be useful in early systems because it requires optimization of catalytic sites for just one reaction. On the other hand, the production of protons and electrons from complete oxidation of glucose potentially encompasses a lengthy biochemical chain which possibly may be implemented using sets of immobilized enzymes in the manner of mitochondrial respiration. (Mitochondrial power density is ~10¹⁻¹⁰⁵ watts/m³.)31,786

There are limits to efficiency set by mismatches between redox potentials at various steps in a respiratory chain. Given a single anodic chamber with a common electric potential and pH, concentrations of intermediates must be unequal to compensate for the redox potential mismatches at various steps. If these concentrations are too unequal, then the lowest concentration in the chain becomes the rate-limiting intermediate for the entire process. Throughput may be optimized by careful choice of intermediates, couplings between reactions, and catalyst design. To oxidize glucose, in biology, glycolysis requires ~18 enzymes, the TCA cycle ~9 enzymes, and the pentose phosphate pathway ~9 enzymes.526 Soreff suggests that a slightly different strategy is to have separate oxidation/proton chambers for each stage in the catalytic pipeline, allowing intermediates to diffuse through semipermeable membranes or to be transported via molecular sorting rotors. This transfers the problem of matching redox potentials into the electrical domain where it is more easily resolved.

Oxyhydrogen fuel cells are another possibility, though probably at much lower power density because hydrogen must first be produced from the glucose fuel. This occurs naturally in the energy-generating hydrogenosomes of the trichomonads (which use the enzymes pyruvate:ferredoxin oxidoreductase and hydrogenase) and in the laboratory using related enzymes (from bacteria that live near hot underwater vents) to convert glucose into hydrogen gas and water. In a modern 350 K oxyhydrogen fuel cell, O₂ and H₂ are fed into adjacent chambers separated by a proton exchange membrane. This membrane allows only H⁺ ions to flow to the O₂ side, making water as the sole waste product and establishing a net negative charge on the H₂ anode and a net positive charge on the O₂ cathode, an electrical potential of ~0.6 volts under load. This chemolectric process is at least ~50% efficient but typically achieves only ~10¹⁻¹⁰⁶ watts/m³ partly due to hydrogen’s low volumetric energy density and partly due to the crude membranes currently in commercial use. Bacterium-based biofuel cells have also been investigated.2427,5351

6.3.5 Electrical Energy Conversion Processes

Microscale electrostatic motors have been fabricated since the late 1980s and are the subject of great current research interest.556,557 Commercially available micromachined rotors -1 micron thick and ~100 microns in radius respond to electric field intensities >10⁴ volts/m producing motive torques of ~10-pN·m by converting electrical to mechanical energy, achieving power densities in the 10⁴⁻¹⁰⁵ watts/m³ range with device lifetimes approaching 10⁷ rotations.556

Drexler10 has described a class of submicron direct-current (DC) electrostatic motors capable of converting between mechanical and electrical power in either direction at a power density of ~4 x 10¹⁴ watts/m³ and an estimated efficiency >99%. (Typical macroscale brush DC motors can have 55%-75% efficiencies, while DC brushless motors can have efficiencies as high as 95%).
In one implementation (Fig. 6.6), electric charge is placed on the rim of a rotor as the rim passes within a dee electrode. This charge is then transported across a ~3 nm gap to the interior of the opposite dee electrode, where it is removed and replaced by a charge of opposite sign. Applying a voltage of proper sign across the dees causes the charges in transit to apply a torque to the rotor (like a Van de Graff generator operating in reverse), converting electrical into mechanical power. As a specific example, a motor 390 nm in diameter and 25 nm thick driven by an electric field strength of ~0.2 x 10^9 volts/m spins with a near-maximum rim velocity of ~1000 m/sec; a current of ~110 nanoamperes and an electric potential of ~10 volts yields a power output of ~1.1 microwatts. Single electron-transfer events at nanometer electrodes have been generated and measured experimentally, and macroscale "Dirod" electrostatic generators have been operated as electrostatic motors in a student laboratory.

Electrical energy is also converted into mechanical energy via the inverse piezoelectric effect. Piezoelectric materials such as lead-zirconate-titanate ceramics (PZT) and natural mechanical resonators such as quartz, when subjected to an electric field, expand or contract depending on the orientation of the field to the piezoelectric polarization. For example, applying 100 volts across a 750 micron thick piezoelectric crystal causes it to deform by ~37 nm, or ~0.4 nm/volt. There are also piezoelectric rotary actuators, with rotations measured in ~arcsec. The feasibility of fabricating ~1 micron thick quartz resonators by etching techniques was demonstrated decades ago. PZT crystals are commonly used to provide movement of samples being examined in scanning probe microscopes such as STMs and AFMs. Piezoelectric drivers used in bimorph configurations as electromechanical transducers can produce large mechanical displacements but relatively small forces. Dielectric induction micromotors 2 microns thick and 100 microns in diameter operating in water can be driven at 250 rev/sec at 110 volts applied potential, converting electrical to mechanical motion and producing a torque of 0.3 pN-m. Electroactive polymers (EAPs) and electrostrictive polymers such as i-PMMA are being investigated for use in "artificial muscles"; an exceptionally high electrostrictive response (~4%) has been observed in

Fig. 6.5. Schematic of oxyglucose biofuel cell with proton exchange nanomembranes.
Electrostatic actuators can produce substantial mechanical forces. One electromechanical transducer described by Drexler is a capacitor with one plate fixed and connected to a signal source, with the other plate grounded via a tunneling junction, leaving it free to move within a small range of displacements. One design with a 1-nm stroke length and a (12 nm)$^2$ plate area transduces an electrical voltage varying from 0-5 volts into a mechanical signal ranging from 0-1 nN. A similar implementation for electromechanical transduction involves drawing a dielectric slab of area $A_{\text{lab}}$ and dielectric constant $\kappa_{\text{lab}}$ into a gap $d_{\text{gap}}$ between the plates of a charged capacitor at voltage $V_{\text{cap}}$ producing a field-dependent force, making a simple mechanical dielectric drive. Applying a time-varying input voltage of frequency $v_\nu$ (Hz) produces a mechanical output power of:

$$P_n = \left(\frac{\varepsilon_0 A_{\text{lab}}}{2 d_{\text{gap}}^2} \right) \left(1 - \frac{1}{\kappa_{\text{lab}}} \right) v_\nu V_{\text{cap}}^2$$

For $\varepsilon_0 = 8.85 \times 10^{-12}$ coul/N-m$^2$ (vacuum permittivity constant), $A_{\text{lab}} = 100$ nm$^2$, $d_{\text{gap}} = 2$ nm, $V_{\text{cap}} = 1$ volt, $\kappa_{\text{lab}} = 5.7$ for diamond at 300 K, and $v_\nu = 5000$ Hz, $P_n = 0.001$ pW (power density 5 x 10$^9$ watts/m$^3$) and the force applied on the slab during $\Delta x = 1$ nm displacements is $P_n / (\Delta x v_\nu) = 1$ nN.

Electrothermal energy conversion occurs during simple joule heating, or during Seebeck effect or Peltier effect cooling, when current passes through a bimetallic junction causing one side to get hot and the other side cold.

Electrochemical transduction may occur in enzymes with electric field-sensitive conformational states, in voltage-gated ion channels and nanomembranes (Section 3.3.3), and in electrolytic cells which are chemoelectric cells (Section 6.3.4.5) operated in reverse. Quantum-well nanostructures can be useful in fabricating electrooptical transducers such as light-emitting diodes, electro-luminescent displays and submicron-scale solid-state lasers (Section 5.3.7), and electrochromic glass with voltage-dependent transmittance permits optical gating. Electric current sent through carbon nanotubes causes both field emission and luminescence, about one photon per 10$^6$ electrons—even a single nanotube makes a faint but visible glow, another case of electrooptical energy transduction.

Magnetomechanical transduction has been also employed to remotely power centimeter-size simple legged robots using magnetostrictive alloys as magnetic field-driven actuators in an 80 Hz 1-tesla field, but this approach may find only limited utility in nanomedical systems because molecular-scale drives suffer from the adverse scaling properties of electromagnet fields. Magnetic sensors (Section 4.7.2), possibly using molecular magnets, could be employed as power transducers.

### 6.3.6 Photonic Energy Conversion Processes

Visible light falling on a blackened surface is quickly absorbed and thermalized, the simplest example of photothermal transduction. Such transduction is reversed in solid-state optical refrigerators, which achieve cooling with 2% energy efficiency by flooding a sample with infrared photons of a particular frequency that will dampen molecular vibrations, then allowing the object to shed energy as higher-energy fluorescent light.

Optomechanical transduction has been demonstrated by using low-power laser light impinging upon an optofluidic converter to create a pressure pulse of sufficient magnitude to operate a pneumatic actuator in a push-pull mode, and an optical thermal "Brownian ratchet" uses rotating laser light to rectify the Brownian motion of a micron-scale plastic bead. Laser-heated gas micro-actuators have been studied, and scattering forces from optical tweezers have been used to rotate small particles and thus may provide a means to power nanomachines by nonmechanical means. Photoacoustic transduction has been induced directly in piezoelectric indium gallium arsenide optical signal delay devices and in "photostriction" materials such as PLZT and indirectly in aqueous suspensions of 30-nm carbon particles illuminated by a pulsed laser. A 200-micron long, 600-nm thick silicon nitride beam with a 250-nm thick coating of a glassy material made of arsenic and selenium (chalcogenide) contracts ~1 nm when illuminated by photons polarized along the length of the beam, and expands ~1 nm when illuminated by photons polarized across the width of the beam. A -300-micron optomechanical motor with diffraction grating developed for the DARPA MOEMS program operates at 20 volts and spins up to ~80 Hz.

Light-driven proton pumps achieve optochemical transduction when optical photons shining on a photosensitive micelle activate a proton pump that pumps hydrogen ions across a membrane, thus building up a large ionic charge differential across the membrane and chemically storing the energy of the incident light. Experimental photoelectrolysis cells producing hydrogen and oxygen from sunlight and water with 12.5% energy efficiency have been built. Green plants are optochemical transducers, but only 1-3% of the photonic energy falling upon a green plant is converted to chemical energy because most of the incident radiation is lost by reflection, transmission, or absorption by non-chloroplast pigments. Of the photons falling directly upon active photosynthetic pigments, up to 35% is converted to chemical energy. Certain bacteria also employ photochemical transduction. Activation of unimolecular chemical reactions by ambient blackbody radiation has been demonstrated, and 10-100 GHz rotation in nanotube gears with dipole charges using laser-generated alternating electric fields has been simulated computationally.

Photoelectric cells are a well-known technology for converting optical to electrical energy. Commercial solar panel arrays currently achieve ~10$^3$ watts/m$^3$, although polymer thin-film...
(100-200 nm) photovoltaic cells may reach $-10^6$ watts/m$^2$ at only 5.5% efficiency. Muller595 Multi-layer multi-band-gap photovoltaic cells can convert solar energy to electrical energy with efficiencies >30%, and photocells and various natural photo receptors such as the LH2 (light harvester) molecular complex of the Rhodopseudomonas acidophila bacterium can achieve efficiencies of 60%-90% at selected monochromatic wavelengths.3532,3533 Photovoltaic dendrimeric light-harvesting molecular antennas have been designed and synthesized.2586-2591

In the 400-700 nm optical band only, direct (cloudless noon-time) sunlight energy flux is 100-400 watts/m$^2$. Taking the 100 watts/m$^2$ figure and applying a 30% conversion efficiency, power supply $P_a$ to a nanodevice of circular cross-section of radius R is $P_a = 100 \times 30\% \pi R^2$ watts in direct sunlight requiring a 10-pW nanorobot. Fullerene sheets exhibit interesting opto-electronic effects: high-intensity light shined on one side of a sheet induces an electrical voltage potential on the other face. Progress is also being made in developing optically pumped (optophotonic) molecular lasers.990

Low-frequency photons such as radiofrequency signals are readily converted into electrical current using millimeter-scale loop antennas (Section 6.4.2). For example, a 2-mm x 10-mm device implanted subdermally with a hypodermic needle electrically stimulates muscle using rf telemetry for control and power, in order to restore human motor functions.563

### 6.3.7 Nuclear Energy Conversion Processes

The concept of using nuclear energy as a power source for nanorobots has not been widely discussed or explored13294, in part for sociocultural reasons. This Section describes a feasible radionuclide power system which emits essentially no radiation into the surrounding tissue. This is followed by a brief examination of the challenges in harnessing fission and fusion in medical nanodevices, and a few other somewhat speculative possibilities.

#### 6.3.7.1 Radionuclides

The most common mode of fissile energy release occurs when an unstable radioactive atomic nucleus emits a small particle such as an electron ($\beta^-$ decay), a positron ($\beta^+$ decay), a helium nucleus (alpha decay), or a high-energy photon (gamma decay), often transforming the nucleus into another element. Selection of an optimum radioactive fuel is guided primarily by safety criteria.

**A. Penetration Range** — The mass of the alpha-particle is ~7000 times greater than that of an electron, so the velocity and hence the range of alpha-particles in matter is considerably less than for beta-particles of equal energy. Consequently the optimum radionuclide for medical nanorobots is predominantly an alpha emitter. A simple “stopping power” approximation67 for the range of 1-10 MeV alpha-particles is:

$$R_\alpha (\text{microns}) \sim 430 Z^{1.3} E_{\alpha\text{MeV}}^{10.6} \rho_{\text{absorb}}$$

where $E_{\alpha\text{MeV}}$ = energy per alpha decay (MeV) and $\rho_{\text{absorb}}$ = absorber density (21,450 kg/m$^3$ for Pt). Thus a 3 MeV alpha-particle (initially traveling at ~0.04c) has a range of ~2 cm in air, ~30 microns in water, ~10 microns in germanium, and ~5 microns in platinum, with a fairly sharp cutoff and only ~1% straggling (fluctuation of particle range around mean range). By contrast, as a crude approximation the range of beta-particles of equal energy (initially traveling at ~0.99c) is ~10$^3$ $R_\alpha$ and the range of gamma photons and fast neutrons of equal energy is ~10$^3$ $R_\alpha$, although “range” for these particles is imprecisely defined because beam intensity is never truly reduced to zero but rather decays exponentially with increased shielding. Paritly because of the heavy shielding required, $\alpha$-powered beta batteries developed in the 1970s achieved only ~30 watts/m$^3$,3528 while $\gamma$-powered thermoelectric batteries produced up to ~58 watts/m$^3$.288 over 2300 nuclear-powered cardiac pacemakers were implanted without mishap during 1970-76,633 but new implantations apparently ceased in 1983.692

The absorption of alpha-particles in matter results almost entirely from collisions with electrons, or ionization reactions, which create 10$^5$-10$^6$ ion pairs per -MeV alpha-particle (depending upon absorber material), slowing the particle almost to a halt. For some light-nucleus absorbers, there may also be an extremely small contribution from direct interactions with the nucleus. Alpha-particles with energies as low as 4.8 MeV (e.g., from $^{88}$Ra$^{226}$) were observed to transmute atoms of all elements from $^\beta$ up to $^\gamma$ (except for $^\alpha$C and $^\alpha$O) with the emission of photons.1007 Low-energy alpha-particles can also occasionally transmute light-atom nuclei to produce neutrons—one of the strongest such reactions involves 4.6 MeV alpha-particles emanating from $^{88}$Rn$^{222}$ striking a $^\alpha$Be$^9$ absorber, wherein one alpha-particle out of every 5000 enters the beryllium nucleus and sends out a neutron (with a much longer range than an alpha-particle of equal energy), a 0.02% reaction probability.1008 Thus the worst case (e.g., shield materials are poorly chosen) is one low-energy proton or neutron emitted per 10$^5$ - 10$^6$ ion pairs generated (e.g., ~1 every 2000 sec at an alpha-particle current of 1 picoamp), though most such secondary reactions have far lower probabilities.

However, naturally emitted alpha-particles generally cannot penetrate the nuclei of heavier elements and therefore generally cannot create significant secondary radiation in heavy-nucleus absorbers, hence shielding should be made from heavier elements. To first order, overcoming Coulomb repulsion and entering the nucleus requires an alpha-particle energy $E_{\alpha\text{pulse}} > 4 \pi Z^2 Z' / 4 \pi r_{\text{nucl}}^2$ (joules),1005 where $Z$ is atomic number (nuclear charge number), $e = 1.60 \times 10^{-19}$ coul (elementary charge), $e_8 = 8.85 \times 10^{-12}$ farad/m (permittivity constant), and Rutherford’s classical formula for nuclear radius $r_{\text{nucl}} \sim Z^{1/3}$1006 where $\rho_{\text{nucl}} = 1.6 \times 10^{-15}$ m. Thus for example, using a $^{78}$Pt absorber, $r_{\text{nucl}} \sim 7 \times 10^{-15}$ m and $E_{\alpha\text{pulse}} > 33$ MeV; with a $^{52}$Ge absorber, $r_{\text{nucl}} \sim 5 \times 10^{-15}$ m and $E_{\alpha\text{pulse}} > 18$ MeV. In either case, the 2-4 MeV low-energy natural-emission alpha particles likely to be employed in nanomedical nucleoelectric systems (Table 6.5) have energies about an order of magnitude smaller than $E_{\alpha\text{pulse}}$.

**B. Gamma Rays** — Gamma rays are toxic to most biomaterials, have the greatest penetrating power, dissipate useful fissile energy, and may foster erosion of shielding due to electron-positron pair creation. Thus the ideal medical nanorobot radionuclide fuel emits no gamma-rays during disintegration.

**C. Decay Chain** — Radioactive elements typically decay to progressively lighter elements which may also be radioactive. Ideally, each of the daughter products of the ideal radionuclide should meet the above criteria for the entire decay chain down to stable nuclei. Since each decay event produces a fixed amount of energy, power density $P_{\text{rad}}$ of a radionuclide of density $p$ is proportional to the number of disintegrations per second, or inversely proportional to half-life, as:

*Some exotic artificial nuclei such as Li$^{11}$ do not obey the classical Z$^{1/3}$ rule.1129
where \( AW = \) atomic weight (gm/mole), \( t_{1/2} = \) half-life (sec), and \( k_{\text{rad}} = 10^9 \ln(2) \times N_A \times E_{\gamma} \), where \( N_A \) is Avogadro’s number and \( E_{\gamma} = 160 \text{ J/eV} \). To be useful in medical nanorobots, \( t_{1/2} \) should be at least 10 days (~10^6 sec) or longer for convenience of storage and use—ideally, commensurate with the anticipated mission duration to preclude the need for potentially hazardous refueling operations in vivo. Table 6.5 gives power densities \( P_{\text{rad}} \) for several candidates and data for a few other benchmark radionuclides.

Among all gamma-free alpha-only emitters with \( t_{1/2} > 10^8 \) sec, the highest volumetric power density is available using Gd\(^{148}\) (gadolinium) which \( \alpha \)-decays directly to Sm\(^{144}\) (samarium), a stable rare-earth isotope. A solid sphere of pure Gd\(^{148}\) (~7900 kg/m\(^3\)) of radius \( r = 95 \) microns surrounded by a 5-micron thick platinum shield (total device radius \( R = 100 \) microns) and a thin polished silver coating of emissivity \( e_r = 0.02 \) suspended in vacuo would initially maintain a constant temperature \( T_2 \) (far from a surface held at \( T_1 = 310 \) K) of:

\[
T_2 = \left( \frac{T_1^4 + \frac{P_\gamma}{3 e_r R^4}}{4 \sigma R^2} \right)^{1/4} - 600K \quad \text{(Eqn. 6.27)}
\]

with a 75-year half-life, initially generating 17 microwatts of thermal power which can be converted to 8 microwatts of mechanical power by a Stirling engine operating at ~50% efficiency. (Smaller spheres of Gd\(^{148}\) run cooler.) While probably too large for most individual nanorobot designs, such spheres could be an ideal long-term energy source for a swallowable or implantable “power pill” (Chapter 26) or dedicated energy organ (Section 6.4.4). A ~0.2 kg block of pure Gd\(^{148}\) (~1 inch\(^3\)) initially yields ~120 watts, sufficient in theory to meet the complete basal power needs of an entire human body for ~1 century (given suitable nuclearchemical energy conversion and load buffering mechanisms, and a sufficiently well-divided structure). Nuclear powered energy organs are discussed further in Section 6.4.4.

A 1 micron\(^3\) block of radioactive gadolinium yields a useful ~3 pW of thermal power. However, a minimum of 5 microns of Pt shielding is still required, so the minimum possible diameter of a zero-emissions Gd\(^{148}\)-powered nanorobot is ~11 microns, reducing power density to ~10\(^4\) watts/m\(^3\),** The low energy density would produce a surface temperature of 320 K, allowing only ~3% thermal conversion efficiency.** Higher efficiencies may be achieved using semiconductor junction \( \alpha \)-particle nucleoelectric transducers that convert the linear ionization trail of electron-hole pairs directly into electrical current in ~10\(^{-9}\) sec (Fig. 6.7). Such ionization in silicon and germanium is well-studied: the average energy loss per ion pair produced by a passing \( \alpha \)-particle is 3.6 eV (580 zJ) for Si, 3.0 eV (480 zJ) for Ge;\(^{567}\) an 11-micron thick Ge wall is required to stop the 3.18 MeV \( \alpha \)-particles. A (1 micron\(^3\)) cube of Gd\(^{148}\) produces ~5 \( \alpha \)-particles/sec, yielding an output current of ~1 picoampere at ~3 volts (e.g., ~5 pW). Unlike most nanomachinery for which a

### Table 6.5 Volumetric Radioactive Power Density of Various Radionuclides\(^{63}\)

<table>
<thead>
<tr>
<th>Radionuclide</th>
<th>Half-Life (sec)</th>
<th>Alpha Emissions (MeV)</th>
<th>Beta Emissions (MeV)</th>
<th>Gamma Emissions (MeV)</th>
<th>Other Emissions</th>
<th>Volumetric Power Density (watts/m(^3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( ^{64}\text{Ge}^{142} )</td>
<td>( 1.6 \times 10^3 )</td>
<td>1.50</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 3.0 \times 10^6 )</td>
</tr>
<tr>
<td>( ^{99}\text{Ir}^{135} )</td>
<td>( 2.2 \times 10^6 )</td>
<td>4.35</td>
<td>---</td>
<td>---</td>
<td>0.074-0.385</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{83}\text{Bi}^{210m} )</td>
<td>( 8.2 \times 10^3 )</td>
<td>4.90</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{96}\text{Cm}^{240} )</td>
<td>( 9.4 \times 10^4 )</td>
<td>2.70</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 1.9 \times 10^4 )</td>
</tr>
<tr>
<td>( ^{96}\text{Cm}^{242} )</td>
<td>( 1.8 \times 10^5 )</td>
<td>0.156</td>
<td>---</td>
<td>0.187-0.64</td>
<td>---</td>
<td>( 1.5 \times 10^5 )</td>
</tr>
<tr>
<td>( ^{96}\text{Cm}^{244} )</td>
<td>( 5.1 \times 10^9 )</td>
<td>4.77</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 4.8 \times 10^5 )</td>
</tr>
<tr>
<td>( ^{64}\text{Gd}^{148} )</td>
<td>( 2.4 \times 10^9 )</td>
<td>3.18</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 8.7 \times 10^4 )</td>
</tr>
<tr>
<td>( ^{61}\text{Pm}^{147} )</td>
<td>( 8.3 \times 10^7 )</td>
<td>---</td>
<td>0.225</td>
<td>---</td>
<td>0.043-0.15</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{94}\text{Pu}^{238} )</td>
<td>( 2.8 \times 10^9 )</td>
<td>5.48</td>
<td>---</td>
<td>---</td>
<td>0.043-0.875</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{88}\text{Ra}^{226} )</td>
<td>( 5.5 \times 10^9 )</td>
<td>5.79</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 2.0 \times 10^4 )</td>
</tr>
<tr>
<td>( ^{64}\text{Gd}^{150} )</td>
<td>( 3.1 \times 10^9 )</td>
<td>6.02</td>
<td>---</td>
<td>---</td>
<td>0.043</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{15}\text{P}^{32} )</td>
<td>( 7.5 \times 10^9 )</td>
<td>---</td>
<td>0.167</td>
<td>---</td>
<td>---</td>
<td>( 3.6 \times 10^4 )</td>
</tr>
<tr>
<td>( ^{35}\text{S}^{18} )</td>
<td>( 8.0 \times 10^7 )</td>
<td>6.11</td>
<td>---</td>
<td>---</td>
<td>0.043-0.10</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{84}\text{Po}^{210} )</td>
<td>( 6.3 \times 10^9 )</td>
<td>6.23</td>
<td>---</td>
<td>---</td>
<td>0.045</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{64}\text{Gd}^{148} )</td>
<td>( 1.4 \times 10^9 )</td>
<td>6.10</td>
<td>---</td>
<td>---</td>
<td>0.04409</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{97}\text{Po}^{210} )</td>
<td>( 1.2 \times 10^9 )</td>
<td>5.30</td>
<td>---</td>
<td>---</td>
<td>0.79</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{84}\text{Po}^{212} )</td>
<td>( 2.3 \times 10^8 )</td>
<td>6.24</td>
<td>---</td>
<td>---</td>
<td>1.71</td>
<td>SF</td>
</tr>
<tr>
<td>( ^{64}\text{Gd}^{146} )</td>
<td>( 1.5 \times 10^9 )</td>
<td>7.17</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 6.5 \times 10^4 )</td>
</tr>
<tr>
<td>( ^{84}\text{Po}^{212} )</td>
<td>( 3.0 \times 10^7 )</td>
<td>8.78</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>( 8.6 \times 10^2 )</td>
</tr>
</tbody>
</table>

SF = Spontaneous Fission

\*The orbit radius \( r_{\text{circ}} \) of an \( \alpha \)-particle of mass \( m = 6.68 \times 10^{-27} \) kg, charge \( q = 2 \times (1.6 \times 10^{-19}) \) coul and energy \( E_{\text{kin}} = 3.18 \text{ MeV} \) circling in a uniform magnetic field \( B = 1 \text{ tesla} \) in vacuo is \( r_{\text{circ}} = \left( k_0 E_{\text{MeV}} m / q B \right)^{1/2} = 0.3 \text{ meter} \) (\( k_0 = 3.2 \times 10^{15} \)), about as wide as a human body, so more compact in vivo nanocyclotron storage of \( \alpha \)-particle emissions to allow a controlled energy release is not feasible.

**Plutonium dioxide radioactive thermal generator (RTG) systems developed by NASA and DOE in the 1980s produced thermal power of ~10\(^5\) watts/m\(^3\) and electrical power of ~10\(^6\) watts/m\(^3\), a ~10% conversion efficiency. Conventional large nuclear power plants average ~32% conversion efficiency, according to U.S. Federal Energy Information Agency statistics.
single-point failure would ordinarily be fatal to machine function (Chapter 13), the design presented here should be extremely resistant to such failures. The C-C bond energy is also ~3.4 eV, so direct nucleochemical transduction of α-particle energy into chemical form is theoretically possible — conversion of carbonaceous material to diamond crystal by fission fragment irradiation has been reported.1029

Are Gd148-powered nanorobots safe? Yes, if their shielding remains intact. There are no β or γ emissions, and given adequate shielding, no α-particles can escape. Significant shield erosion is unlikely because:

1. almost all α-particle interactions are with orbital electrons, not nuclei, of shield atoms;
2. α-particle emission rates are low (e.g., ~5/sec);
3. radioactive impurities need not be present in nanomanufactured structures; and
4. background radiation is not a major reliability issue for homogeneous components larger than ~10 nm (Chapter 13).10

In the extremely unlikely event of complete shield removal, maximum range of α-particles in water or soft tissues is ~30 microns or ~one cell width. The radiation from the radioactive cores of ~1 billion Gd148-powered nanorobots, if all their shielding was destroyed, would deliver a ~500 rad lifetime lethal radiation dose in ~1 day; an unshielded ~0.2 kg block of Gd148 (~100 watt output) delivers an LD50 (see below) dose in ~10 seconds. Gd148 is a non-fissionable material, hence fissionable chain reactions or explosions are not possible. Indeed, gadolinium has the highest neutron absorption cross-section of any known element and is more useful in control rods or as a nuclear shield. As for chemical toxicity in the event of a breach, ionic Gd is rapidly converted to colloidal hydroxide and phosphates in the blood, which is then rapidly taken up by the reticuloendothelial system mostly in the liver and spleen;598 half-time in the lung is ~2 hours.600 In macrophages, the rare earths localize in lysosomes as insoluble phosphates.599 Rare earths such as Gd or Sm are not particularly toxic — LD50 is typically 0.5-5 grams/kg body weight,601 or 35-350 gm for a 70 kg human adult. No chemical carcinogenicity or mutagenicity has been found.601 In 1998, Gd148 could be purchased from Los Alamos National Laboratory for $0.50/micron.5,2315 This cost must be significantly reduced for Gd148-powered nanorobots to become economically feasible.

If greater operating power densities are required, radionuclide-powered nanorobots with power cores that are hotter or are smaller than ~1 micron3 (but with shield diameters still >11 microns) will require switching to a less-safe material of higher energy density, such as Po210, which decays to stable Pb206 with a 5-month half life. For example, a Po210 spherical power core of radius 100 nm has a surface temperature of ~600 K and produces ~6 pW which may be tapped via heat engine as described earlier. However, such a heat source also produces one 0.79 MeV gamma photon every ~10^4 sec; unshielded, ~300 such photons represent a lethal dose (~500 rad or ~5 joules/kg soft tissue) to a single mammalian cell, giving a dose rate of ~10 milliR/hr vs. ~5 milliR/hr for the early implantable Pu238 nuclear thermoelectric batteries.274 Thus the lifetime exposure* of any cell visited by Po210-powered nanorobots is restricted for safety to at most ~1 nanorobot-month, which may be acceptable in many short-duration applications. No such restriction would apply to the Gd148 device, discussed earlier.

Larger amounts of energy (up to ~0.1% of rest mass) may be released by the fission of a heavy atomic nucleus into nuclei of two lighter elements of roughly equal mass, followed by any of ~50 distinct decay chains down to stable nuclei for the unstable fission fragments, as for example:

\[
\text{U}^{235} \rightarrow \text{mixed fission products} + 200 \text{MeV} \ (3.2 \times 10^{10} \text{zJ})
\]

[Eqn. 6.28]

which achieves ~0.091% mass conversion to energy. During fission, typically one neutron is absorbed by a nucleus but 1-3 neutrons are emitted, making possible a chain reaction. In effect, neutrons act as fission catalysts. Nuclear species which may support a net-energy-producing fission chain reaction include U232, U233, U235 (critical mass ~3.6 kg), Pu239, Am241, and Am242 (thermal or fast neutrons) and Th232, Pa231 and U238 (fast neutrons only). From a

![Fig. 6.7. Schematic of semiconductor-junction nucleoelectric transducer.](Image)

*The rad is the standard unit of absorbed dose, defined as 1 rad = 0.01 joule/kg. LD50 is the absorbed dose that produces lethality in 50% of all exposures, typically 400-500 rads for human tissue. The traditional LD50 for a 70-kg human body (~500 rads) thus requires the absorption of ~350 joules. For a single ~10 nanogram human liver cell, assuming the traditional LD50 (~500 rads) would require the absorption of ~50 pJ of ionizing radiation energy. To rigorously assess the safety of radiation emitted from nanorobots, the effects of the particular radiation on specific cells or organs must be evaluated.1070-1074 There are large differences in organ sensitivity, ion-mass sensitivity, and environmental sensitivities such as variations in iron, Vitamin C or melatonin levels. Lethality of cells is at one end of a continuum which also includes lower-dose effects (e.g., cancer transformation rates) that disturb the cell’s function without killing it, thus harming the body the cell is serving. The availability of cellular repair devices vastly increases both LD50 and traditional “lifetime exposure” radiation limits that may be tolerated by humans (Chapter 24). Substantial genomic, biotechnological, and nanotechnological engineering to increase the radiation tolerance of cells (Chapter 24) also could help to minimize cancer risks and birth defects, and to allow human beings to more easily live and work in space. Nevertheless, a fundamental nanomedical design principle is to avoid using devices whose normal behavior causes additional damage to the body (Chapter 11).
6.3.7.2 Nuclear Fusion

Still larger energies (up to ~1% of rest mass) are available from nuclear reactions in which light nuclei fuse to form heavier nuclei, although it appears that conventional nuclear fusion is unlikely to be practical for in vivo nanomedical applications. The deuterium-tritium (D-T) fusion reaction requires the lowest temperature to ignite—141 million K, which is 12.1 KeV or 1.94 x 10^-15 joules per nuclei—and achieves 0.38% mass conversion to energy:

\[ \text{D}^2 + \text{T}^3 \rightarrow \text{He}^4(3.5 \text{ MeV}) + n(14.1 \text{ MeV}) \]  

(Eqn. 6.29)

A D-T atom pair has volume \(10^{-3} \text{ nm}^3\); maximum diamondoid piston (controlled) compression at \(10^9 \text{ atm}\) puts \(10^{20}\) joules into this volume, too low by a factor of \(\sim 10^6\). However, target bombardment may allow micron-scale fusion ignition. Although magnetic cyclotron confinement of 12.1 KeV particles at \(B = 1\) tesla requires a storage ring of radius \(r_{\text{circ}}\sim 3\) cm (Section 6.3.7.1), the length of a linear particle accelerator scales as \(-K.E./qE\) (Section 4.7.4) so a \(10^9\) volt/m electric field requires only \(-12\) microns to accelerate an ionized triton or deuteron to the ignition temperature.

The first major problem with the D-T reaction from a nanomedical standpoint is the dangerously high flux of 14.1 MeV fast neutrons, which in human tissue interact with hydrogen to produce deuterium plus a 2.2 MeV γ-ray, and with nitrogen to produce radioactive C14 plus a 0.6 MeV proton. Neutrons, like γ-rays, are highly penetrating radiation, so using the D-T reaction will require much thicker shielding (see below). One alternative is to switch to a cleaner-burning (aneutronic fusion) but highly penetrative radiation, such as the classical Cockcroft-Walton reaction involving proton bombardment of lithium which achieves 0.23% mass conversion to energy:

\[ \text{Li}^7 + p \rightarrow \text{Be}^7 + n \]  

(Eqn. 6.30)

(A competing neutron-emitting reaction \((\text{Li}^7 + p \rightarrow \text{Be}^7 + n)\) occurs only at incident proton energies >1 MeV, and radiative absorption of the proton, producing a radiated γ-ray, occurs as a sharp resonance absorption at a proton bombardment energy of 440 KeV.\(^{2463}\) A -10^9 volt/m linear accelerator -125 microns in length can produce 125 KeV protons, and from Eqn. 6.25 a platinum shield ~18 microns thick absorbs ~80% of the 8.36 MeV α-particles produced, thermalizing the energy.

The second major problem with micron-scale fusion is its low cross-section reaction. Cross-section is a measure of the probability that a given nuclear reaction will take place. If \(I_0\) incident protons/sec provoke \(I_1\) reactions/sec in a sample of thickness \(x\) with nuclear cross-section \(\sigma_{\text{nuc}}\) barns (1 barn = 10^-28 cm^2) and target density \(N_{\text{target}}\) (atoms/cm^3) = \(\rho\ N_A/\text{AW}\), then

\[ \frac{I_1}{I_0} = 1 - e^{-N_{\text{target}}\sigma_{\text{nuc}}x} \]  

(Eqn. 6.31)

To get more energy out of the fusion reaction than is put into the protons (e.g., energy breakeven), the reaction probability \((I/I_0)\) > (125 KeV / 16.73 MeV) = 0.7%. For \(\text{Li}^7\), \(A = 7\) and \(\rho = 0.534\) gm/cm^3 so \(N_{\text{target}} = 4.6 \times 10^{22}\) atoms/cm^3; with \(\sigma_{\text{nuc}} = 0.01-0.001\) barns for the reaction over this energy range,\(^{662}\) then to achieve energy breakeven the lithium target must be at least \(x > 10^2-160\) cm thick. However, the range of 125 KeV protons in lithium is only ~0.0003 cm,\(^{667}\) so energy breakeven cannot be achieved. As for the D-T reaction, one calculation for laser compression of D-T fuel pellets indicated that the minimum pellet diameter to achieve (explosive) ignition is ~4.2 mm and requires a total incident energy of 11.2 megajoules\(^{602}\) in a ~nanosecond pulse. Smaller pellets could not achieve energy breakeven.\(^{603}\) It is estimated that the minimum dimension of a high-temperature fusion reactor may be on the order of centimeters.\(^{603}\) But even if a -10^11 joule/m^2 energy density can be obtained with diamondoid materials allowing the reactor core to be as small as (5 cm)^3, a 2.3-cm thick Pt shield for the 14.1 MeV neutrons only reduces the incident intensity by half. Thus the smallest medically safe D-T fusion reactor is probably at least ~1 meter in diameter (e.g., to reduce incident fast neutron intensity by ~10^-4 upon exit).

D-D fusion produces either neutrons or radioactive tritium effluent, only 3-4 MeV per reaction, and achieves only ~0.1% mass conversion to energy. The D-He\(^+\) reaction produces 0.39% mass conversion with H and He\(^3\) effluent (both particles charged, hence field-manipulable), but is harder to ignite (~60 KeV). There are always unavoidable neutron-producing reactions due to D-D reactions in the primary fuel and D-T reactions between the primary deuterium and the secondary tritium, although Petrie\(^{2462}\) noted that neutron reactions can be less than ~5% of the D-He\(^+\) reactions. He\(^3\) is also very rare on Earth. The H-B\(^1\) reaction may produce C\(^12\) and is probably exergic, but γ-rays are emitted.\(^{1006,535}\)

6.3.7.3 Exothermal Nuclear Catalysis

One solution to low nuclear reaction probability and a poor energy balance is to employ a nuclear catalyst. A number of possibilities have been investigated, as described below.

In principle, the two deuterons in a deuterium molecule can spontaneously fuse to form tritium + proton or He\(^3\) + neutron, liberating 4 Mev of energy. The two electrons in the D\(_2\) molecule act as a catalyst, holding the deuterons together so they can react. According to quantum mechanics, the deuterons can tunnel toward each other through the classically forbidden region of repulsion until they get so close (~2 x 10^-15 m) that the strong force dominates and fusion occurs.\(^{605}\)

In practice the rate of this reaction is very small, ~10^-74 molecule^-1 sec^-1.\(^{609}\) But if an electron of mass \(m_e\) is replaced by a heavier negatively charged particle such as a muon (\(M_{\text{muon}} \approx 207\) Me), forming a muonic molecule, the required tunneling distance shortens by the ratio of the masses—in this case, from \(5 \times 10^{-11}\) m to \(2 \times 10^{-13}\) m, making penetration of the barrier much more likely and dramatically raising the reaction rate to ~10^-3 molecule^-1 sec^-1.\(^{507}\)

Muon catalysis of the proton-deuteron reaction, initially proposed theoretically by Frank in 1947,\(^{606}\) was first observed experimentally in 1957 by Alvarez.\(^{604}\) It has since been shown to be an effective means of rapidly inducing fusion reactions in low-temperature

*More recent hydrodynamic simulations of a 10-micron D-0-vapor-doped D\(_2\) gas bubble compressed to 0.38 micron in 0.5 micrometer using a 5-atom triangular spike superimposed on a 1-atm 27.6 KHz oscillating driving pressure produces a 2.2 KeV peak central temperature ~11 picoseconds before the bubble reaches minimum radius, which should be hot enough to generate at least a small number of thermonuclear D-0 fusion reactions in the bubble.\(^{322}\)
(<1200 K) mixtures of hydrogen isotopes, with D-T reaction rates of \(-10^{-9}\) sec\(^{-1}\). The field has had its own technical journal, *Muon Catalyzed Fusion*, since 1987, and there are excellent recent review articles. But muon catalysis has long been considered impractical for large-scale fusion reactors because the muon is relatively short-lived (2 microsec) and is quickly captured by a helium nucleus formed in a fusion reaction. Typically the number of fusions catalyzed by a muon during its lifetime is ~150 in liquid D\(_2\)/T\(_2\), but ~1000 are needed to achieve energy breakeven given the energy cost of artificial muon production. More than 70% of the cosmic ray flux at Earth’s surface consists of positive and negative muons, but the cosmic-ray induced fusion rate is still impractically low, ~10\(^{-7}\) watts/micron\(^3\) in liquid deuterium targets.

It has also been demonstrated experimentally that various changes in chemical composition, pressure, or electric fields can also act as nuclear catalysts, increasing the rates of nuclear transformations:

A. Chemical Composition — The rate of electron capture for Be\(_7^+\) is 0.08% greater in BeF\(_2\) than in metallic Be.

B. Mechanical Pressure — The electron capture (EC) decay rates of Te\(_{49}^+\) and Ba\(_{131}^+\) are measurably altered at a pressure of 100,000 atm. At 230,000 atm there is a 0.35% increase in the electron density at the nucleus of the free Be atom; the observed increase in the EC decay constant of Be\(_7^+\) oxide with pressure is so linear that it may be used as a method of pressure measurement in diamond anvil experiments in which optical access is impossible. It has been suggested that very high pressures could induce fusion. One experiment in which Pd and Ti immersed in D\(_2\O\) were bombarded with intense ultrasound apparently produced above-background levels of He\(_4\), an expected endproduct of D-D fusion processes.

C. Fracture Deformations — Fracto-fusion experiments have detected neutron emission when a crystal of lithium deuteride or heavy ice is mechanically fractured, believed to be the consequence of deuteron acceleration by >10 KeV electric fields generated by a propagating crack in the crystal, consistent with D-D fusion. Heating, cooling, or fracturing metal specimens exposed to high-pressure D\(_2\) (e.g., deuterated titanium) frequently produces statistically significant bursts of neutrons and emission of charged particles, rf signals and photons. It is proposed that crack growth results in charge separation on the newly formed crack surfaces, accelerating D\(^+\) ions in the electric field across the crack tip to energies >10 KeV sufficient to significantly raise the D-D fusion probability. Neutrons are reportedly generated when fragments of titanium are crushed with steel balls in a bath of heavy water. It has also been speculated that the core of the spherical acoustic shock wave generated during sonoluminescence, if it remains stable to a 10-nm radius, might reach temperatures adequate to fusion \(\cong 10^6\) K.

D. Electric Fields — Claytor et al at Los Alamos National Laboratory passed a current of 2.5 amperes at 2000 volts through 200-micron diameter palladium wires in a glow discharge tube of D\(_2\) gas at 0.3 atm for ~100 hours apparently producing ~10 nanocuries of tritium, with great care being taken to eliminate possible sources of contamination. Deuterium-saturated LiTaO\(_3\) crystals in a 75 KV/cm AC field exhibit elevated neutron emission attributed to D-D fusion.

Wires of LiD exploded by high current pulses also emit fusion neutrons.

E. Metallic Deuterides — Most controversial is the speculative possibility of metallic deuteride catalyzed fusion at temperatures between 300 K–1100 K, first reported (then later partially retracted) in the years 1926-27. Positive results are reported for a comprehensive series of experiments conducted at SRI International for the Electric Power Research Institute (EPRI) during 1989-94 and for another comprehensive series of experiments conducted by the U.S. Naval Air Warfare Center at China Lake during 1989-96 and U.S. patents have been issued for such devices (e.g., Patterson and Cravens, U.S. #5,607,563 on 4 March 1997). In another class of experiments [Edmund Storms, personal communication, 1996], a hydrogenophilic metal such as palladium is loaded with deuterium at effective pressures \(\cong 10^{-10}\) atm, giving molecular loadings of D/Pd = 85%-95%. Palladium deuteride normally exists either as Pd\(_2\)D or PdD, but it is believed that the highest loadings may give rise to significant concentrations of PdD\(_4\) \((x=1-2)\), asserted to be the “nuclear active phase.”

Superstoichiometric palladium hydride \((x = 1.33)\) at ~50,000 atm has been observed experimentally in x-ray diffraction studies, although preliminary molecular simulation studies of deuterium-entrained metal lattices have given pessimistic results.

Upon applying a current of ~1 nannoamper/micron\(^3\) at ~1-10 volts to a superstoichiometric metallic deuteride, significant heat energy in excess of the electrical input is said to be developed as the deuterium is consumed, on the order of \(10^6-10^7\) watts/m\(^3\). He\(_4\) is claimed to be produced at the expected rate of \(~10^{-11}\) He\(_4\) atoms/sec/watt, with neutrons, tritons (tritium nuclei), \(\gamma\)-rays and X-rays missing or detected in amounts far too small to account for the excess energy, which is asserted to be evidence of a catalyzed D-D anionic process at work. If the results of these experiments were confirmed, it might become possible to use diamondoid pistons to maintain continuously high deuterium loadings in an active catalytic crystal and thus to develop 1-1000 pW of aneutral thermal energy in a precisely nanomanufactured porous 1 micron\(^2\) metal-deuteride reactor with He\(_4\) (23.85 MeV) as the principal (and benign) effluent, achieving storage densities >10\(^{16}\) joules/m\(^3\) which would allow a completely self-contained >10 year fuel supply to be carried aboard a 10 pW nanorobot in a ~1 micron\(^3\) fuel tank.

6.4 Power Transmission

Power may be stored in, and converted among, many different forms, but the ultimate energy resource must be received from some external source by medical nanodevices (e.g., chemical energy from ingested food). The following discussion concentrates on acoustic (Section 6.4.1), electromagnetic (Section 6.4.2), and tethered (Section 6.4.3) power transmission, and closes with a brief description of dedicated energy organs (Section 6.4.4).

6.4.1 Acoustic Power Transmission

Acoustically powered medical nanorobots may derive their energy either from sources indigenous to the human body (Section 6.3.3 and Table 6.3) or from artificial sources placed in or on the human body. Large arrays of microscale acoustic radiators could create focused coherent sonic beams. Such artificial sources are likely to involve ultrasonic frequencies designed to minimize any aural
discomfort to the user. According to the official statement issued by the American Institute of Ultrasound in Medicine (AIUM) in 1978, no significant biological effects have been reliably observed in mammalian tissues exposed in vivo to unfocused –1 MHz ultrasound with intensities of 1000 watts/m² or less, although the onset of continuous-wave ultrasound-induced lysis of human erythrocytes has occasionally been detected experimentally at intensities as low as 60 watts/m² at 1.6 MHz when micron-size gas bubbles are also present. There are also no significant biological effects observed after continuous exposures exceeding 1 second in duration for total energy transfers of 500 kilojoule/m² or less. (The comparable hazard threshold for UV excimer laser light is ~0.5 kilojoule/m².)

Exposures of any duration >500 kilowatts/m² may cause cavitation and other harmful effects in biological tissue (see below). On the basis of additional experimental results, the AIUM in 1988 slightly revised its statement to allow intensities as high as 10,000 watts/m² for exposures to highly focused sound beams, which is about the highest that human volunteers can tolerate. (In 1985 the FDA allowed ultrasound intensities up to 7300 watts/m² for cardiac use, 15,000 watts/m² for peripheral vessels, and 1800 watts/m² for fetal, abdominal, intraoperative, pediatric, cephalic, and small-organ (breast, thyroid, testes) imaging.)

Figure 6.8 summarizes current standards in simplified graphical form.

Deleterious biological effects may occur in at least five ways:

1. **Transient Cavitation** — Transient bubbles which implode produce temperature increases of ~10³ K and pressure spikes of ~10⁴ atm localized in regions of a few microns in radius. Normal or transient cavitation requires ~10⁵ watts/m² (-5.4 atm) at 30 KHz or ~10⁶ watts/m² (~17 atm) at 1 MHz to form in water. Intensities less than ~10⁴ watts/m² will not produce transient cavitation in any tissue.

2. **Stable Cavitation** — Small pre-existing bubbles surrounded by water resonate in synchrony with the acoustic field, with the liquid acting as the oscillating mass and the gas serving as the compliant component. For an air bubble of radius \( r_{\text{bubble}} \) (meters) in water, resonant frequency \( f_{\text{res}} = 0.3 / r_{\text{bubble}} \) (Hz) for hard-shell-less bubbles up to ~1 MHz, thus a 6-micron bubble resonates at ~600 KHz. Resonating bubbles have been reported in therapeutic beams at power intensities as low as 6800 watts/m² at 750 KHz.

3. **Heating** — Dissipation of vibrational energy can initially heat tissues ~1 K/minute if applied at, say, 50,000 watts/m² at 3 MHz, in the hazard zone of Figure 6.8. Continuous exposure to 2000-6000 watts/m² at 0.1-10 MHz raises human tissue temperature by 1 K at equilibrium, which is considered safe. Ultrasound intensities of ~10³ watts/m² at 750 KHz may give rise to small voltages in bone via the piezoelectric effect.

4. **Acoustic Torque and Fluid Streaming** — Physical fluid motions are driven by ultrasonic radiation pressure, typically ~0.001 N/watt, or ~1 pN/micron² at 1000 watts/m². Forces of this magnitude or larger can cause damaging shear stresses in molecules or cells and may give rise to small voltages in bone via the piezoelectric effect.

5. **Shock Wave Formation** — Shock waves most easily form in liquids having low attenuation such as urine in the bladder or amniotic fluid. A 3-MHz 10-atm pulse shows a shock waveform after passing through 5 cm of water.

Consider a micron-scale nanorobot located somewhere deep within the human body that wishes to receive acoustic power from an artificial external source. A “safe” \( I_{\text{power}} = 1000 \text{ watt/m}² \) source pressed against the skin suffers three major reductions in intensity before its emanations reach the nanorobot. First, the power transferred to the entire body is reduced by a geometrical power reduction factor (PRF) which may be approximated as the ratio of the area of the transmitter in contact with the skin to the largest planar section through the torso, ~1500 cm². Second, after conversion of incident power intensity to pressure amplitude using Eqn. 4.53, transmitted amplitude is reduced by reflection losses (Eqn. 4.54), which may range from a 90% loss for a poorly-coupled source to just 10% loss for a well-coupled transmitter. Third, the pressure amplitude of the power signal is attenuated by absorption, reflection, and scattering as it passes through human tissue, as described by Eqn. 4.52.

Thus attenuated, the power signal finally reaches the nanorobot and is converted to mechanical power using a piston-type transducer as described by Eqn. 6.13. Combining all the above factors gives received power \( P_n \), which is a cubic and exponential function of acoustic frequency \( \nu_p \) and an exponential function of acoustic path length \( X_{\text{path}} \), which factors interact to produce an optimum acoustic power transmission frequency for various path lengths as shown in Figure 6.10. \( P_n \) is given by Eqn. 6.13, with \( \Delta V = V_l^3 \) and \( \Delta P = (1 – R_{\text{loss}})^2 \rho_{\text{sound}} I_{\text{power}} \text{PRF} \left( \frac{2 \rho \nu_{\text{sound}} I_{\text{power}} \text{PRF}}{1} \right)^{1/2} e^{-\alpha_{\text{transmit}} A_{\text{transmit}}/A_{\text{incident}} \nu_{\text{sound}} X_{\text{path}}}, \) where \( \rho = 993.4 \text{ kg/m}³, \nu_{\text{sound}} = 1500 \text{ m/sec in water at } 310 \text{ K (Eqn. 4.53), and } R_{\text{loss}} = 1 – A_{\text{transmit}}/A_{\text{incident}} \) from Eqn. 4.54.

In operating table scenarios wherein acoustic transmissions from the table may reach all parts of the body within a very short path length (~10–30 cm) and the 1000 watt/m² transmitter may be well-coupled to ~40% of the patient’s body surface giving PRF >1.00(100%), the power available to a ~0.3 micron² nanorobot receiver may be quite large, on the order of ~10⁴ pW or greater even at sub-MHz frequencies at the 1000 watts/m² incidence level. Similar power levels might also be achieved using carefully-designed permanently-implanted acoustic radiator organs which are themselves powered by some external source such as induced emf, backpack batteries or even tethered household electrical current.

The greatest engineering challenge arises when whole-body acoustic power is to be supplied by some small extradermal source...
the size of a wristwatch, wallet, belt buckle, ankle bracelet, headband, or amulet worn around the neck. The power curves in Figure 6.9 assume a wristwatch-band-sized radiator (~30 cm²) with a PRF of 2% and a single 0.3 micron receiver on the nanorobot. For a conservative design the longest acoustic path length in the human body (~200 cm) would govern the calculation, giving an optimum acoustic power transmission frequency of ~60 KHz, providing >40 pW in this configuration for a path entirely through soft tissue (α_{tissue} = 8.3 x 10^{-6} sec/m; Table 4.2). Thus a 10-pW nanorobot may be powered on just a ~25% duty cycle, allowing plenty of “quiet time” for acoustic communications, sensing, and navigational activities. A 1000 watt/m² transmitter with 30 cm² of contact area provides 3 watts of acoustic power, enough to supply ~10^{11} 10-pW nanorobots; ~10^{13} 10-pW nanorobots could be supported in the operating table scenario. For comparison, typical average power outputs of medical ultrasound scanners are 1-70 milliwatts at 0.5-20 MHz.\textsuperscript{506,628}

As a practical matter, acoustic power transmission may require slightly higher frequencies (e.g., >110 KHz) to ensure that some patients do not hear an annoying high-pitched noise that would be perceived as being centered in the head. Ultrasonic hearing via bone or body-fluid conduction has been reported at least up to 108 KHz in humans, probably mediated by the saccule, an otolithic organ that normally responds to acceleration and gravity.\textsuperscript{1372} Physical discomfort,\textsuperscript{3536} intra-articular pain,\textsuperscript{3537} and a lowering of electrical pain sensation threshold\textsuperscript{3538} due to ultrasound exposure have also been reported in humans.

Figure 6.10 quantifies the effects of transdermal power reflection and receiver size, neither of which significantly influence the choice of optimum acoustic frequency although larger receivers have a steeper frequency cutoff due to rapidly rising piston inertia losses with frequency (Section 6.3.3).

![Figure 6.9. Received acoustic power vs. acoustic frequency for various path lengths in the human body.](Image)

Net power received by medical nanorobots is also strongly influenced by the type of tissue lying across the acoustic path because of the exponential dependence of received power on the absorption coefficient α_{tissue} (Fig. 6.11). Similarly, Table 6.6 shows that very highly absorbent tissues such as bone, organs containing gas bubbles such as the stomach or bowel, and especially the air-filled lungs efficiently scatter and reflect large amounts of acoustic energy. Nanorobots positioned such that these highly attenuative tissues lie between them and the source should still receive adequate power due to internal acoustic reflection, since 99.95% of the acoustic amplitude reaching the skin-air interface from inside the body is reflected back into the body (Eqn. 4.54).

Available power declines rapidly for nanodevices located inside bone, bowel, or lung tissue, creating an “energy shadow”. Fortunately the path lengths are sufficiently short within these regions of the body that such losses should not become severe. For example, a 1 micron\textsuperscript{3} receiver located 200 cm from a 1000 watt/m² transmitter at 60 KHz with PRF = 2% and reflection loss = 10% at the skin receives 153 pW if the acoustic path passes through soft tissues only, 46 pW if the 200 cm acoustic path includes 1 cm of bone (with a 66% reflection loss at the tissue-bone interface), or 7 pW if the path length includes 5 cm of lung tissue (with an 80% reflection loss at the tissue-lung interface). If two or more acoustic power sources are simultaneously employed, destructive interference might occur within spatially periodic regions of smallest width on the order of one wavelength, or ≥ v_{sound}/ν_{p} = 1 cm (>>r_{nano}) for 150 KHz waves in water at 310 K, reducing locally available power.

Energy concentration, the opposite of energy shadowing, can also occur. If all nanorobot activity is confined to a specific organ or other small volume, a concave focusing transducer can shape pulse waves so that they arrive at a specific focal point within the body at high intensity. Spheroidally shaped regions in which the speed of sound is slower than in the surrounding medium (e.g., lower bulk modulus, or higher density or elasticity, as for instance a fat globule inside an organ; Eqn. 4.30 and Table 6.7) will tend to concentrate acoustic energy towards the center by refractive focusing.\textsuperscript{928} Sound waves also bend toward the cold side of a thermal gradient, the low-pressure side of a pressure gradient, and the low-concentration side of a salinity gradient. Large tubular tissue systems with areal cross-sections >(–λ)\textsuperscript{3} can act as acoustic waveguides; e.g., for 60 KHz waves in water, λ ~ 2.5 cm, roughly the diameter of the human aorta.

In general, liquids within the body are only weakly absorbing, allowing maximum transmission. Common sonolucent fluids include urine, aqueous humor, vitreous humor, amniotic and cystic fluids. In medical ultrasound, water immersion scanning is commonplace (99.77% transmission at water-tissue interface; Table 6.6) and a full bladder is a standard technique for obtaining an ultrasound window into the uterus. Muscle tissue acoustic absorption is anisotropic; a difference of a factor of 2.5 has been reported between the attenuation across and along its fibers.\textsuperscript{628}

6.4.2 Inductive and Radiofrequency Power Transmission

Electrically-powered nanorobots may tap natural electric fields already present in the body (Section 4.7.1) or may seek to acquire electromagnetic power from external sources. Transdermal magnetic induction and radiofrequency energy transfer has been employed in radio tracking and biomedical telemetry since at least the
Basic Capabilities • Power

1920s, 634, 635, 637, 638, 3510 and commercially available miniaturized implantable transmitters are now commonplace in laboratory work,639 but using this technique to power submillimeter devices is a relatively recent concept.630

The biological effects of radio waves, microwaves, and infrared rays are usually equivalent to the effects of heating, although non-thermal rf-induced vasodilation, possibly due to Ca++ flow manipulation, apparently has been observed in frogs.3328 Radio waves mainly induce thermal agitation of molecules and excitation of molecular rotations, while infrared rays excite vibrational modes of large molecules and may release fluorescent emission as well as heat. Both types of radiation are preferentially absorbed by unsaturated fats. In the United States the maximum permissible continuous occupational exposure level to microwave radiation has been 50 watts/m² for the testes and 100 watts/m² for the whole body—essentially double the thermal radiance of a 100-watt (~2000 Kcal/day) ~2 m² human body, e.g., ~ 50 watts/m² across the skin, or roughly the intensity of direct sunlight on the skin (Section 4.9.4). Figure 6.12 compares several generally accepted international occupational and population exposure limits that have been adopted.824 Michaelson823 carried out extensive investigations and found no evidence for hazard

<table>
<thead>
<tr>
<th>Tissue Interface</th>
<th>Energy Reflected</th>
<th>Energy Transmitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kidney/Liver</td>
<td>0.0036%</td>
<td>99.9964%</td>
</tr>
<tr>
<td>Liver/Muscle</td>
<td>0.032%</td>
<td>99.968%</td>
</tr>
<tr>
<td>Muscle/Blood</td>
<td>0.07%</td>
<td>99.93%</td>
</tr>
<tr>
<td>Soft tissue/Water</td>
<td>0.23%</td>
<td>99.77%</td>
</tr>
<tr>
<td>Fat/Kidney</td>
<td>0.64%</td>
<td>99.36%</td>
</tr>
<tr>
<td>Lens/Vitreous humor</td>
<td>0.91%</td>
<td>99.09%</td>
</tr>
<tr>
<td>Fat/Liver</td>
<td>1.00%</td>
<td>99.00%</td>
</tr>
<tr>
<td>Lens/Aqueous humor</td>
<td>1.04%</td>
<td>98.96%</td>
</tr>
<tr>
<td>Fat/Muscle</td>
<td>1.08%</td>
<td>98.92%</td>
</tr>
<tr>
<td>Bone/Muscle</td>
<td>41.23%</td>
<td>58.77%</td>
</tr>
<tr>
<td>Soft tissue/Bone</td>
<td>43.50%</td>
<td>56.50%</td>
</tr>
<tr>
<td>Bone/Fat</td>
<td>48.91%</td>
<td>51.09%</td>
</tr>
<tr>
<td>Soft tissue/Lung</td>
<td>63.64%</td>
<td>36.36%</td>
</tr>
<tr>
<td>Diamond/Water</td>
<td>90.80%</td>
<td>9.20%</td>
</tr>
<tr>
<td>Muscle/Air</td>
<td>98.01%</td>
<td>1.99%</td>
</tr>
<tr>
<td>Water/Air</td>
<td>99.89%</td>
<td>0.11%</td>
</tr>
<tr>
<td>Soft tissue/Air</td>
<td>99.90%</td>
<td>0.10%</td>
</tr>
<tr>
<td>Diamond/Air</td>
<td>99.997%</td>
<td>0.0025%</td>
</tr>
</tbody>
</table>

*45% by volume in water
**pneumonitis

**Table 6.7 Velocity of Sound in Various Organs and Biological Materials at 310 K**

<table>
<thead>
<tr>
<th>Biological Material</th>
<th>Velocity of Sound (m/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air bubbles*</td>
<td>20</td>
</tr>
<tr>
<td>Lung (collapsed)**</td>
<td>320</td>
</tr>
<tr>
<td>Dry air</td>
<td>353</td>
</tr>
<tr>
<td>Lung (normal)</td>
<td>630</td>
</tr>
<tr>
<td>Silastics</td>
<td>950</td>
</tr>
<tr>
<td>Ethanol</td>
<td>1160</td>
</tr>
<tr>
<td>Fat</td>
<td>1470</td>
</tr>
<tr>
<td>Water</td>
<td>1500</td>
</tr>
<tr>
<td>Aqueous humor</td>
<td>1510</td>
</tr>
<tr>
<td>Castor oil</td>
<td>1520</td>
</tr>
<tr>
<td>Vitreous humor</td>
<td>1540</td>
</tr>
<tr>
<td>Soft tissue (avg)</td>
<td>1540</td>
</tr>
<tr>
<td>Brain</td>
<td>1550</td>
</tr>
<tr>
<td>Liver</td>
<td>1565</td>
</tr>
<tr>
<td>Whole blood</td>
<td>1570</td>
</tr>
<tr>
<td>Kidney</td>
<td>1575</td>
</tr>
<tr>
<td>Muscle</td>
<td>1580</td>
</tr>
<tr>
<td>Optic nerve</td>
<td>1615</td>
</tr>
<tr>
<td>Lens of eye</td>
<td>1630</td>
</tr>
<tr>
<td>Glycerol</td>
<td>1880</td>
</tr>
<tr>
<td>Collagen (across axis)</td>
<td>2940</td>
</tr>
<tr>
<td>Bone (typical)</td>
<td>3550</td>
</tr>
<tr>
<td>Collagen (along axis)</td>
<td>3640</td>
</tr>
<tr>
<td>Skull bone</td>
<td>4090</td>
</tr>
<tr>
<td>Enamel (tooth)</td>
<td>5800</td>
</tr>
<tr>
<td>Diamond</td>
<td>~17,300</td>
</tr>
</tbody>
</table>

* ~45% by volume in water
** pneumonitis
at 100 watts/m², though there is conclusive evidence of potentially hazardous effects at levels above 1000 watts/m².

Single exposures to levels up to ~100 times these values may be briefly tolerated without injury, as for example, photosetting dental fillings which require ~20 sec exposures to visible intensities of ~10⁵ watts/m². Optical tweezers also demonstrate that monochromatic exposures of ~10¹¹ watts/m² are tolerated by biological macromolecules immersed in an optically transparent aqueous medium due to fast thermal equilibrium at the micron scale.¹⁶⁴ On the other hand, the threshold of pain for ink-blackened human skin exposed to a radiant heat stimulus is ~10,000 watts/m² for 3-second exposures.⁵⁸⁵

An electromagnetic wave passing through the human body declines in intensity as it heats tissues and is attenuated approximately according to:

\[ I = I_0 e^{-\alpha_d} \]  \hspace{1cm} \text{(Eqn. 6.32)}

where \( I = \) power per unit area transmitted through the tissue, \( I_0 \approx 100 \text{ watts/m}^2 \) (maximum safe incident intensity), \( d \) is depth of tissue penetration (meters), and \( \alpha_d \) (meter⁻¹) is the total attenuation factor including scattering and absorption, the inverse of the mean free path. For radio frequency and microwave radiation, \( \alpha_d = \alpha_E \nu E^2 \), where \( \nu \) is incident frequency (Hz) and \( \alpha_E \) ranges from 2 x 10⁻³ sec⁻¹/² m⁻¹ for muscle to 10 x 10⁻³ sec⁻¹/² m⁻¹ for vitreous humor but averages ~5 x 10⁻³ sec⁻¹/² m⁻¹ for soft tissue.⁶³⁵ At 100 MHz, 99% of incident power is removed in a 5-cm path length of soft tissue; at 1000 MHz (microwaves), 99% attenuation occurs within ~15 mm; at 10⁶ MHz (far infrared), ~0.5 mm of tissue removes 99% of the energy. On the other hand, 53% of the incident energy at 0.1 MHz passes through a 20-cm thick human body uns scattered and unab sorbed, thus remains available for utilization by medical nanodevices seeking electrical power. The dependence of \( \alpha_E \) on \( \nu \) varies significantly across the electromagnetic spectrum,⁵⁶⁹,⁶⁷⁶,⁷²⁷ additionally, \( \alpha_E \) for different tissues may deviate by up to a factor of 10 from the average values shown in Figure 6.13, especially in the IR-UV region.

Heetderks⁶⁸⁶ has analyzed the feasibility of transmitting electrical power into the body using submillimeter receiver coils—a magnetic transcutaneous link, which is preferred when large power-transfer rates are required. In one of several designs, Heetderks starts with a 14-cm diameter transmitter coil (diameter of the neck) with 11 turns (~40 microhenries inductance) and a 10 volt peak driving voltage at a frequency of 2 MHz delivering 4 watts RMS into the coil, a high but probably acceptable incident intensity of ~260 watts/m² at the skin. Inside the body, Heetderks uses a 400 micron diameter ferrite-core receiver coil with 60 turns spaced 50 microns apart, making a total receiver solenoid length of 3000 microns with ~0.89 microhenries inductance and Q ~ 40 for both transmitter and receiver, giving a fractional energy loss per cycle of 2\( \pi \)T / Q ~ 16%. The receiver produces 1.1 volts and 320 microwatts in the ideal case where the receiver lies coplanar within the circumference of the transmitter coil. Large noncoplanar displacements may produce significant performance declines.

Each of Heetderks’ receivers develops a power density of ~10⁶ watts/m². Up to 600 receivers may be present within the volume of operation of each transmitter, collectively drawing up to 10% of unloaded transmitter coil power consistent with maintaining Q ~40. In a nanomedical application, each receiver could act as a microscopic power station, absorbing the externally supplied electromagnetic energy and then transducing it into other forms of energy which may more directly be tapped by in vivo medical nanorobots. In theory, a 100 watt/m² flux imposed on the entire ~2 m² human body surface should allow the transdermal importation of ~20 watts, enough to support a population of ~10¹² 10-pW nanorobots assuming 50% energy transduction efficiency among the widely dispersed “power stations”.

How much smaller could Heetderks’ receivers be made? A detailed design analysis is beyond the scope of this book, but a simple scaling estimate may be made as follows. For a receiver coil of magnetic inductance \( L_m \) (henries), resistance \( R_c \) (ohms), and rf waves of angular frequency \( \omega \) (rad/sec), a coupling coefficient (\( k_{\text{coupling}} \)) between transmitter and receiver coils may be crudely approximated as the ratio of the power stored in the magnetic field of the LR oscillator to the total power delivered to the receiver coil by the transmitter, or (dividing out the terms for current) \( k_{\text{coupling}} = \frac{\rho_{\text{wire}} L_m}{(R_c + \rho_{\text{wire}} L_m)} \).

For \( \nu_E = \nu_L / 2\pi \), \( R_c = 1.26 \times 10^{-6} \text{ henry/m} \), \( \mu_r \approx 5 \) (a self-inductance factor when the receiver coil of \( N_{\text{coil}} \) loops contains a ferrite core⁶⁸⁶), \( \rho_{\text{wire}} \text{ wire resistivity} \approx 3 \times 10^{-3} \text{ ohm-m} \), and \( L \) is the characteristic size of the receiver (meters), then \( L_m = \mu_0 \mu_r N_{\text{coil}}^2 L^3 \), \( R = \rho_{\text{wire}} N_{\text{coil}}^2 / L \), and

\[ k_{\text{coupling}}^{-1} = 1 + \frac{L_m}{\nu_L L^4} \]  \hspace{1cm} \text{(Eqn. 6.33)}

where \( L_m = \rho_{\text{wire}} / 2\pi \mu_0 \mu_r \). For discrete-component (macroscopic) radios that receive broadcast signals, \( L \approx 1 \text{ cm} \) and \( \nu_E \approx 1-100 \text{ MHz} \), giving \( k_{\text{coupling}} = 0.9-0.999 \), which is very good. For Heetderks’ submillimeter receiver designs, \( k_{\text{coupling}} = 10^{-3} - 10^{-5} \); for example, the receiver described above has \( L \approx 700 \text{ microns} \), \( \nu_E = 2 \text{ MHz} \), giving \( k_{\text{coupling}} = 5 \times 10^{-4} \), which is very poor. With such poor coupling, very large transmitter intensities are required to induce very small receiver power levels. Troyk and Schwanz⁶⁸⁶ have suggested that printed thin-film coils integrated directly with sensor circuitry driven by a special transmitter coil topology could provide adequate amounts of power to implanted coils having coupling coefficients up to two orders of magnitude lower than Heetderks’ designs. However, even \( k_{\text{coupling}} \approx 10^{-2} \) at 2 MHz only reduces receiver size to \( L \approx 500 \text{ microns} \), which is not much improvement. The L⁴ scaling of \( k_{\text{coupling}} \) suggests that magnetic transcutaneous power receivers smaller than ~500 microns may be impractical, given the negligible amounts of magnetic field energy that can be stored in nanoscale volumes (Section 6.2.4) and the likelihood that small AC circuits will be heavily damped.¹⁰ (See also Section 7.2.3.)
DE / \text{watts/m}^2 \) where \( c = 3 \times 10^8 \text{ m/sec} \) (speed of light) and \( n = \) transmitter field is \( DE = \frac{1}{2} E_{e} \alpha_{bob}^{1/2} \) in the receiver transmits to the nanorobot a received power of \( P_n \approx 0.8 \text{ pW} \) delivered to the nanorobot producing a power density within the housing of \( D_n = P_n / L_{trav} L^2 \approx 10^6 \text{ watts/m}^3 \). In a practical system, these receivers must be aligned precisely with the incident field to maximize energy coupling, possibly requiring dynamic orientation control of each onboard receiver. Resonant beams of ~30 microns in size that can serve as rf filtering and oscillator elements have been demonstrated experimentally at ~15 MHz. High-frequency sub-micron electrometers with charge per unit bandwidth sensitivity of ~0.1 electron/Hz/\text{coul/m}^2 have been demonstrated, and submicron rf mechanical resonators have been discussed by Cleland and Roukes.

In 1999, Pharma Seq was developing laser-photon-powered cubic ~250-micron rf microtransponders for use with oligonucleotide probes in DNA assays; suspended in slurry, each device had an integrated circuit storing the sequence of the oligo attached to it, and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule bounded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a coded signal (the probe's serial number) to a nearby receiver when a fluorophore-labeled target DNA molecule binded and emitted a....
operating temperature $T_{\text{max}}$ then from the standard electrical formulas and Eqn. 6.19:

$$R = \frac{\rho}{A_W}$$  \hspace{1cm} \text{(Eqn. 6.36)}

$$I = \frac{V}{R} = \left(\frac{P}{R}\right)^{1/2} \left(4 \frac{e^2 \sigma A_{W} L^{2} (T_{\text{max}} - T_{0})}{\rho_{e}}\right)^{1/2}$$  \hspace{1cm} \text{(Eqn. 6.37)}

$$V = IR = \left(\frac{P}{R}\right)^{1/2} \left(4 \frac{e^2 \sigma A_{W} L^{2} (T_{\text{max}} - T_{0})}{\rho_{e}}\right)^{1/2}$$  \hspace{1cm} \text{(Eqn. 6.38)}

$$\rho_{e} = \frac{4 \frac{e^2 \sigma A_{W} L^{2} (T_{\text{max}} - T_{0})}{\rho_{e}}}{P_{e}}$$  \hspace{1cm} \text{(Eqn. 6.39)}

$$D = \frac{\rho_{e}}{A_W} = \frac{4 \frac{e^2 \sigma A_{W} L^{2} (T_{\text{max}} - T_{0})}{\rho_{e}}}{P_{e}}$$  \hspace{1cm} \text{(Eqn. 6.40)}

where $\sigma = 5.67 \times 10^{-8}$ watt/m$^2$K$^4$ (Stefan-Boltzmann constant). Thus for a silver wire 1 micron in diameter and 1 meter long, with $\rho_e = 1.6 \times 10^{-8}$ ohm-m, $\epsilon = 0.02$ for polished silver, $T_{\text{max}} = 373$ K (boiling point of water) and $T_{0} = 310$ K (body temperature), then $R = 16,000$ ohms, $I = 50$ microamps (current density $I_d = 5 \times 10^7$ amps/m$^2$), $A_W = 1$ micron$^2$, $A = 4$ mm$^2$, $V = 0.9$ volt, $P = 50$ micro-watts (sufficient power to operate a load of 10 electrostatic motors $\sim 390$ nm in diameter, as described in Section 6.3.5), and $D = 5 \times 10^7$ watts/m$^3$. Note that the minimum quantum unit of resistance in atomic- or molecular-scale wires is $\hbar/2e^2 \sim 13,000$ ohms, $h = 6.63 \times 10^{-34}$ joule-sec (Planck's constant) and $e = 1.6 \times 10^{-19}$ joule/eV, due to quantized conductance (e.g., Coulomb blockade) in narrow channels whose characteristic transverse dimension approaches the electronic wavelength.\(^{1740}\) By 1998, metallic wires down to $\sim 20$ nm in width had been fabricated.\(^{1517}\)

Wires with electrical and thermal insulation can safely carry higher currents in vivo (Section 6.5.5). If the wire in the preceding example is wrapped with a thermal insulator of thickness $d_{\text{insul}}$, then the maximum current increases to:

$$I = \left(\frac{K_{e} (T_{\text{max}} - T_{0}) A_{W} d_{\text{insul}}}{L \rho_{e}}\right)^{1/2}$$  \hspace{1cm} \text{(Eqn. 6.42)}

Using water as the thermal insulator ($K_{e} = 0.623$ watt/m-K at 310 K), a $d_{\text{insul}} = 1$ micron layer allows the wire to rise to 373 K while the outside of the insulating jacket remains near 310 K. A 100-micron thick water jacket allows electrical current and voltage to safely rise by a factor of 10, thus power to rise by a factor of 100.

Another possibility is the fullerene (pure carbon) conductors (Section 10.2.2.1). For example, 7-12 nm nanotube ropes 200-500 nm in length at $\sim 310$ K conduct $\sim 50$ nanoamperes at $\sim 1$ millivolt, transmitting a total power of $\sim 50$ pW.\(^{641}\) Boron nitride coated carbon nanotube wires should be highly oxidation-resistant.\(^{1308}\) Field emission from room-temperature carbon atomic wires has been measured as 0.1-1 microamperes at 80 volts, or $\sim 10$ micro-watts.\(^{643}\) Nanowires a few mm wide and hundreds of microns long have already been fabricated, and continuous meter-length threads (and longer)

have been proposed; current densities $\sim 10^{14}$ amps/m$^2$ have been pulled through a single chain of carbon atoms.\(^{643}\)

If fullerene nanowires are doped with metal atoms, adding electrons or holes so that the charge carrier density is high, electrical conductivity may be precisely engineered from semiconductor levels up to equal or better than copper in the plane of the graphene sheet. Multitude cables may display a quantum confinement effect that sharply reduces the resistance, possibly giving an electrical conductivity 10-100 times that of copper at room temperatures, and may provide shielding comparable to coax. Doped electron-rich conductors are easily oxidized on the nanometer scale, but wrapping such conductors in a second concentric undoped outer nanotube provides an insulating jacket that is effectively chemically impermeable under normal conditions. By 1998, coaxial nanocables $\sim 30$ nm in diameter had been fabricated in lengths up to 50 microns.\(^{2016}\) Boron-nitrogen (BN) equivalents of fullerene nanotubes are physically strong and good insulators. Low-temperature superconductivity has been demonstrated in doped\(^{644}\) and undoped\(^{626}\) fullerenes.

An interesting alternative is a conveyor belt system for direct mechanical charge transport, crudely analogous to the charging system of a Van de Graaff generator. Nanoscale conveyor belts (Section 3.4.3) could transport $\sim 10^6$ electrons/sec at a belt speed of 4 mm/sec, comparable to the $\sim 1$ cm/sec drift speed of electrons flowing in macroscale wires (Section 6.5.5). This gives a 0.2 picoampere current through a device with cross-sectional area $\sim 10^{-6}$ m$^2$, a modest current density of $\sim 10^3$ amperes/m$^2$ and electrostatic belt stress of $\sim 1$ atm between adjacent charges spaced $\sim 4$ nm apart. If a much smaller belt can be designed of a scale comparable to the 150-micron long computer data storage tape (Section 7.2.6) described by Drexler,\(^{10}\) a belt speed of $\sim 10^6$ m/sec and charge density of $\sim 1$ electron/nm$^2$ ($\sim$-amino acid density) on a 1 nm$^2$ belt gives a current of $\sim 2$ nanoamperes and a current density of $\sim 10^3$ amperes/m$^2$, comparable to the electromigration-limited $\sim 10^3$ amperes/m$^2$ maximum current density in bulk aluminum,\(^{*}\) with an electrostatic belt stress of $\sim 3$600 atm. Net belt mechanical energy dissipation may be as low as $\sim 10^{-6}$ zJ/nm of travel for each $\sim$nanometer-scale charge carrying device (Section 3.4.3).

6.4.3.2 Electromagnetic Tethers

Optical energy may be piped from one place to another by allowing light to enter a narrow solid fiber of clear plastic or fused silica. Light undergoes total internal reflection at the glass-air boundary and follows the contours of the light pipe, with losses due to scattering and absorption in the material as each photon undergoes up to $\sim 10^5$ reflections/meter during transit. The scattering minimum in commercial ultrapure silica fiberoptic cable occurs at $\sim 1500$ nm, or $\sim 2 \times 10^{-14}$ Hz, in the near infrared. These losses are roughly equivalent to transmission in clear air, negligible over distances of relevance in nanomedicine: measurable amounts of energy may be transmitted through single fibers $\sim 100$ km long.

At the lowest frequencies, the cutoff frequency for electromagnetic waveguides of diameter $d_{\text{guide}}$ filled with material of dielectric constant $\kappa_{e}$, below which frequency the waveguide cannot transmit photons (velocity $c = 3 \times 10^8$ m/sec in vacuo, and slightly lower in various dielectric materials at various frequencies), is $\nu_{\text{cutoff}} = c / 2 \kappa_{e} d_{\text{guide}}$ $\sim 7.5 \times 10^{13}$ Hz (for $\kappa_{e} = 1$ in the near infrared—close to the scattering minimum at $\sim 1500$ nm—for $d_{\text{guide}} = 2$ microns and $\kappa_{e} = 1$).

This cutoff applies only to waveguides; coax (shielded single wire) or triax (shielded twisted pair) lines can conduct frequencies all the

*This “maximum” may be quite conservative: “A single crystal metal wire surrounded by a strongly bonded, lattice-matched sheath should be stable at far greater current densities than a conventional wire of the same material”\(^{10}\).
way down to DC and all the way up to near-infrared (Section 7.2.5.1).

More specifically, if a nonmagnetic weak dielectric (-air) fills the space between two coaxial cylindrical conductors, with inside wire of radius \( r_{in} \), outer jacket of radius \( r_{out} \) and line current \( I_{line} \), then the average transmitted power\(^{272}\) is given by:

\[
P_E = 30 I_{line}^2 \ln \left( \frac{r_{out}}{r_{in}} \right) \text{ (watts)} \quad \text{[Eqn. 6.43]}
\]

In coax theory,\(^{272}\) \( \frac{r_{out}}{r_{in}} = 1.65 \) allows maximum power to be carried at a given breakdown voltage gradient across the dielectric (most appropriate for power transmission), whereas \( \frac{r_{out}}{r_{in}} = 3.6 \) gives the lowest attenuation due to conductor losses for a given outer diameter (most appropriate for communications applications). Thus for a 1-micron power coax, \( r_{out} = 0.5 \) micron implies \( r_{in} = 0.3 \) micron; assuming \( I_{line} = 10^6 \) amps/m² on the inside wire, \( P_E < 10^4 \) pW at -1 millivolt and the line has a characteristic impedance of 30 ohms.

For the highest frequencies, optical fibers may be as small as \(-0.5\) micron in diameter, roughly the photon wavelength—a dielectric rod can serve as a waveguide crudely analogous to the hollow metal pipes used at rf frequencies. Blue photons (\(-400 \) nm) carry \(-500 \) zJ of energy, almost enough to break C-C bonds (\(-550 \) zJ). Ultraviolet (UV) wavelengths shorter than \(-300 \) nm are greatly attenuated for fiber lengths \( >1 \) meter due to heavy absorption, and intense prolonged UV irradiation of silica creates defects (color centers) in the material that lead to further absorption of the laser light in the fiber.

In biomedical applications, silica fibers are used to conduct photons from excimer lasers, the brightest known sources of UV radiation.\(^{359,3560}\) Typical maximum continuous power intensities are \(-30,000 \) watts/m² for corneal sculpting, \(-10^6 \) watts/m² for bile duct cholangiocarcinoma tumor surgery, and \(-10^6 \) watts/m² for arterial debulking, laser dental machining and laser lithotripsy for kidney stones,\(^{645,646}\) so \(-0.01-1 \) microwatts may be delivered to medical nanodevices using a single \(-1 \) micron\(^2\) optical tether transmitting UV photons. The maximum transmittable power intensity is approximated by Eqn. 6.40; taking \( c_L = 0.80 \) for silica, \( A_{n/2} = 1 \) mm diameter fiber, \( L = 1 \) meter and \( T_{max} = T_0 = 20 \) K at room temperature for a handheld surgical instrument, then \( P_{max} = 4 \times 10^9 \) watts/m², in line with the above figures for excimer lasers assuming a \(-40\)% duty cycle. For a nanomedical optical power cable with \( A_{n/2} = 1 \) micron diameter fiber, \( L = 1 \) meter, \( T_0 = 310 \) K and \( T_{max} \leq 373 \) K, then \( P_{max} \leq 1 \times 10^9 \) watts/m² and a maximum of \( P \) \(-1 \) milliwatt may be delivered down the fiber. However, at this high fluence there is tremendous risk of localized tissue incineration, should the fiber detach in vivo and the photon flow is not immediately halted—a good argument for limiting optical power tether intensities to \(-10^8 \) watts/m² or less. A thin adjacent backchannel fiber could serve as a fuse, rupturing simultaneously with the main fiber and cutting off the feedback control signal.

### 6.4.3.3 Hydraulic and Acoustic Tethers

Power is easily conveyed to medical nanodevices through tethers by simple hydraulic means. For example, a virtually leakproof thick-walled pullerene nanotube measuring \( 2r_{tube} = 1 \) micron in diameter and \( l_{tube} = 1 \) meter long could safely transport pressurized fluid of absolute viscosity \( \eta = 6.9 \times 10^{-4} \) kg/m/sec \((310 \) K water\) at a fluid pressure \( P_{tube} \approx 5 \) atm to drive a mechanical turbine or valved reciprocating piston system, establishing an energy density of \( 0.5 \) pl/micron\(^3\) and delivering \( P_n = \pi r_{tube} P_{tube} / 8 \eta l_{tube} \sim 10 \) pW to a \( V_n = 1 \) micron\(^3\) nanorobot power plant at a fluid flow velocity of \( v_{fluid} = 2 \pi r_{tube} P_{tube} / 8 \eta l_{tube} \approx 20 \) microns/sec and producing a power density of \(-10^5 \) watts/m², assuming Poiseuille flow.

Acoustic waves could be delivered by tether, but slightly larger tubes are required. Consider a water-filled pipe with dimensions as defined in the previous paragraph which is excited at one end by a vibrating piston operating at a frequency \( v_{p} \approx v_{sound} / (2 \pi f_{tube} / 1 \) GHz for \( 2r_{tube} = 1 \) micron). For such small tubes, the inertia and kinetic reaction of the fluid may be neglected in favor of the frictional force (e.g., Poiseuille flow; Eqn. 9.25), since the compressions and rarefactions of the fluid are practically isothermal on account of the almost perfect heat conduction, and the power transmission is given by:

\[
P_n = P_0 \exp(-2 \pi r_{tube} l_{tube}) \text{ (watts)} \quad \text{[Eqn. 6.44]}
\]

with the attenuation coefficient given by Rayleigh’s classical formulation:\(^{649,650}\)

\[
\alpha_{tube} = \frac{(8 \pi \gamma \eta v_p)}{\rho v_{sound}^2 r_{tube}^{1/2}} \text{ (cm/m)} \quad \text{[Eqn. 6.45]}
\]

where \( \gamma \) is the ratio of specific heats (1.004 for water, 1.009 for seawater), \( v_{sound} = 1500 \) m/sec and \( \rho = 993.4 \) kg/m³ for water at 310 K. Maximum energy transfer (like a rigid bar) occurs at integral multiples of half-wavelengths of the incident sonic waves, in other words, at \( v_p = n v_{sound} / 2 l_{tube} \), where \( n = 1 \) is the fundamental frequency or first harmonic, \( n = 2 \) is the first overtone or second harmonic, and so forth. Minimum attenuation occurs at \( n = 1 \), the first harmonic.

A periodic source pulse of an intensity low enough to avoid cavitation in pure water (\(-10^9 \) watts/m²; Section 6.4.1) applied as input power \( P_0 \) across a tube area of \( \pi r_{tube}^2 \) produces a 10 pW output power for \( l_{tube} \approx 300 \) microns at \( v_p = 2.5 \) MHz and \( r_{tube} = 0.5 \) microns with \( P_0 = 8000 \) pW and \( \alpha_{tube} = 11,000 \) m⁻¹. To get \( l_{tube} = 1 \) meter length requires \( r_{tube} = 15 \) microns and \( v_p = 750 \) Hz to deliver 10 pW at the output. (A three-phase “alternating current” acoustical power transmission system using 100 Hz waves traversing three water-filled pipes was patented by M. Constantinesco in 1920.)\(^{650}\)

Diamond rods can also make nearly lossless acoustical power transmission lines (Section 7.2.5.3).

### 6.4.3.4 Gear Trains and Mechanical Tethers

Lengthy nanoscale gear trains may transmit power over great distances because mechanical efficiency for steric nanogear pairs may be \(-99.997\%\) (Section 6.3.2). Thus the initial mechanical input power only falls to 94% after passing through a sequential gear train of 2000 units, a 10-micron long train assuming each gear is \(-5 \) nm in diameter.

Torque power may also be transmitted via a long rotating cable inside a fixed tubular sheath, somewhat resembling an automobile speedometer cable or a dentist’s drill (\(-10\)–100 Hz). (Reciprocating cables are briefly treated in Section 7.2.5.4.) Two modes of power transmission are readily distinguished—first, a twist-and-release or “AC” strategy, resulting in the propagation of a torsion wave which looks locally like a shear wave, and second, a constant-twist or “DC” strategy, in which a driver end turns the cable at a constant rate (up to the bursting velocity) and the cable is maintained at constant torque (up to the shear strength limit). In the following analysis, we consider a cylindrical transmission cable of radius \( r_{cable} \), length \( l_{cable} \).

Basic Capabilities • Power 167
and density $\rho$, with shear modulus $G$ and working stress $\sigma_w$. The author thanks J. Soreff for helpful clarifications.

The AC case may be thought of as the propagation of a shear sound wave with maximum shear $\sigma_w$ and maximum strain $s_{\text{max}} = \sigma_w / G$, traveling at $v_{\text{shear}} = (G / \rho)^{1/2}$, the transverse wave velocity in the cable medium. The maximum power passing through the cable is $P_{\text{AC}} = \pi r_{\text{cable}}^2 D E v_{\text{shear}}$, where energy density $D E = G s_{\text{max}}^2 = \sigma_w^2 / G$, hence:

$$P_{\text{AC}} = \frac{\pi r_{\text{cable}}^2 \sigma_w^2}{\rho^{1/2} G^{1/2}} \quad \text{(watts)} \quad \text{(Eqn. 6.46)}$$

limiting maximum cable operating frequency to $-v_{\text{shear}} / L_{\text{cable}}$ or:

$$v_{\text{AC}} \leq \frac{G^{1/2}}{\rho^{1/2} L_{\text{cable}}} \quad \text{(Hz)} \quad \text{(Eqn. 6.47)}$$

In the DC case, the maximum surface speed of a spinning cable is the flywheel bursting speed $v_{\text{burst}} = (\sigma_w / \rho)^{1/2}$ and the maximum shear stress is $-\sigma_w$. Ignoring the minor complication that a real cable cannot withstand maximum static shear and maximum tangential velocity simultaneously, the maximum power passing through the cable is approximated by $P_{\text{DC}} = \pi r_{\text{cable}}^2 \sigma_w v_{\text{burst}}$, so:

$$P_{\text{DC}} = \frac{\pi r_{\text{cable}}^2 \sigma_w^{3/2}}{\rho^{1/2}} \quad \text{(watts)} \quad \text{(Eqn. 6.48)}$$

limiting maximum cable operating frequency to $-v_{\text{burst}} / 2 \pi r_{\text{cable}}$ or:

$$v_{\text{DC}} \leq \left( \frac{\sigma_w}{2 \rho^{1/2} r_{\text{cable}}} \right)^{1/2} \quad \text{(Hz)} \quad \text{(Eqn. 6.49)}$$

Hence $P_{\text{DC}} / P_{\text{AC}} = (G / \sigma_w)^{1/2} - 7$, taking $G = 5 \times 10^{11}$ N/m$^2$ and $\sigma_w = 10^{10}$ N/m$^2$ for a diamondoid cable, so the initial conclusion is that the DC strategy allows transmission of -1 order of magnitude higher mechanical power than the AC strategy, for a given cable size and material. A cable, secured at one end and twisted through a maximum angle $\theta_{\text{max}}$ at the other end, acquires at maximum stress an energy $E_{\text{cable}} = (1/2) k_{\text{ torsion}} \theta_{\text{max}}^2$, where $k_{\text{ torsion}} = \pi G r_{\text{cable}}^4 / 2 L_{\text{cable}}$; setting $E_{\text{cable}} = \pi r_{\text{cable}}^2 L_{\text{cable}} DE$ gives $\theta_{\text{max}} = 2 L_{\text{cable}} \sigma_w / r_{\text{cable}} G - 4$ radians for an $L_{\text{cable}}$ = 1 micron and $r_{\text{cable}}$ = 10 nm diamondoid cable.

The AC cable may suffer transmission losses due to shear wave radiation from an oscillating torque. However, thermal safety of in vivo mechanical power tethers is also of paramount concern. Under severe braking, loading or jamming of a cable, significant heat may be released. For a cable in vivo, a temperature rise of 63 K at the outer jacket surface (e.g., in contact with body fluids) boils water. Even a $\Delta T > 10$ K may constitute an unacceptable risk in a conservative nanomedical design, as this is more than enough to trigger biological responses, for example by heat-shock proteins (HSPs). (Activation of HSPs in some cases may be beneficial to human health, but conservative nanomedical design requires minimizing all such unplanned side-effects.)

In the most optimistic scenario, power is immediately shut off the instant a fault is detected, instantly converting a DC cable into a half-cycle AC cable under relaxation. To prevent stored energy from causing damage in the event of a heat pulse, a DC cable should never be operated faster than the fastest physically equivalent AC cable, e.g., $v_{\text{DC}} \leq v_{\text{ACm}}$.

In the most pessimistic scenario, the cable jams at a single pointlike defect and then radiates the entire power flux from a sphere of radius $-r_{\text{cable}}$; the practical effect is to preclude power cables altogether, because total cable power could not be allowed to exceed the heat flux from a droplet of water of radius $-r_{\text{cable}}$ whose temperature had been raised by $\Delta T$.

An intermediate, yet conservative, scenario allows that the entire power flux (not just the energy already stored in the cable) must be dissipated, but that the whole length of the cable may be employed as the radiator during a dissipative event—as, for example, if power continues to be transmitted after the fault but the cable jacket retains physical integrity. In this scenario, a cable carrying power flux $I_{\text{power}} \left(\text{W/m}^2\right)$ through a medium of heat capacity $C_V$ and thermal conductivity $K_t$ overheats the immediate environment in a time $t_{\text{overheat}} = L_{\text{cable}} C_V / I_{\text{power}}$, but thermal equilibration time $t_{\text{EQ}} = X_{\text{bio}} / K_t$, where $X_{\text{bio}}$ is a characteristic thermal conduction path length (e.g., minimum-size biological elements of size $X_{\text{bio}}$ = 1 micron). Requiring for safety that $t_{\text{overheat}} > t_{\text{EQ}}$, then $I_{\text{power}} < (K_t L_{\text{cable}} / t_{\text{EQ}})^2$, which defines considerably more restrictive operating frequency limits for cables.

For an AC cable in the intermediate thermal scenario, the energy density per cycle is $-(1/2) D E$; giving at maximum frequency $v_{\text{AC}}$ a power flux of $I_{\text{AC}} = v_{\text{AC}} L_{\text{cable}} \sigma_w w^2 / 2 G$, hence:

$$v_{\text{AC}} \leq \frac{2 G K_t \Delta T}{\sigma_w X_{\text{bio}}^2} \quad \text{(Hz)} \quad \text{(Eqn. 6.50)}$$

For a DC cable in the intermediate thermal scenario, power flux is $I_{\text{DC}} = \sigma_w v_{\text{burst}} = 2 \pi r_{\text{cable}} \sigma_w v_{\text{burst}}$, hence:

$$v_{\text{DC}} \leq \frac{L_{\text{cable}} K_t \Delta T}{2 \pi r_{\text{cable}} \sigma_w X_{\text{bio}}^2} \quad \text{(Hz)} \quad \text{(Eqn. 6.51)}$$

Adopting the intermediate scenario, the general conclusion is that very small cables tend to be thermally limited, while very large cables are both thermally and mechanically limited. In the following configurations, we assume a diamondoid cable with $\Delta T = 10$ K, $\rho = 3510$ kg/m$^3$ for diamond, $K_t = 0.623$ W/m$^K$ for water at 310 K, and $X_{\text{bio}} = 1$ micron.

An AC cable carrying a power throughput of $\pi r_{\text{cable}}^2 I_{\text{AC}} = 1000$ pW (near-peak nanorobot requirement) with $r_{\text{cable}} = 5$ nm and $L_{\text{cable}} = 2$ microns is thermally limited to an operating frequency of $v_{\text{AC}} \leq 60$ KHz ($< v_{\text{ACm}} = 6$ GHz), giving a power density of $I_{\text{AC}} / L_{\text{cable}} = 6 \times 10^{12}$ W/m$^2$. The same cable in DC mode is thermally limited to an operating frequency of $v_{\text{DC}} \leq 40$ KHz ($< v_{\text{DCm}} = 70$ GHz) with the same power density. If thermal limits...
were ignored, the cable could carry 0.2 milliwatts in AC mode and 1 milliwatt in DC mode.

Similarly, an AC cable carrying a power throughput of \( \pi r_{cable}^2 \)

\( I_{AC} = 100 \text{ watts (human basal requirement)} \) with \( r_{cable} = 4 \text{ microns} \) and \( L_{cable} = 0.4 \text{ meter} \) is mechanically limited to an operating frequency of \( v_{ACm} \approx 30 \text{ KHz} \) (< \( v_{ACm} < 60 \text{ KHz} \)), again giving a power density of \( \frac{I_{AC}}{L_{cable}} = 6 \times 10^6 \text{ W/m}^3 \). The same cable in DC mode is thermally limited to an operating frequency of \( v_{DCm} \approx 10 \text{ MHz} \) (< \( v_{DCm} < 100 \text{ MHz} \)) with the same power density. If thermal limits were ignored, the cable would still carry only 100 watts in AC mode but could carry up to 700 watts in DC mode.

The refractive index profile of a sufficiently wide (Section 6.4.3.2) rotating diamond cable could be engineered, and might possibly be made sufficiently transparent, to also act as a photonic channel simultaneously conveying optical power or communications signals.

6.4.3.5 Chemical Tethers

Chemical energy may also be delivered by tether at power densities > \( 10^9 \text{ watts/m}^3 \). Single tethers could bring ATP, glucose, or synthetic energy storage molecules (Section 6.2.3) directly to nanodevices in vivo. Double-tether cables could carry two hypergolic fuels, or a separate fuel and oxidizer, in a paired cable under pressure with the components mixed on site to release their stored energy. At 1000-10,000 atm pressures, most organic substances should undergo a hypergolic (spontaneous) reaction with oxygen piped in at equivalent pressure; safety would be a primary concern with this approach. Chemical tethers may also be used to resupply macroscopic energy organs (Section 6.4.4), or to exchange chemicals with an artificial nanoliver or other chemical modulating, processing, or synthesizing implanted nano-organisms.

6.4.3.6 Power Tether Configurations In Vivo

If an application requires only a very small number of nanodevices operating in a very limited tissue volume for a brief period of time, then direct tethered power supply might be preferred. Tethered devices may employ simpler onboard energy conversion systems, hence may be appropriate for devices of more primitive design such as may exist in the early years of nanomedical technology development. Tethered power systems could be useful when working inside a tissue volume that has been drained of indigenous useful chemical energy resources and which has no bloodflow or active homeostatic processes, such as amputated limbs, massively ischemic organs, or cryogenically preserved tissues or bodies. Tethered power might also be acceptable for operations inside tissues with incomplete, non-existent, or low-density vascularization, such as the epidermis, or inside tissues with incompetent or compromised immune systems, such as an embryo or fetus. In cyto slaved nanodevices may transmit/receive power or data to/from an extracellular master nanorobot via transmembrane tethers (Section 9.4.5.6).

The two principal drawbacks to tethered supplies are physical vulnerability and tissue irritation (see also Section 7.3.3). Tethers may kink, break, or become tangled; failures in rotating cable tethers may include "drilling" and "weed-whacker" modes, chemical tethers can detach and leak, and so forth. If billions or trillions of medical nanodevices are deployed, these vulnerabilities are enhanced. In addition, tether volume may equal or exceed nanorobot volume, increasing intrusiveness and reducing nanorobot mobility. For instance, a 10 cm long tether measuring 20 nm in diameter that might be used to power a -1 micron³ device has a volume of ~30 micron³ - a clear case of the tail wagging the dog. Limb/organ macromotions and tissue micromovements may produce a microscale sawing action of tether fibers against tissues, causing irritation, inflammation (including leukokine release with macrophage and fibroblast mobilization), and possibly a granulomatous reaction unless tethers can be made with variable compliance to match the stiffness of the tissues through which they pass, adding greatly to design complexity. Mobile tethers deployed inside cells can mechanically excite cytoskeletal elements, eventually triggering unwanted gene expression cascades in the nucleus. Tethers deployed through blood vessel walls or cell membranes may face similar challenges regardless of the chemical biocompatibility of their exterior or sheathing surfaces.

The simplest tethered configuration for powering large numbers of nanorobots is to transmit external power into an energy organ (Section 6.4.4) using a single tether connection. A complete analysis of networks and configurations of more complex tethered nanorobot power supplies is beyond the scope of this book. However, several simple configurations include Cartesian, fractal, and hub systems. A Cartesian system is laid out along a regular gridwork using a rectangular, cylindrical, spherical, helical, or other coordinate system as appropriate to the biotopology. A fractal system resembles a branching vascular-like "power tree" inside the body, with tendrils following blood or lymph vessels and natural tissue graining. A hub system (aka "mother ship") employs an array of isolated distribution nodes, with numerous individual nanorobots directly connected to one or another node.

Tethers must be directed entirely through tissue; placement in the active bloodflow may precipitate prompt thrombosis from microturbulence shear forces or netting action with formed elements. Macrophagic responses characteristic of wound repair (Chapter 24) would likely be stimulated within hours of tissue entry as cellular shear force sensors activate cytokine signaling mechanisms in response to mechanical stresses. To avoid massive granulomatous reaction, tethers cannot be abandoned after use and probably must be spooled out during deployment and respooled during retraction, adding additional potential failure modes and increasing the time required for deployment or retraction, a major drawback in emergency situations.

It might be possible to carefully install a biocompatible electrical or fiberoptic power network throughout a patient’s body (Section 7.3.1) as a permanent augmentation without causing irreparable damage. The presence of great numbers of isolated current-carrying wires creates numerous stray electrical fields that could affect cellular processes and possibly stimulate a macrophagic response; however, electric fields are minimal outside a well-shielded 1-micron-diameter coax or twisted-pair triax cable. If ~1 micron strand-free cables or optical fibers are used, a total network volume of ~10 cm³ of fiber (~0.01% of human body volume) has ~13,000 kilometers of length; assuming 1-meter strands and a simple Cartesian distribution, fibers have a mean lateral separation of ~80 microns between nearest neighbors (~4 cell widths).

6.4.4 Dedicated Energy Organs

Delivery of energy to in vivo nanorobots via tethers necessarily involves a hierarchical distribution network of some kind, ranging from maximum span (e.g., maximally horizontal hierarchy: all units directly connected to the external source) to minimum span (e.g., maximally vertical hierarchy: all units connected in series). Many intermediate network topologies are readily imaginable. Networks typically are most efficient at some optimum mix of spans and levels. To support this optimum mix, dedicated energy storage organs positioned at network nodes can add stability and longevity to the system and provide load buffering. Energy organs may act as distribution or routing devices, simply passing power down the line, or they may act as output nodes which are periodically encountered...
or revisited by in vivo nanorobots to recharge or refuel depleted onboard energy stores.

For example, we may envision a permanent macroscopic implanted device with a transdermal connector port. Through this port, the user recharges the internal energy organ by plugging in a source of compressed gas, by connecting through the port to a battery or a household electrical wall outlet, or by attaching an antenna system to facilitate the receipt of acoustic, mechanical, inductive, rf, infrared, or optical power signals. Chemical tethers could also resupply these energy organs: Assuming Poiseuille flow (Eqn. 9.25), ~40 mm^3/sec of stoichiometric oxyglucose fuel supply at 1000 atm pressure differential flowing (separately) through a double-tether -1 meter in length and ~60 microns in diameter supplies the entire 100-watt human basal power requirement assuming 50% chemical energy conversion efficiency at the receiving end, with the glucose supplied in a saturated 70% aqueous solution at 310 K. Chemical tether power density - 4 x 10^10 watts/m^3.

Inside the body, components of the implant may protrude into tissues or the bloodstream to supply power to internal nanorobots. These protrusions could act in broadcast mode, emitting acoustic or electrical power intracorporeally. Circumferential current-carrying loops could energize bloodborne nanorobots as they passed through the vascular solenoid, although energy transfer efficiency may be low (Section 6.4.2) and possible biological effects of these electric fields on cellular systems should be studied further. Energy organ protrusions could release manufactured artificial energy-rich molecules (Section 6.2.3) or microscale fuel tankers containing energetic compounds directly into the bloodstream. Protrusions might act as “energy teats” to which nanorobots could dock and refuel or recharge, although this will require navigational beacons, docking mechanisms, transfer conduits, etc., thus increasing system complexity. Energy organ protrusions could also serve as catalytic or conversion nodes, transducing one form of energy into another, or as manufacturing nodes, converting less energetic chemicals into more energetic chemicals through the application of externally supplied energy. In all these cases, biocompatibility is a critical issue that must be fully addressed (Chapter 15).

Dedicated energy organs may range in size from hundreds of microns up to several centimeters, depending on the position in the distribution hierarchy and the task to be performed, and may themselves be recharged from other larger energy organs positioned higher up in the hierarchy, or directly from external sources. Given the maximum conventional energy storage density of ~10^{11} joules/m^3 (Section 6.2), the human basal energy requirement of ~2000 sec of free glucose energy available solely from internal stores without requiring external resupply, assuming ~50% energy conversion efficiency. Biological power densities range from 10^2-10^6 watts/m^3, with the typical tissue cell operating at about 3800 watts/m^3. The active flight muscle of the honeybee has the highest known power density in biology (3.4 million watts/m^3).

6.5 Design Energetics Assessment

6.5.1 Power in Biological Cells

Table 6.8 summarizes the total power outputs and power densities of various cells in the human body, adapted from numerous experimental studies on tissue metabolism using assumed cell volumes as listed. Relatively inactive red cells produce only ~0.01 pW, while thermogenic fat cells may release 8 million times more energy during shivering. Muscle cells and platelets have very low basal rates, but are capable of brief bursts of high power density. The typical human tissue cell operates at ~30 pW and at this power level has ~2000 sec of free glucose energy available solely from internal stores without requiring external resupply, assuming ~50% energy conversion efficiency. Biological power densities range from 10^2-10^6 watts/m^3, with the typical tissue cell operating at about 3800 watts/m^3. The active flight muscle of the honeybee has the highest known power density in biology (3.4 million watts/m^3).

6.5.2 Thermogenic Limits In Vivo

Consider nanorobots at work inside a single biological cell of volume V_{cell}, and thermal conductivity K_{t} = 0.623 watts/m-K (water) immersed in a large isothermal heat sink. The maximum temperature differential ΔT_n between core and periphery caused by the intracellular release of P_{cell} (watts) due to nanorobot activity is:

\[ ΔT_n = \frac{P_{cell}}{6L_{cell}K_{t}} = \frac{D_{cell} L_{cell}^2}{6K_{t}} \text{ (kelvins)} \]  

where D_{cell} is total cellular energy density (watts/m^3). A typical ~20 micron diameter human tissue cell produces 30-480 pW, giving ΔT_{cell} = 0.3-6.0 microkelvins. The data in Table 6.8 indicates that individual cells appear to safely tolerate internal energy releases up to D_{cell} = 10^6 watts/m^3, which leads to the following proposed thermogenic limit for in cyto operations:

\[ P_{cell} ≤ 10^6 V_{cell} \text{ (watts)} \]  

Thus, for example, using the proposed limit an isolated 20-micron tissue cell can safely host in its interior a set of active nanorobotic machinery that in total generates P_{cell} = 8000 pW and raises cell temperature by ΔT_{cell} = 100 microkelvins, which seems extremely conservative.

If many cells in a given block of tissue contain active nanomachinery, as will often be the case, then the assumption of an isothermal heat sink no longer holds because the nanorobots will collectively warm up their operating environment. In the limit of an entire human body, a 100-watt total nanorobot power expenditure giving a whole-body 1000 watt/m^3 incremental power density (Table 6.8) seems quite conservative since 50,000-100,000 watts/m^3 is considered the safe therapeutic range for medical diathermy. A temperature rise in the tissues triggers an increase in blood flow, which in turn produces a cooling effect. The maximum thermal load that an adult can dissipate under room temperature conditions without an increase in body temperature is a whole-body exposure of 100 watts/m^3; thus a 100-watt whole-body nanorobot power budget normally will produce no measurable increase in body temperature.
Operated at its peak output of 1600 watts, the human body’s core temperature rises ~3.5 K. Even assuming a simple linear relation, a whole-body 100-watt nanorobot power budget would correspond to an increase in core body temperature of at most ~0.2 K, far smaller than the mean normal diurnal variation of ~1 K (Section 6.3.1) and less than the ~0.2-0.5 K load error in the human thermoregulatory control system. Since glucose-powered nanorobots may compete with tissues for fuel, a 100-watt total withdrawal also represents ~2000 Kcal/day, near the limit of what can be conveniently replaced dietetically on a long-term basis.

A proper comprehensive analysis of energy flows requires a detailed model of body geometry, incident radiation, skin emissivity, convection currents, heat exchange by conduction, evaporative cooling, respiratory heat losses, physiological and behavioral thermoregulatory responses, work done by the patient, and the precise distribution, clumping patterns and activities of in vivo nanorobots, all of which is very complex and quite beyond the scope of this book. However, a simple log-linear interpolative relation with correct endpoints ($D_{tiss} = 10^6$ watts/m³ at $L_{tiss} = 20$ microns and $D_{tiss} = 10^3$ watts/m³ at $L_{tiss} = 0.5$ meters) suggests the following proposed crude

### Table 6.8 Estimated Power Output and Measured Power Density of Biological Cells and Human Tissues

<table>
<thead>
<tr>
<th>Organelle, Cell or Object</th>
<th>Estimated Power Output (picowatts)</th>
<th>Assumed Volume (micron³)</th>
<th>Measured Power Density (watts/m³)</th>
<th>Compiled, Computed or Estimated from: (References)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Myosin muscle motor</td>
<td>0.000001</td>
<td>$5 \times 10^{-7}$</td>
<td>$2.0 \times 10^6$</td>
<td>578</td>
</tr>
<tr>
<td>Bacterial flagellar motor</td>
<td>0.0001</td>
<td>$5 \times 10^{-8}$</td>
<td>$2.0 \times 10^6$</td>
<td>581</td>
</tr>
<tr>
<td><em>S. faecalis</em> bacterium (basal)</td>
<td>0.00035</td>
<td>0.2</td>
<td>$1.8 \times 10^6$</td>
<td>798</td>
</tr>
<tr>
<td>Platelet (resting)</td>
<td>0.003-0.09</td>
<td>3</td>
<td>$0.1-3.0 \times 10^6$</td>
<td>791,792,796</td>
</tr>
<tr>
<td>Red blood cell</td>
<td>0.008</td>
<td>94</td>
<td>$8.5 \times 10^5$</td>
<td>789</td>
</tr>
<tr>
<td><em>E. coli</em> bacterium (basal)</td>
<td>0.05</td>
<td>2</td>
<td>$2.5 \times 10^5$</td>
<td>797</td>
</tr>
<tr>
<td>Mitochondrion organelle</td>
<td>0.1-1.1</td>
<td>1</td>
<td>$0.1-1.1 \times 10^5$</td>
<td>781,786</td>
</tr>
<tr>
<td><em>S. faecalis</em> bacterium (maximum growth)</td>
<td>0.23</td>
<td>0.2</td>
<td>$1.2 \times 10^6$</td>
<td>798</td>
</tr>
<tr>
<td>Chordocyte</td>
<td>0.3</td>
<td>670</td>
<td>$4.0 \times 10^5$</td>
<td>744</td>
</tr>
<tr>
<td>Platelet (activated)</td>
<td>0.7-7.0</td>
<td>3</td>
<td>$0.2-2.3 \times 10^6$</td>
<td>790</td>
</tr>
<tr>
<td>Skin cell</td>
<td>1-3</td>
<td>1000</td>
<td>$1.0-3.1 \times 10^3$</td>
<td>744,783,794,818</td>
</tr>
<tr>
<td>Skeletal muscle cell (resting)</td>
<td>1-10</td>
<td>2000</td>
<td>$0.5-4.9 \times 10^3$</td>
<td>526,744,818</td>
</tr>
<tr>
<td>Neutrophil leukocyte</td>
<td>2-9</td>
<td>450</td>
<td>$0.4-2.0 \times 10^4$</td>
<td>753,792,796</td>
</tr>
<tr>
<td>Osteocyte (bone)</td>
<td>2-38</td>
<td>30,000</td>
<td>$0.08-1.3 \times 10^3$</td>
<td>744,818</td>
</tr>
<tr>
<td>Pancreatic cell</td>
<td>9</td>
<td>1000</td>
<td>$8.9 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>Relaxed resting cardiac muscle cell</td>
<td>16</td>
<td>8000</td>
<td>$2.0 \times 10^6$</td>
<td>788</td>
</tr>
<tr>
<td>Diaphragm muscle cell</td>
<td>20</td>
<td>2000</td>
<td>$1.0 \times 10^6$</td>
<td>744</td>
</tr>
<tr>
<td>Liver cell (hepatocyte)</td>
<td>45-115</td>
<td>6400</td>
<td>$0.7-1.8 \times 10^4$</td>
<td>744,745,782,783,818</td>
</tr>
<tr>
<td><strong>Typical Tissue Cell (basal)</strong></td>
<td>30</td>
<td>8000</td>
<td>$3.8 \times 10^6$</td>
<td>744,745,782</td>
</tr>
<tr>
<td>T-Cell lymphocyte (basal)</td>
<td>40</td>
<td>200</td>
<td>$2.0 \times 10^5$</td>
<td>470,777</td>
</tr>
<tr>
<td>Bone marrow cell</td>
<td>45</td>
<td>8000</td>
<td>$5.6 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>Intestine/Stomach cell</td>
<td>46-52</td>
<td>8000</td>
<td>$5.7-6.5 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>Testicular cell</td>
<td>46-149</td>
<td>8000</td>
<td>$0.6-1.9 \times 10^6$</td>
<td>745</td>
</tr>
<tr>
<td>Lung cell</td>
<td>56-78</td>
<td>8000</td>
<td>$7.1-9.7 \times 10^5$</td>
<td>744,745,782</td>
</tr>
<tr>
<td>Brown fat cell (resting)</td>
<td>60</td>
<td>200,000</td>
<td>$3.0 \times 10^6$</td>
<td>745,818</td>
</tr>
<tr>
<td>Spleen cell</td>
<td>60-80</td>
<td>8000</td>
<td>$7.5-9.9 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>Neuron cell (basal)</td>
<td>70-110</td>
<td>14,000</td>
<td>$5.0-7.9 \times 10^5$</td>
<td>744,745,747,779,782</td>
</tr>
<tr>
<td>Thymus cell</td>
<td>74</td>
<td>8000</td>
<td>$9.3 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>Heart muscle cell (typical)</td>
<td>87-290</td>
<td>8000</td>
<td>$1.1-3.6 \times 10^6$</td>
<td>744,782,783,795,818</td>
</tr>
<tr>
<td>Skeletal muscle cell (max., voluntary)</td>
<td>111</td>
<td>2000</td>
<td>$5.7 \times 10^6$</td>
<td>818</td>
</tr>
<tr>
<td>Thyroid cell</td>
<td>130</td>
<td>8000</td>
<td>$1.6 \times 10^5$</td>
<td>745</td>
</tr>
<tr>
<td>T-Cell lymphocyte (antigen response)</td>
<td>130</td>
<td>200</td>
<td>$6.5 \times 10^5$</td>
<td>470,777</td>
</tr>
<tr>
<td>Adrenal cell</td>
<td>150</td>
<td>8000</td>
<td>$1.9 \times 10^6$</td>
<td>745</td>
</tr>
<tr>
<td>Kidney cell</td>
<td>155-346</td>
<td>8000</td>
<td>$1.9-4.3 \times 10^6$</td>
<td>7 4 4 , 7 4 5 , 7 8 2 , 8 1 8</td>
</tr>
<tr>
<td>Neuron cell (maximum)</td>
<td>255-330</td>
<td>14,000</td>
<td>$1.8-2.4 \times 10^6$</td>
<td>746,779</td>
</tr>
<tr>
<td><strong>Typical Tissue Cell (maximum)</strong></td>
<td>480</td>
<td>8000</td>
<td>$6.0 \times 10^6$</td>
<td>745,780,782</td>
</tr>
<tr>
<td>Skeletal muscle cell (max., tetanic)</td>
<td>2300</td>
<td>2000</td>
<td>$1.2 \times 10^6$</td>
<td>526</td>
</tr>
<tr>
<td>Honeybee flight muscle cell</td>
<td>3400</td>
<td>1000</td>
<td>$3.4 \times 10^6$</td>
<td>786</td>
</tr>
<tr>
<td>Heart muscle cell (maximum)</td>
<td>3500-5000</td>
<td>8000</td>
<td>$4.6-6.3 \times 10^5$</td>
<td>783,787</td>
</tr>
<tr>
<td>Pancreatic islet (multi-cell)</td>
<td>50,000-90,000</td>
<td>8,000,000</td>
<td>$0.6-1.1 \times 10^6$</td>
<td>793</td>
</tr>
<tr>
<td>Brown fat cell (thermogenic)</td>
<td>64,000</td>
<td>200,000</td>
<td>$3.2 \times 10^5$</td>
<td>786</td>
</tr>
<tr>
<td>Human brain</td>
<td>15-25 watts</td>
<td>$1.4 \times 10^3$ m³</td>
<td>$1.1-1.8 \times 10^4$</td>
<td>746,774,786,875</td>
</tr>
<tr>
<td>Human body (basal)</td>
<td>100 watts</td>
<td>0.1 m³</td>
<td>$1.0 \times 10^5$</td>
<td>780</td>
</tr>
<tr>
<td>Human body (maximum)</td>
<td>1600 watts</td>
<td>0.1 m³</td>
<td>$1.6 \times 10^4$</td>
<td>780</td>
</tr>
<tr>
<td>Gasoline-powered automobile</td>
<td>200,000 watts</td>
<td>10 m³</td>
<td>$2.0 \times 10^5$</td>
<td>Author’s 1969 Oldsmobile</td>
</tr>
<tr>
<td>The Sun</td>
<td>$3.92 \times 10^6$ watts</td>
<td>$1.41 \times 10^7$ m³</td>
<td>0.28</td>
<td>1662</td>
</tr>
</tbody>
</table>
maximum thermogenic limits for nanorobot operations in human soft tissue:

\[ P_{\text{tiss}} \leq V_{\text{tiss}} \frac{D_{\text{tiss}}}{1800} \text{ (watts)} \]  
\[ D_{\text{tiss}} = 10^{[2.77 - 0.687 \log_{10} (L_{\text{tiss}})]} \text{ (watts/m}^3\text{)} \]  
\[ \Delta T_{\text{tiss}} = \frac{P_{\text{tiss}}}{1800 L_{\text{tiss}}^{3/2}} \text{ (kelvins)} \]

where \( V_{\text{tiss}} = L_{\text{tiss}}^3 \) for the total volume of soft tissue in which all deployed nanorobots are active. (Eqns. 6.52 through 6.56 are summarized in Figs. 6.14 and 6.15). Thus, for example, medical nanorobots deployed solely in the thyroid gland (\( L_{\text{tiss}} \sim 2.6 \text{ cm}; \) Table 8.9) should be restricted to a maximum total power budget \( P_{\text{tiss}} \sim 0.1 \text{ watt} \) which elevates thyroid temperature by \( \Delta T_{\text{tiss}} \sim 0.02 \text{ K} \). Higher in vivo power densities may sometimes be justified in hospital or emergency medical situations (Chapter 24), and a few other special circumstances (e.g., Chapter 21).

### 6.5.3 Nanorobot Power Scaling

As an initial, crude order-of-magnitude estimate of nanorobot power consumption, allometric scaling of metabolism in biology for whole organisms follows a 3/4 power law. Normalizing to \( P = 100 \text{ watts} \) for an \( m = 70 \text{ kg} \) human body mass and assuming water density for nanorobots, then \( P = (4.13) m^{3/4} = 23 \text{ pW} \) for a 1 micron\(^3\) nanorobot, a power density of \( D_{n} \sim 2 \times 10^7 \text{ watts/m}^3\).

Surface power intensity considerations drive maximum nanorobot onboard power density. For instance, the maximum safe intensity for ultrasound is 100-1000 watts/m\(^2\) (pain threshold for human hearing \( \sim 100 \text{ watts/m}^2 \) (Section 6.4.1) and the conservative maximum safe electromagnetic intensity is also \( \sim 100 \text{ watts/m}^2 \) (Section 6.4.2). Additionally, a surface at 373 K (boiling water) relative to a 310 K environment radiates 100-1000 watts/m\(^2\) at emissivities \( e_r \sim 0.1-1 \) (Eqn. 6-19). Conservatively taking 100 watts/m\(^2\) as the maximum safe energy flux across the surface of a spherical nanorobot of radius \( r_n \), then the maximum nanorobot power density is

\[ D_{n} = \frac{300}{r_n} - 10^9 \text{ (watts/m}^3\text{)} \]

for a device -1 micron in diameter—the largest safe whole-device power density that should be developed in vivo. A \( 10^9 \text{ watt/m}^3 \) maximum implies a \( \sim 1000 \text{ pW} \) limit for 1 micron\(^3\) nanorobots. Interestingly, the energy dissipation rate required to disrupt the plasma membrane of \(-95\%\) of all animal cells transported in forced turbulent capillary flows is on the order of \(-10^9 \text{ to } 10^{10} \text{ watts/m}^3\).

For chemically powered foraging nanorobots, another fundamental constraint on power density is imposed by diffusion limits on fuel molecules. From Eqn. 3.4, the maximum diffusion current of glucose molecules of energy content \( E_{\text{fuel}} = E_{\text{glu}} (4765 \text{ zJ}) \); Egn. 6.16, burned at efficiency \( e% \sim 0.50 \) (50\%) in oxygen, to the surface of a spherical nanorobot of radius \( r_n \) will support a maximum onboard power density of:

\[ \frac{D_{\text{fuel}}} {e% E_{\text{fuel}} C} = \frac{D_{\text{diff}}}{r_n^2} \text{ (watts/m}^3\text{)} \]

For \( D = 7.1 \times 10^{-10} \text{ m}^2/\text{sec} \) for glucose in water at 310 K (Table 3.3), \( C = (0.67) 3.5 \times 10^{24} \text{ molecules/m}^3 \) in (newborn and) adult

---

*Since oxyglucose foraging nanorobots are not seriously glucose-limited, there is little to be gained by enabling energy organs or cooperative nanorobot populations to secrete insulin/glucagon hormones (mimicking the pancreas), cortisol hormones, etc. to artificially manipulate serum glucose levels. Available oxygen also may be artificially manipulated using nanorobotic compressed gas dispensers or by other means (Chapter 22).

**Note that as nanodevices get smaller, their surface/volume ratio expands. Thus the Square/Cube law predicts that smaller nanodevices can admit more energy (e.g., chemical, acoustic, electromagnetic) through their surfaces per unit enclosed volume of working nanomachinery, hence can have higher power densities, as illustrated by Eqn. 6.58.
human blood plasma (Appendix B), and \( r_m = 0.5 \) micron, \( D_{\text{diff}} = 7 \times 10^{-9} \) watts/m³. However, because oxygen dissolves only slightly in blood plasma and interstitial fluid, the oxyglucose engine is more severely diffusion-limited by its oxygen requirements than by its glucose requirements. Applying Eqn. 6.58 and using (for oxygen) \( D = 2.0 \times 10^{-9} \) m²/sec (Table 3.3), \( C = 7.5 \times 10^{21} \) molecules/m³ in arterial blood plasma (Appendix B), \( E_{\text{glu}} = 6 \), and \( r_m = 0.5 \) micron, \( D_{\text{diff}} = 7 \times 10^{-9} \) watts/m³. Once again, \(-10^{9}\) watts/m³ appears to be a correct upper limit for whole nanorobots.***

The disposition of combustion byproducts, particularly \( CO_2 \), may provide another weak constraint on chemical systems. For example, a 10-pW glucose-burning nanorobot generates \(-10^{-9}\) molecules/sec of \( CO_2 \). Pressurized to 1000 atm, this production rate fills ~0.001 \( CO_2 \) concentration rises by \(-2 \times 10^{-7} \) M/sec, reaching 0.0003 M total power draw for a population of oxygen-unrestricted mean interdevice separation). Recall that the maximum diffusion-limited may develop \(-10 \) pW (roughly in line with biologically-derived devices (Chapter 21), but the typical simple micron-scale nanorobot for the largest and most sophisticated repair and defensive in vivo systems, it is natural to assume that electrical power would be the most practical choice for onboard energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

For \( e_r = 0.97 \) (e.g., carbon black) to maximize heat emission at the lowest possible temperature and \( T_{\text{environment}} = 300 \) K, then \( R_{\text{crit}} = 0.22 \) mm for \( f_n = 100 \). Assuming a full cold start, critical time to incineration is \( t_{\text{crit}} = C V (T_{\text{burn}} - T_{\text{environment}}) / f_n D_p = 1.4 \) sec for \( f_n = 100 \) if nanorobot heat capacity \( C_V = 1.8 \times 10^6 \) joules/m³-K (-diamondoid). Decreasing \( D_p \) to a more reasonable \(-10^7 \) watts/m³ or simply switching off 99% of the nanorobots \( (f_n = 1\%\) increases \( R_{\text{crit}} \) to \(-22 \) cm.

### 6.5.4 Selection of Principal Power Source

In selecting the ultimate source of power for a given application, the appropriate energy acquisition strategy must first be decided. Two overall strategies are readily distinguished: (1) nonrefuelable and (2) refuelable.

A nonrefuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.

A refuelable strategy implies that all nanorobot power is drawn from internal energy storage. Such a nanorobot simply stops when its internal energy stores are depleted. The relevant parameter is onboard energy storage density, which should be maximized. From Section 6.2, chemical energy storage provides the highest energy density, up to \(-10^3 \) joules/m³. (Nuclear storage has a much higher theoretical energy density, but in medically "safe" fissions the rate of energy withdrawal from the store cannot be precisely regulated using known technology, thus limiting the effective energy density of this potential source.) A 1 micron³ store of chemical energy at a conservative \(-10^{10}\) joules/m³ provides a 10 pW nanorobot for \(-10^{-3}\) sec, which may be sufficient in some applications.
energy of choice to drive nanomedical systems. However, it is clear that mechanically powered nanorobots are competitive because mechanical energy may be transmitted at very high efficiency over nanoscale distances. Fully mechanical motors, pumps, actuators, manipulators, and even computers have been designed.\textsuperscript{10}

The principal disadvantages of electrically-powered nanorobots include:

A. Bioelectric Interactions — Electrical systems can create stray fields which may activate bioelectric-based molecular recognition systems in biology. While electrical systems allow ready coupling with electrochemical systems in the body, stray fields also may provoke unintended electrokinetische interactions. For example, galvanic sources have been found to provoke a significant host reaction, including formation of layers of necrotic debris, free neutrophils, granulation tissue and complete fibrous connective tissue encapsulation of long-term implants.\textsuperscript{590,3512} Microelectrophoretic interactions with possible natural rf oscillations of cells\textsuperscript{683} could present additional complications. Stray high-frequency vibrations from purely mechanical systems may be less provocative.

B. Electrical Interference — Electrical nanorobots are susceptible to electrical interference from external sources such as rf or electric fields, EMP pulses, and stray fields from other in vivo electrical devices. (Digital cellular telephones have been reported interfering with implanted cardiac pacemakers, causing the pacemakers to speed up, slow down, or even turn off.)\textsuperscript{3495-3498} Cosmic rays can provoke arc discharges in systems operated at high electrical potential.

C. Thick Insulators — Very thick insulation is required to prevent electron leakage, especially serious at the smallest sizes where significant quantum mechanical tunneling can occur. Without careful design, the high conductivity of the in vivo medium* can cause sudden power loss, e.g., by “shorting out.”

D. Thick Wires — Relatively thick wires are needed to conduct significant power levels without overheating, although future room-temperature superconductors might reduce this disadvantage.

The principal advantages of electrically-powered nanorobots include:

E. Speed of Operation — Electronic field configurations have a velocity near the speed of light (c - 3 x 10\textsuperscript{8} m/sec), much faster than mechanical logic rods which may move as slow as 1 m/sec. Even the classical electron drift speed v\textsubscript{d} in a bulk metallic conductor at high current density I\textsubscript{d} may be faster than mechanical rod motions: v\textsubscript{d} = I\textsubscript{d}/n\textsubscript{e}e, where n\textsubscript{e} = 8.4 x 10\textsuperscript{24} electrons/m\textsuperscript{3} for copper and e = 1.6 x 10\textsuperscript{-19} coul/electron, giving v\textsubscript{d} ~ 100 m/sec at I\textsubscript{d} ~ 10\textsuperscript{10} amps/m\textsuperscript{2}, near the maximum current density. (A more typical I\textsubscript{d} ~ 10\textsuperscript{4} amps/m\textsuperscript{2} gives v\textsubscript{d} ~ 1 cm/sec.) And the mobility of an electron confined to a one-dimensional wire or patterned diamondoid channel can be much higher than in the bulk material, up to ~10\textsuperscript{6} m/sec or -0.3\textsuperscript{97,1098} Thus applications demanding the greatest speeds may require electrical systems.

F. Electromagnetic Coupling — Detection of photons and electrical and magnetic fields is more easily mediated by an electrical transmission device than a mechanical device. For instance, a photocell or rhodopsin antenna absorbs photons to directly create moving charges. This is probably more efficient than using a piezoelectric transducer backwards as a detector, but involves handling higher-energy packets of energy which is inherently riskier.

G. Transmission Attenuation — Electrical signals at sub-MHz frequencies are only modestly attenuated during passage through human tissue.

6.5.6 Power Analysis in Design

An important part of any nanodevice design exercise is a detailed assessment of power requirements and heat dissipation. Such an assessment might typically include the following elements:

A. Molecular Transport — Sorting rotors, internal transport mechanisms, sieving or nanocentrifugation, and receptor-based transport have specific power requirements (Chapter 3). Sorting rotors operating at full speed (~10\textsuperscript{5} rev/sec) dissipate ~10\textsuperscript{10} watts/m\textsuperscript{3}. However, it should be possible to boost efficiency in some continuous molecular exchange systems by recovering most of the sorting energy by compressing (or concentrating) one species using energy derived largely from the decompression (or deconcentration) of the other via differential gearing.

B. Chemical Transformations — Drexler\textsuperscript{10} estimates that the energy dissipation caused by chemical transformations involving carbon-rich materials is -1000 zJ/atm or -10\textsuperscript{8} joules/kg of final product using readily-envisioned irreversible methods in systems where low energy dissipation is not a design objective, but may in theory be as low as -1 zJ/atm or -10\textsuperscript{5} joules/kg “if one assumes the development of a set of mechanochemical processes capable of transforming feedstock molecules into complex product structures using only reliable, nearly reversible steps.” However, R. Merkle notes that essentially all current proposals are quite dissipative. Most interactions with biological molecules may also have relatively low energy densities because common cut-and-repair operations can involve only a small number of covalent bonds on a single macromolecule. For example, forgoing just one new 1000-zJ covalent bond on a single 500-residue protein molecule containing ~10,000 atoms has a net molecular transformation cost of only ~0.1 zJ/atm. Transformations involving noncovalent bonds can have even lower per-atom energy costs.

C. Mechanical Operations — As a crude rule, physically active nanocomponents such as rotating roller bearings generate ~10\textsuperscript{-5} watts/m\textsuperscript{3} of waste heat; inactive parts produce no waste heat. Larger mechanical assemblages such as nanomanipulator arms dissipate ~10\textsuperscript{-3} watts/m\textsuperscript{3} in normal continuous operation,\textsuperscript{10} or ~1 megawatt/kg for active diamondoid nanomachinery. Energy required for locomotion, actuation and manipulation is described in Chapter 9; energy requirements for shape-changing or other metamorphic activities must also be assessed (Chapter 5).

*The purest water (e.g., distilled and deionized) has a specific resistance of 2.5 x 10\textsuperscript{15} ohm-m.\textsuperscript{390} It is normally difficult to obtain and store water with resistivity exceeding ~10\textsuperscript{15} ohm-m because of the absorption of CO\textsubscript{2} and other air gases, and of alkali and other electrolytes leached from glassware; ordinary distilled water in equilibrium with air has resistivity ~1000 ohm-m.\textsuperscript{390} These values may be compared to 1.59 x 10\textsuperscript{17} ohm-m for silver at 293 K,\textsuperscript{353} 3.5 x 10\textsuperscript{15} ohm-m for amorphous carbon at 293 K,\textsuperscript{1662} 0.0893 ohm-m for 1M KCl in water at 298 K and 68.1 ohm-m for 0.001M KCl solution\textsuperscript{390} (human blood is ~0.15M NaCl), 10\textsuperscript{-5}-10\textsuperscript{-6} ohm-m for synthesized diamond (depending on the growth process used), and >10\textsuperscript{15} ohm-m dark resistivities in natural diamond.\textsuperscript{137}
D. Communication and Navigation — There must be a design assessment of the energy required, if any, for internal navigation (Chapter 8) and for all communication tasks (Chapter 7) including intradevice and interdevice signaling, inmessaging and outmessaging, plus allocation of systemic overhead for communications and navigational networks.

E. Computation — Mechanical nanocomputers using rod logic in the design described by Drexler\textsuperscript{10} dissipate \(-10^{12}\) watts/m\(^3\) at -GHz clock rates, yielding \(-10^{10}\) MIPS/watt or \(-10^9\) instructions/sec per pW (\(-10^{28}\) instructions/sec-m\(^3\)). R. Merkle suggests that the use of nearly reversible computational operations\textsuperscript{73} may reduce energy requirements per instruction by at least two orders of magnitude, at the cost of a slower clock cycle. With standard logic, dissipative irreversible operations can approach a minimum energy dissipation of \(\ln(2) kT \approx 3 \text{ zJ}\) per operation at 310 K, and in principle with reversible logic a computer can dissipate arbitrarily less energy per logic operation (Section 10.2.4.1). Nevertheless, total nanorobotic power demand will often be dominated by computational energy requirements (Section 10.2).

F. Component Assembly — Fabricating nanoscale diamondoid parts using hydrogen-rich organic feedstock molecules and oxygen generates energy, because the process involves a controlled, but highly-exoergic, combustion reaction. In a well-known example given by Drexler,\textsuperscript{10} “burning” hydrocarbon to make diamondoid liberates 17 MJ/kg of gross energy, less 1.7 MJ/kg of local entropy decrease in the reactants, giving \(-15\) MJ/kg of free energy; assuming 3 MJ/kg dissipated in the mills, rotors, and computers needed to drive the manufacturing process, a net output of 12 MJ/kg of surplus energy is produced. However, subsequent higher-order assembly processes using this diamondoid material are endoergic. In the worst case, assembly of larger diamondoid structures from 1-nm cubes may cost \(-9\) MJ/kg, although with proper technique most of this energy can be recovered as mechanical work, reducing the cost of block assembly of larger structures to as low as \(-0.5\) to \(-1.0\) MJ/kg.\textsuperscript{10} Leaving aside the initial exoergic energy production, then depending upon computation, materials handling, and other requirements in a particular implementation, assembling prefabricated diamondoid part-like or module-like components into larger structures will probably cost \(-1.3\) MJ/kg of energy\textsuperscript{10} (Chapter 19).

6.5.7 Global Hypsithermal Limit

Finally, it is possible to derive a limit to the total planetary active nanorobot mass by considering the global energy balance. Total solar insolation received at the Earth's surface is \(-1.75 \times 10^{17}\) watts (\(\text{I}_\text{Earth} \approx 1370 \text{ W/m}^2 \pm 0.4\%\) at normal incidence\textsuperscript{887,1945}). Global energy consumption by mankind reached an estimated \(1.2 \times 10^{13}\) watts (\(-0.02\) W/m\(^2\)) in 1998. This latter figure may also be regarded as the total heat dissipation of all human technological civilization worldwide, as distinct from the \(-10^{12}\) watt metabolic output of the global human biomass. The evidence for global warming remains controversial,\textsuperscript{7} but it is clear that as the waste heat from human-built machinery continues to grow, the climate will begin to change. Technological heat plumes from asphalt-covered major cities already have demonstrable impact on local weather and thermal patterns. Leaving aside all considerations of changing concentrations of various atmospheric components (e.g., \(\text{CO}_2, \text{H}_2\text{O}, \text{O}_3\)), one might ask at what point anthropogenic energy releases could begin to seriously affect the global energy balance. (Climatologists sometimes call this the “hypsithermal limit.”) Global warming data remain inconclusive but certainly suggest that the present \(-10^{13}\) watts may lie within an order of magnitude of an important threshold. The power dissipation of all terrestrial vegetation is \(-10^{15}\) watts,\textsuperscript{*} and it is well-known that such vegetation plays a major role in the planetary energy equation. Climatologists have speculated that an anthropogenic release of \(-2 \times 10^{15}\) watts (-1% solar insolation) might cause the polar icecaps to melt. A high upper limit is provided by Venus, which receives \(-3.3 \times 10^{15}\) watts at its cloudtops, producing a surface temperature of \(-700\) K and a hellish \(-100\) atm sulfuric atmosphere—despite its close geological and size similarity to Earth. As a fairly liberal estimate,\textsuperscript{3100,3017} the maximum hypsithermal limit for Earth is taken to be \(-10^{17}\) watts.

The hypsithermal ecological limit in turn imposes a maximum power limit on the entire future global mass of active nanomachinery or “active nanomass.” Assuming the typical power density of active nanorobots is \(-10^{-9}\) W/m\(^3\), the hypsithermal limit implies a natural worldwide population limit of \(-10^7\) m\(^3\) of active functioning nanorobots, or \(-10^{14}\) kg at normal densities. Assuming the worldwide human population stabilizes near \(-10^{10}\) people in the 21st century and assuming a uniform distribution of nanotechnology, the above population limit would correspond to a per capita allocation of \(-10\) kg of active continuously-functioning nanorobots, or \(-10^{14}\) active nanorobots per person (assuming 1 micron\(^3\) nanorobots developing \(-10\) pW each, and ignoring non-active devices held in inventory). Whether a \(-10\)-liter per capita allocation \((-100\) KW/person) is sufficient for all medical, manufacturing, transportation and other speculative purposes is a matter of debate.

The hypsithermal active nanorobot population limit cannot be defeated by heroic artifices of planetary engineering such as giant Earth-orbiting sunscreens or planetary-scale heat pumps, which would devastate global photosynthetic activity or climate while not significantly reducing thermal energy density at the Earth's surface. Likewise, significantly reducing the atmospheric concentration of the most important natural greenhouse gases, especially \(\text{CO}_2\) and \(\text{H}_2\text{O}\), on a worldwide basis is not possible if the present ecology is to remain essentially undisturbed, but even the complete elimination of such gases would raise the hypsithermal limit by at most a factor of 10-20. Specifically, if Earth had no atmosphere at all, then to maintain the global mean surface temperature at its current value of \(T_{\text{Earth}} \approx 288\) K,\textsuperscript{3097} an additional nanoborgenic power of \(P_{\text{nanoim}} = \pi R_{\text{Earth}}^2 \times 4 \times 10^{-16}\) watts could be released at the Earth's surface, where \(R_{\text{Earth}} = 6.37 \times 10^6\) m (radius of Earth) and \(\sigma = 5.67 \times 10^{-8}\) W/m\(^2\)-K\(^4\) (Stefan-Boltzmann constant). Still more unrealistically, if Earth's surface were an atmosphere-free black body surrounded by a perfect mirror (e.g., \(\text{albedo} = 1.00\), vs. \(-0.31\) for natural Earth), an additional nanoborgenic power release (underneath the mirrored barrier) of \(P_{\text{mirror}} \leq 4 \pi R_{\text{Earth}}^2 \times (\pi T_{\text{Earth}}^4 - L_{\text{gol}}) \approx 2 \times 10^{17}\) watts would allow maintenance of the current mean global surface temperature \(T_{\text{Earth}}^\text{gol} \approx 0.05\) W/m\(^2\).\textsuperscript{3123} The mean geological heat flow at Earth's surface due to radioactive decay (e.g., \(U^{238}, K^{40}\) in the crust). Reductions in nanorobot waste heat made possible by low-dissipative molecular manufacturing and reversible computing may be offset by popular

\textsuperscript{*}Global photosynthetic fixation of \(\text{CO}_2\) by plants is \(-3.85 \times 10^9\) kg \(\text{CO}_2/\text{yr}, -46\%\) by ocean flora;\textsuperscript{3172} \(\text{CO}_2\) fixation requires \(-794\) zJ/molecule (Eqn. 6.16) or \(-1.1 \times 10^7\) J/kg \(\text{CO}_2\), giving a global vegetative total of \(-1.4 \times 10^9\) watts.

\textsuperscript{**}Note that a global population limit of \(-10^{26}\) 10-pW nanorobots represents only \(-100\) moles of active devices; for 1000-pW devices (\(-10^6\) W/m\(^3\) power density), the upper population limit is only \(-1\) mole of 1-micron\(^3\) active nanorobots worldwide. Facile talk of controlling localized “mole quantities” of nanodevices is thus highly misleading.
insistence on adherence to a more conservative global limit, perhaps \(-10^{15}\) watts (today's level), to better preserve the terrestrial habitat. Consistent with long-term ecological maintenance the hypsithermal limit ultimately can only be avoided by continuing human technological progress in space, which seems an excellent idea in any case.
**CHAPTER 7**

**Communication**

### 7.1 Nanorobot Communications Requirements

Communication is an important fundamental capability of medical nanorobots. At the most basic level, nanomachines must pass sensory and control data among internal subsystems to ensure stable and correct device operation. They must also exchange messages with biological cells, communicating with the human body at the molecular level. Nanodevices must be able to communicate with each other in order to:

1. coordinate complex, large-scale cooperative activities,
2. pass along relevant sensory, messaging, navigational, and other operational data, and
3. monitor collective task progress.

Finally, nanorobots must be able to receive messages from, and transmit messages to, both the human patient and external entities including antennas and telecommunication links, laboratory or bedside computers, and attending medical personnel.

It is instructive first to examine the gross information flows likely to be required in typical nanomedical situations. Nanorobot onboard computers are likely to generate from $10^4$ to $10^6$ bits/sec for the simplest systems to $10^9$ bits/sec for the most complex systems and tasks (Chapter 12), which sets broad limits on internal communications requirements. Assuming onboard data storage of $10^5$ to $10^9$ bits, rewriting an entire nanorobot memory in $1$ second demands information flows of $10^5$ to $10^9$ bits/sec. Communications with human cells may take place at many levels, ranging from 10-1000 bits/sec for individual neuronal impulses (Section 4.8.6) up to brief bursts at $10^6$ bits/sec ($\sim$MHz frequencies) for intracellular processes involving enzyme action or molecular gating. Inter-nanorobot communications are unlikely to require data transfer rates exceeding $10^5$ to $10^6$ bits/sec; explicit exchanges between nanodevices and the human user are restricted by the maximum data processing rate of the conscious mind, variously estimated as 1-1000 bits/sec (Chapter 25), although dermal and retinal displays may transfer visual information to the patient at up to $10^7$ bits/sec (Section 7.4.6). Communications with external entities may include monitoring or data transfer operations. For example, the $525 \times 390 = 204,750$ pixel/frame of standard halftone black-and-white broadcast television transmitted at 30 frames/sec with $1$ bit/pixel mandates a $-6.1 \times 10^6$ bit/sec (6.1 MHz) digital transfer rate (6.1 MHz transmission bandwidth); downloading an entire uncompressed human genome during cytogenetic repair operations (Chapter 20) in ~1000 sec requires a $-10^9$ bits/sec transfer rate.

According to classical information theory for channel capacity on a dissipative transmission line with additive equilibrium thermal noise, the minimum erasure energy required per transmitted bit is

$$E_{bit} \geq kT \ln(2) = 3 \, \text{zJ/bit} \quad \text{(Eqn. 7.1)}$$

where $k = 0.01381 \, \text{zJ/K}$ (Boltzmann constant), $T = 310 \, \text{K}$, and 1 zeptojoule ($\text{zJ}$) = $10^{-21} \, \text{joule}$. Thus the maximum $f_{max} = 10^9$ bits/sec bandwidth requirement noted earlier must draw $\geq 3$ picowatts ($\text{pW}$), well within the anticipated 1-1000 pW power budget of typical in vivo medical nanodevices (Section 6.5.3). Slower bit rates can draw even less power. The design challenge is to closely approach this minimum theoretical limit.

In this Chapter, an analysis of the most common nanorobotic communication modalities (Section 7.2) and communication network architectures (Section 7.3) is followed by a brief discussion of the many specific communication tasks to be performed (Section 7.4).

### 7.2 Communication Modalities

As a general principle, any method by which materials or power can be transferred into, around, or out of the human body also may be employed as a mode of communication by imposing a time-varying modulation on the flow. For in vivo communications, the leading candidates are free-tissue chemical, acoustic, and electromagnetic broadcast, nanomechanical and cable systems, and dedicated communicytes. An extensive treatment of data protocol systems is followed by a brief discussion of the many specific communication tasks to be performed (Section 7.4).

#### 7.2.1 Chemical Broadcast Communication

Chemical broadcast communication is widely used in the human body, as for example, in cellular communications using cytokines and other mechanisms (Section 7.4.5), hormones, neuropeptides, immune system components and pheromones. Consider the simple case of in vivo nanorobots attempting to communicate with each

---

* Although treated as approximately equal for simplicity in this book, the digital transfer rate is distinct from the transmission bandwidth. The bandwidth needed to handle a specific transfer rate is related to the actual coding and modulation technique selected. The simplest modulation techniques can indeed approach $-1$ bit/sec per Hz of bandwidth, but much more efficient techniques are available. For instance, a typical 1998-vintage PC modem used complex coding and modulation schemes to transmit 33 KB/sec over a voice-grade transmission line that was only 4 KHz wide, thus achieving a transfer rate of $-8$ bits/sec per Hz of bandwidth; $-100:1$ data compression algorithms are currently available for voice transmissions.

** Landauer points out that in theory there is no minimum energy required to transfer a bit; but see also Levitin and the discussion of reversible computing in Section 10.2.4.1.
other exclusively by chemical means. In this model, the sender releases a coded messenger molecule of some type into the cytosol, interstitial fluid, or bloodstream, and this molecule travels via diffusion and convective transport to its intended recipient, whereupon the messenger molecule is recognized, absorbed and decoded.

### 7.2.1.1 Ideal Messenger Molecule

The ideal chemical messenger molecule will have several important characteristics. First, its structure will contain a distinctive “head” or “flag” that permits easy recognition and binding by nanorobot molecular receptor systems such as molecular pumps or sorting rotors (Section 3.4), so that the entire message need not be read in order to identify the intended recipient. Second, it will be relatively bioinactive, thus not readily broken down by natural processes before the message is likely to be received. Third, it should be easily eliminated from the body, thus preventing potentially toxic accumulations; however, the molecule and any likely breakdown products should also be inherently nontoxic in anticipated maximum concentrations.* Finally, the molecule should be capable of easy extension to larger sizes, permitting significant entire messages to be written on the molecule, as the statistical nature of the transport process implies a relatively long time between assured detections of the messenger molecule.

One possible candidate messenger is the partially fluorinated polyethylene molecule originally suggested by Drexler for nanocomputer tape bulk memory systems, although others have been investigated. Such molecules can store one bit per carbon atom using an H atom to represent a “0” and an F atom to represent a “1” on one side of the carbon-chain backbone, with all H atoms occupying the other side of the backbone to facilitate easy reading. The message may be a single linear chain or may include branching structures representing conditionals or prioritizations embedded in the message. Assuming molecules (averaging 50%/50% H/F atoms) on the read end are of the form CH₃(CHX)ₓCH₃ (with X = H or F atoms) and can store 1 bit per unit (with n units/molecule), then message molecule density ρmessage = 1000 kg/m³ and message molecule volume Vmessage = (3.82 × 10⁻³⁰)ₙ(meter³), giving an information density of Dmessage = 26 bits/nm³ (3-3 bits/nm of linear message molecule length). (By comparison, linear DNA achieves ~1 bit/nm³.) Message molecule molecular weight MWmessage ≈ n MWunit (kg/mole) for n >> 1, where MWunit = 0.023 kg/mole for (CHX), ignoring the end caps and again assuming 50%/50% H/F atoms.

Hydrofluorocarbon message molecules with n > 20 may be strategically crosslinked in a standardized pattern across the all-H backbone and folded into a maximally compact spherical configuration in preparation for transmission. Alternatively, the polymer could be wound up onto a bobbin or reel, allowing more convenient reading in preparation for transmission. Alternatively, the polymer could be wound up onto a bobbin or reel, allowing more convenient reading.

## 4.2.2.5 Hydrofluorocarbon Delivery

Hydrofluorocarbon molecules are biocompatible? Fluorinated hydrocarbon perfusants with n < 20 such as Fluosol-DA or the commercial solvent polyfluoro-octobromide (Perflubron) are FDA-approved and have been used as reversible oxygen carriers in artificial blood formulations for years. Such fluorocarbons are characterized by high chemical and biochemical inertness, absence of metabolism, and rapid excretion. The rate of excretion has been shown to decrease exponentially with increasing molecular weight, with the exception of fluorocarbons containing a lipophilic substituent such as Br (e.g., Perflubron) or a hydrogenated fragment in their molecular structure. These water-insoluble compounds usually consist of 8-10 carbon atoms with molecular weights of ~450-500 daltons; clinical administrations typically reach maximum bloodstream concentrations of 70-400 gm/liter. (cf. LD50 ~ 700 gm/liter estimated from acute single-dose toxicity studies), many orders of magnitude in excess of the levels anticipated for nanodevice chemical communication applications. Interestingly, mice survive immersion in fluorocarbon through which oxygen is bubbled, and rats breathing 95% oxygen have survived total blood replacement with fluorocarbon fluid. The boiling point of perfluoropentane (C₅F₁₂) is 29.2°C, but is >37°C for n>5—e.g., 56.6°C for perfluorohexane (C₆F₁₄).

Fluorocarbons and fluorocarbon moieties have very strong intramolecular bonds and very weak intermolecular interactions, hence should display low particle aggregation. Fluorinated surfactants are less hemolytic and less detergent than their hydrocarbon counterparts, and fluorosurfactants appear unable to extract membrane proteins. The stability and permeability of fluorinated liposomes has been widely studied. For example, anionic double chain glycophosplolipids with either two hydrocarbon or two fluorocarbon chains, or a mixed double chain (one fluorinated, one hydrogenated), readily give vesicles 30-70 nm in diameter when dispersed in water, with maximum tolerated IV doses up to 0.5 gm/kg body weight in mice (~5 gm/liter blood volume); hemolytic activity sharply decreases with increasing degree of fluorination. Many fluorinated surfaces such as Teflon are simultaneously hydrophobic and lipophobic.

Fluorocarbons are typically administered in the form of emulsions of 0.1-0.2 micron droplets dispersed in a physiologic solution, similar to fat emulsions routinely used for parenteral nutrition. A compact folded messenger molecule of spherical radius rmessage can store approximately

\[
I_{\text{message}} = \frac{4}{3} \pi r_{\text{message}}^3 D_{\text{message}} \quad \text{(bits)}
\]

and may display

\[
I_{\text{surface}} = (3 \pi / 12) I_{\text{message}} / 4 \quad \text{bits}
\]

on a compact spherical packet surface. For rmessage = 0.2 micron, Imessage = 10⁹ bits/molecule (e.g., “genome packets”) with Isurface = 10⁶ bits. Of this number, at most ~1000 bits may be required for recipient identification, time stamping, “destroy by” dating, or other message packet flagging or header information. For the (CHX)ₓ core fragment, Imessage = n (bits).

Hydrofluorocarbon messenger molecules will not be readily metabolized. The smaller molecules are cleared rapidly from the circulation with brief retention in the mononuclear phagocyte system, mainly in the liver, spleen and bone marrow. They are then reintroduced into the circulation by lipid carriers in the blood at a rate that appears to depend on the fluorocarbon’s solubility in fat, with some concentration in adipose tissues. Eventually, they are excreted through the lungs with the expired air as a vapor (for short message segments of low molecular weight), with an intravascular persistence of 4-12 hours for present oxygen-carrying emulsions. (Linear mixed

* M. Krummenacker notes that if a common standard format is established for design economy, then any stray messenger molecules that are released intact into the environment could trigger unwanted action in other bodies, and at inappropriate times, if inhaled or ingested by someone other than the original patient. Such “information toxicity” can be reduced by judicious degradation and deflagging of all messenger molecules prior to final excretion, by recycling, and by employing validity-checking security protocols (Chapter 12) before permitting any action to be stimulated or influenced by received messenger molecules.
fluorocarbon-hydrocarbon amphiphiles have demonstrated at least some bloodstream persistence in vivo of up to 4 months.\textsuperscript{213} Massive single doses or repeated administration can cause saturation and transient blockade of the reticuloendothelial system (RES), temporarily depressing this component of the host-defense system.\textsuperscript{710} Message-carrying fluorocarbons should be extremely inert, although there is some evidence that macrophages which have ingested perfluorocompounds show loss of phagocytic function and possible release of cytokines and other immune mediators.\textsuperscript{708} Fluorosol has been found to elicit anaphylactoid-type reactions is a small percentage of patients at blood concentrations as low as ~0.1 gm/liter.\textsuperscript{711} Irregular Teflon (CF\textsubscript{2})\textsubscript{2} granules measuring 4-100 microns (equivalent to n = 10\textsuperscript{6}-10\textsuperscript{10}) employed clinically in periurethral injection\textsuperscript{344} have produced granulomatous reaction, embolization, and migration to the lungs.\textsuperscript{945-946} This implies a maximum size limit on large naked messenger molecules and a possible requirement for encapsulation in a biocompatible shell prior to release. R. Bradbury suggests that such a shell might include a lipid membrane with “docking tags” similar to viral attachment proteins. These tags should be engineered to be incapable of binding by antibodies, to have minimal affinity for common cell receptors, and to be indigestible by the regular cellular protein breakdown and recycling machinery (and thus be incapable of generating long-term immunity).

A hydrofluorocarbon messenger molecule with a lengthy stretch of exposed hydrogens might display increased toxicity, as some short-chain linear paraffin hydrocarbons (C\textsubscript{n}H\textsubscript{2n+2}) are considered “poisonous”\textsuperscript{2945} with an official NFPA Health Hazard Rating of 1 or “slightly toxic” (scale 0-4).\textsuperscript{2947} Animal toxicity of pure propane (C\textsubscript{3}H\textsubscript{8}) occurs with inhalation exposures to concentrations >10% and includes mild respiratory tract irritation and irregular respiration, cardiac sensitization to catecholamine-induced arrhythmias, analgesia and hypotension.\textsuperscript{2947,2948} although the principal risk is simple asphyxia and cold burn during bulk exposures. Typical OSHA workplace air exposure concentration limits for propane and butane are 0.06%-0.1%.\textsuperscript{2946,2947} Inhalation toxicity to pentafluoroethane (C\textsubscript{2}F\textsubscript{5}H) in rats is 70.9% (4-hour ALC), with no significant toxicological effects under <5% concentration.\textsuperscript{2948} Toxicity mechanisms may differ for short-chain and long-chain hydrocarbon molecules.

This entire area deserves an in-depth toxicology study. If long-term toxicity is found to be unacceptable, the solution may be to engineer “digestible” but non-immunogenic packages, reducing bit density but increasing the available molecular pool of raw materials for building messenger molecules in vivo. As a possible alternative to hydrofluorocarbons, R. Bradbury suggests using unusual protein or polysaccharide chains (which can compactly encode large quantities of data; Section 8.5.2.2)\textsuperscript{3122} that are easily read by nanorobots but also are easily digestible by cells as “food.” If patient immunotype is known, then protein sequences can be designed which cannot be bound by MHC molecules (Section 8.5.2.1), preventing the patient from developing an immunity to them, and which look like self-shapes or are nonbinding to antibodies.

### 7.2.1.2 Diffusion-Limited Broadcast Rate

Given I\textsubscript{message} = 10\textsuperscript{9} bits for r\textsubscript{message} = 0.2 micron from Eqn. 7.2, it is clear that onboard stores of messenger molecules will be quickly exhausted at the highest broadcast bit rates unless the communication system architecture makes substantial use of messenger molecule recycling. Without recycling, emission of hydrofluorocarbon message molecules is fluoride-limited, because carbon and hydrogen atoms are readily available from glucose and water molecules. The bloodstream concentration of fluoride is ~4.5 x 10\textsuperscript{7} gm/cm\textsuperscript{3} (Appendix B) or ~1.4 x 10\textsuperscript{22} F atoms/m\textsuperscript{3}. From Eqn. 3.4, the maximum diffusion current to the entire surface of a spherical nanodevice of radius R = 1 micron and diffusion constant D = 2 x 10\textsuperscript{5} m\textsuperscript{2}/sec (est. from Table 3.3) is \(I = 3.6 \times 10^{10} \text{ F atoms/sec or 1.1} \times 10^{17} \text{ kg/sec, enough fluorine to manufacture 2.7} \times 10^{17} \text{ kg/sec representing a continuous long-term maximum broadcast rate of} \ 7.2 \times 10^{9} \text{ bits/sec of (CHX)}_n \text{ 1-bit/unit hydrofluorocarbon messenger molecules.}

Even in the unlikely event that all 8.3 x 10\textsuperscript{22} atoms of fluorine present in the human body (Table 3.1) were somehow converted to hydrofluorocarbon messenger molecules and released into the bloodstream simultaneously, the bloodstream concentration would reach ~1.2 gm/liter, well below the ~700 gm/liter LD50, though admittedly exceeding the ~0.1 gm/liter limit that occasionally triggers anaphylactoid-type reactions.

In 1963, Bossert and Wilson\textsuperscript{703} completed the first systematic analysis of chemical communication in the context of animal olfactory communication in air, later extended by Bossert\textsuperscript{725} in an analysis of maximum possible bit rates. In the following discussion, their results are adapted to the problem of in vivo chemical communication among medical nanodevices.

#### 7.2.1.3 Instantaneous Stationary Source in Stationary Medium

Consider a chemomessaging nanorobot that releases Q\textsubscript{message} messenger molecules (each carrying I\textsubscript{message} bits/molecule) in a single puff as a point source at time t = 0. This is the ideal design for an alarm-type message or for messages requiring rapid fade-out. For simplicity, the human body is taken to be a continuous, isotropic, unbounded, stationary aqueous medium; a more detailed treatment is beyond the scope of this book. The spatial density of message molecules as a function of time t and distance r from the point source is

\[
U_{\text{message}} = \frac{Q_{\text{message}}}{(4 \pi D t)^{3/2}} e^{-r^2 / (4 D t)} \text{ (molecules/m}^3\text{)} \quad \text{[Eqn. 7.3]}
\]

de where D = the translational diffusion coefficient for message molecules, which are assumed to be roughly spherically packed, estimated from Eqs. 3.5 and 7.2 as

\[
D = \frac{kT}{\eta} \left( \frac{D_{\text{message}}}{162 \pi^2 I_{\text{message}}} \right) \quad \text{[Eqn. 7.4]}
\]

For I\textsubscript{message} = 100 bits, D = 2.2 x 10\textsuperscript{-11} m\textsuperscript{2}/sec; for I\textsubscript{message} = 10\textsuperscript{9} bits, D = 1.0 x 10\textsuperscript{-13} m\textsuperscript{2}/sec. Because the detection of molecules by receptor-based chemical sensors requires a concentration-dependent minimum sensor cycle time t\textsubscript{EQ}, approximated by Eqn. 4.3, then there exists some minimum threshold concentration (c\textsubscript{min}) of message molecules that can be detected by a chemical sensor in some minimum waiting time t\textsubscript{sensor} = t\textsubscript{EQ} given by

\[
c_{\text{min}} = \left( \frac{N_{\text{encounters}}}{2 \pi t_{\text{message}} t_{\text{sensor}}} \right)^{1/2} \left( \frac{I_{\text{message}} MW_{\text{unit}}}{4 \pi kT N_A} \right)^{1/2} \quad \text{[Eqn. 7.5]}
\]

For N\textsubscript{encounters} = 100, 10\textsuperscript{10} N\textsubscript{A} = 6.023 x 10\textsuperscript{23} molecules/mole (Avogadro’s number), t\textsubscript{message} from Eqn. 7.2 and taking t\textsubscript{sensor} = 1 sec, then c\textsubscript{min} = 1 x 10\textsuperscript{8} molecules/nm\textsuperscript{3} for I\textsubscript{message} = 100 bits and c\textsubscript{min} = 9 x 10\textsuperscript{11} molecules/nm\textsuperscript{3} for I\textsubscript{message} = 10\textsuperscript{9} bits. The concentration of message molecules exceeds c\textsubscript{min} within an expanding diffusive sphere. In a time t\textsubscript{esc} = (0.0293 / D) (Q\textsubscript{message} / c\textsubscript{min})\textsuperscript{1/2} this expanding sphere reaches a...
maximum size $R_{\text{max}} = (0.419) \left( \frac{Q_{\text{message}}}{c_{\min}} \right)^{1/3}$ after which it begins to contract as the puff dissipates; the concentration eventually falls below $c_{\min}$ everywhere at $t_{\text{fadout}} = \epsilon \frac{X_{\text{fadout}}}{V}$ where $\epsilon = 2.718...$

10^12 nanorobots uniformly distributed throughout a 0.1 m$^3$ human body have a mean interdevice separation of ~50 microns. For simple messages ($Q_{\text{message}} = 100$ bits) and $R_{\text{max}} = 50$ microns, then $Q_{\text{message}} = 2 \times 10^9$ message molecules emitted, $t_{\text{rec}} - 20$ sec (1 - 5 bits/sec), and $t_{\text{fadout}} - 50$ sec. If $R_{\text{max}} = 1$ mm, then $Q_{\text{message}} = 2 \times 10^{10}$ molecules and $t_{\text{rec}} - 8000$ sec to receive the signal (1 - 0.01 bits/sec). For complex messages ($Q_{\text{message}} = 10^5$ bits), $Q_{\text{message}} - 1 \times 10^5$ molecules but $t_{\text{rec}} - 4000$ sec (-1 hour) for the message to be transported $R_{\text{max}} = 50$ microns ($1 - 3 \times 10^5$ bits/sec).

A simple alarm signal ($Q_{\text{message}} = 100$ bits) released within an individual tissue cell is received everywhere throughout the cytosol—an assumed -20 micron$^3$ cell volume—in $t_{\text{rec}} - 0.8$ sec ($1 - 100$ bits/sec) with $t_{\text{fadout}} - 2$ sec, using a single alarm puff of $Q_{\text{message}} - 17,000$ message molecules (computed with $R_{\text{max}} - 10$ microns) which molecules may be stored in a (40 nm)$^3$ volume per puff. As a practical matter, the cytosol is quite crowded with macromolecules and cytoskeletal components (Section 8.5.3); exclusion effects will increase diffusion times by as yet unknown amounts, to be determined experimentally.

7.2.1.4 Instantaneous Stationary Source in Flooded Tube

For a source at the end of a long water-filled tube with non-absorbent walls, cross-sectional area $A_{\text{tube}}$ and length $X_{\text{max}} = (0.484) Q_{\text{message}} / A_{\text{tube}} c_{\min}^2$ and $t_{\text{fadout}} = \epsilon \frac{X_{\text{fadout}}}{V}$ using $c_{\min}$ from Eqn. 7.5. Hence a simple message ($Q_{\text{message}} = 100$ bits) passing through a tube of area $A_{\text{tube}}$ = 1 micron$^2$ and length $X_{\text{max}} = 50$ microns requires $Q_{\text{message}} = 130$ message molecules and is received in $t_{\text{rec}} - 60$ sec ($1 - 2$ bits/sec).

7.2.1.5 Continuous Stationary Source in Stationary Medium

Consider a source of message molecules emitting continuously at the constant rate $Q_{\text{message}}$ (message molecules/sec) into the idealized stationary medium described in Section 7.2.1.3. Such a source might be useful for status telemetry, navigational beacons, or periodic sampling monitors. If the source continues for a long time then the detectable threshold concentration sphere of message molecules around the point source asymptotically approaches a maximum radius $R_{\text{max}} = Q_{\text{message}} / (4 \pi D c_{\min})$. The time for the expanding detectable concentration sphere to reach a radius $R = f_{\text{R}} R_{\text{max}}$ (where $0 \leq f_{\text{R}} \leq 1$ is the fractional radial expansion of the message sphere), the exact solution for which involves the complementary error function, is approximated reasonably well by $t_{\text{rec}} - (1.1 f_{\text{R}} Q_{\text{message}} / 8 \pi c_{\min} (1-f_{\text{R}}) D^{1/2})^2$ for $0.1 < f_{\text{R}} \leq 1$, using $D$ from Eqn. 7.4 and $c_{\min}$ from Eqn. 7.5. Thus for simple messages ($Q_{\text{message}} = 100$ bits) with $R_{\text{max}} = 100$ microns and $R = 50$ microns ($f_{\text{R}} = 0.5$), then $Q_{\text{message}} - 4 \times 10^9$ message molecules emitted per second and $t_{\text{rec}} - 140$ sec ($1 - 1$ bit/sec). For complex messages ($Q_{\text{message}} = 10^5$ bits), $Q_{\text{message}} - 10$ message molecules emitted per second and $t_{\text{rec}} - 8.4$ hours ($1 - 3 \times 10^4$ bits/sec).

7.2.1.6 Continuous Mobile Source in Stationary Medium

Consider a mobile nanorobot traveling through a stationary aqueous medium at velocity $V_v$ (m/sec), emitting message molecules continuously at the rate of $Q_{\text{message}}$ (molecules/sec). The cigar-shaped molecular trail envelope has a maximum length $X_{\text{fadout}} = Q_{\text{message}} / 4 \pi D c_{\min}$, and at a distance $X_{\text{max}} = X_{\text{fadout}} / \epsilon$ behind the moving source the message molecules diffuse outward to a maximum detectable radius $R_{\text{max}} = (0.342) \left( \frac{Q_{\text{message}}}{V_v c_{\min}} \right)^{1/3}$ as measured from the motion axis; for distances >> $X_{\text{max}}$ the radius of the detectable message molecule envelope declines to zero at $X_{\text{fadout}}$ with $t_{\text{fadout}} = X_{\text{fadout}} / V_v$. Thus, a nanorobot velocity $V_v = 10$ microns/sec and an $X_{\text{fadout}} = 1000$ microns requires $Q_{\text{message}} = 4 \times 10^7$ molecules/sec for simple message molecules ($Q_{\text{message}} = 100$ bits) giving $R_{\text{max}} - 60$ microns at $X_{\text{max}} = 370$ microns and $t_{\text{fadout}} = 100$ sec.

These equations also apply to a continuous stationary source in a nonstationary medium, where the medium moves past the source with a slow, constant velocity with perfect laminar nonturbulent flow in the absence of nearby boundary surfaces, an idealized instance of the case described in Section 7.2.1.7.

7.2.1.7 Continuous Stationary Source in Nonstationary Medium

A nonstationary transporting medium does more than just move message molecules in the direction of flow—it also may create turbulence in the medium, adding a component of turbulent diffusivity that overwhelms simple Brownian diffusion, giving a total diffusivity that is more a property of the flow structure of the medium and the boundary surfaces than of the substance of the medium, a complex analytical problem. Consider a continuous point source emitting a constant $Q_{\text{message}}$ (molecules/sec) from a fixed position near the luminal surface of a blood vessel radius $R_{\text{vase}}$.

From the relations given in and as a crude approximation assuming plume width << boundary wall separation, $Q_{\text{message}} = (0.0396) V_{\text{vase}} c_{\min} X_{\text{fadout}} / 7^{1/4}$ to create an ellipsoidal message molecule plume of maximum detectable length $X_{\text{fadout}}$ in a medium flowing along the x-axis with fluid velocity $V_{\text{vase}}$. The widest part of the plume occurs at $X_{\text{max}} - (3.57) F^{1/3}, Y_{\text{max}} - (0.686) F^{1/2},$ and $Z_{\text{max}} - (3.41) F^{1/2}$, where $F = Q_{\text{message}} / V_{\text{vase}} c_{\min}$. Message receipt time $t_{\text{rec}} = X_{\text{fadout}} / V_{\text{vase}}$.

Thus a simple message ($Q_{\text{message}} = 100$ bits, $c_{\min} - 10^{-9}$ molecules/nm$^3$ for $t_{\text{sensor}} = 1$ sec) broadcast from the luminal surface of the common carotid artery ($V_{\text{vase}} = 0.2$ m/sec, $R_{\text{vase}} = 3$ mm; Table 8.2) with a desired plume length $X_{\text{fadout}} = 100$ microns requires at least $Q_{\text{message}} - 10^9$ molecules/sec, producing a very narrow plume with $Y_{\text{max}} - 40$ microns, $Z_{\text{max}} - 20$ microns (the radial direction), and $t_{\text{rec}} - 0.5$ millisecond ($1 - 100$ bits/sec, limited by $t_{\text{sensor}} = 1$ sec). For a complex message ($Q_{\text{message}} = 10^5$ bits, $c_{\min} - 9 \times 10^{-11}$ molecules/nm$^3$ for $t_{\text{sensor}} = 1$ sec), $Q_{\text{message}} - 7 \times 10^7$ molecules/sec at $t_{\text{rec}} - 0.5$ millisecond ($1 - 10^5$ bits/sec, limited by $t_{\text{sensor}} = 1$ sec).

7.2.1.8 Assessment of Chemical Broadcast Messaging

Chemical broadcast messaging in theory permits information transfers up to $10^9$ bits/sec but appears to be extremely energy-inefficient. The assembly of (CHX)$_n$ messenger molecule units may require 1-1000 zJ/bit (Section 6.5.6 (B)). For instance, replacing C-H (642 zJ) with C-F (876 zJ) costs ~234 zJ, but recycling CH$_3$F molecules into C-CH$_2$F chains actually produces 7 gJ of energy. A feedstock of CH$_3$F and CH$_4$ will release energy when the tape is first assembled, but will cost energy to tear apart into its fundamental subunits. R. Bradbury believes that the net cost of tape production and recycling can be reduced to 10-50 zJ/bit, given sufficient space to store lots of tape subunits: “If the assembly process is efficient at harvesting the energy produced, and the molecules don’t get lost to cellular recycling machinery, then you could produce a system with very high overall efficiency, though you will have to worry about distribution of harvested energy and feedstock.”

However, since $10^6$ molecules must be released to ensure receipt of at least one message molecule by an intended recipient -50 microns away (e.g., Section 7.2.1.3), the energy cost per received bit rises sharply to $10^6-10^9$ zJ/bit, which is >> 3 zJ/bit, the
theoretical minimum (Eqn. 7.1). Thus a 1-pW communications power budget limits steady-state bandwidth to ~1 bit/sec across ~50-micron transmission ranges. Additional disadvantages of chemical communications are slowness of transmission (except in nonstationary media or across short distances) and the ability to be delayed or blocked by semipermeable or impermeable barriers, or by absorptive nanorobotic agents active in the environment. Noise on chemical channels may almost be eliminated by using artificial messenger molecules, reducing the natural background nearly to zero, though noise may still be present due to the presence of old signals, interference or cross-talk from other man-made devices, or even from intentional medical or military jamming. The encoding mechanism of “genome packets” would act less like email, and more like FTP because the amount of information transferred per messenger molecule is so large. Finally, used messenger molecules must be degraded by nanomachines (or other means) after the “destroy by” date has been passed; the number of messenger molecules to be captured and degraded or recycled by the messaging nanorobot population must approximately approximate the number of messenger molecules released.

Nanorobot-to-nanorobot chemical broadcasts may be useful in the case of nanorobots working in close quarters where low emission rates will suffice, as in bacterial quorum sensing. For example, achieving the minimum detectable concentration \( c_{min} \approx 10^{-9} \text{ molecules/nm}^3 \) for 100-bit message molecules inside a (20 micron) \(^3 \) human tissue cell requires emission of only \(-10^{4} \) molecules (Section 7.2.1.3), costing \(-1 \text{ pJ} \) to manufacture and representing just \( 10^{-7} \% \) of tissue cell volume. Chemical communications may also be useful in various special applications such as “silent alarms” using premanufactured messenger molecules vented from onboard storage tanks, thus moving the manufacturing energy cost ex vivo. Assuming 0.1 micron \(^3 \) tankage per nanorobot, each device may store \(-10^{8} \) 100-bit message molecules available for ready release, enough to signal \(-10^{4} \) in cyto alarm events. Each such alarm signal may permeate the entire cytosol in \(-1 \) sec.

Chemical broadcast messaging may be most useful when the high energy cost of creating the message molecules is borne by some external or macroscale agency, thereby requiring individual nanorobots only to receive the chemical signal but not to send it. For example, messenger molecules manufactured in the laboratory, treatment clinic, or implanted dedicated communication organ (Section 7.3.4) may be injected into the bloodstream in sufficient concentration to ensure receipt. Given \( c_{min} = 1 \times 10^{-9} \text{ molecules/nm}^3 \) (\( \text{message} = 100 \text{ bits/molecule} \)) or \( c_{min} \approx 9 \times 10^{-13} \text{ molecules/nm}^3 \) (\( \text{message} = 10^6 \text{ bits/molecule} \)), the minimum number of messenger molecules needed to promptly raise the entire bloodstream concentration to \( c_{min} \) (assuming no absorption by the body) is \( >7 \times 10^{15} \) 100-bit messenger molecules (a \(-0.006 \text{ mm}^3 \) injection) or \( >5 \times 10^{14} \) 100-bit messenger molecules (a 4 cm \(^3 \) injection producing \(-0.9 \text{ gm/liter} \) concentration, somewhat exceeding the \(-0.1 \text{ gm/liter} \) threshold that occasionally produces anaphylactoid-type reactions). By comparison, a single nanorobot with a 1 pW power budget can only manufacture \(-10^{4} \text{ to } 10^{5} \text{ 100-bit message molecules/sec} \).

In 1998, chemical messaging was only starting to be applied in robotics. For example, in one experiment simple pheromone-guided macroscale mobile robots could follow a scent plume using silk-worm moth antennae wired into an electronic neural network.

7.2.2 Acoustic Broadcast Communication

Communication by acoustic waves is a highly useful channel for information transfer in nanomedicine. Previous Chapters have described the reception of acoustic energy in sensing (Sections 4.5, 4.8.2 and 4.9.1) and in the transfer of power (Sections 6.3.3, 6.4.1 and 6.4.3.3). Issues remaining to be addressed here include the generation of ultrasonic radiation by nanorobots (Section 7.2.2.1) and free-tissue acoustic channel bandwidths available in the human body (Section 7.2.2.2). Acoustic cables and transmission lines are treated in Section 7.2.5.3.

7.2.2.1 Acoustic Radiators

The acoustic power generated by any vibrating source is \( P_{out} = U^2 R_A \) (RMS watts), where \( U \) is the rate of oscillatory volume displacement of the fluid, measured in RMS m/sec, and \( R_A \) is the acoustic radiation resistance seen by the source, measured in MKS acoustic ohms. Two basic generators are the vibrating piston and the pulsating sphere, both of radius \( r \). From Stokes Law (Eqn. 9.73) the input power required to propel a circular piston through water at velocity \( v \) is \( P_{in} = 6 \pi r \eta v^3 \), or \( P_{in} = 24 \pi r \eta v^2 \) for a radially oscillating sphere. Hence \( U = \pi r^2 v = (\pi r^3 P_{in} / 6 \eta)^{1/2} \) for a vibrating piston to which \( P_{in} \) watts of mechanical input power are delivered, and \( U = (2 \pi r^3 P_{in} / 3 \eta)^{1/2} \) for the pulsating sphere.

The radiation resistance \( R_A = \pi \rho v^2 / k_v \) for \( v = 2 \text{ (piston)} \) or \( v = 4 \text{ (sphere)} \) and \( v \) is acoustic frequency (Hz). This equation is valid only when \( r/\lambda < 1 \) (e.g., radiator size is very small in comparison to acoustic wavelength \( \lambda = \nu / v \); for \( r = 1 \text{ micron} \) and \( v = 1 \text{ MHz} \), then \( r/\lambda - 0.001 \) in water. Sound energy generated from a source with dimensions small compared with the wavelength of the vibration in the medium produces an intensity uniform in all angular directions; such a generator is considered to be a point source. Formulas for radiators having \( r/\lambda >> 1 \) are also provided by Massa.

Combining these results with Eqn. 4.53, transmitted acoustic pressure \( (A_p) \), output power \( (P_{out}) \), and power intensity \( (I_p) \) at the radiator surface are given by

\[
A_p = \left( \frac{\pi \rho v^2 r}{3 k_v \eta} \right)^{1/2} \text{ (N/m}^2) \]  
\[P_{out} = \frac{k_v \rho v^3 r^3 P_{in}}{24 \eta v_{sound}} \text{ (watts)} \]  
\[I_p = \frac{4 P_{out}}{k_v \eta \pi^2} \text{ (watts/m}^2) \]

For example, a vibrating piston radiator of radius \( r = 1 \text{ micron} \) and input power \( P_{in} = 10 \text{ pW operating at v = 1 MHz in vivo produces A}_{p} = 0.0007 \text{ atm radiation pressure, P}_{out} = 0.005 \text{ pW} \) (giving \( \eta = P_{out} / P_{in} = 0.0005 \) (0.05%) and an acoustic intensity of \( I_p = 0.002 \text{ watts/m}^2 \) at the surface of the radiator, assuming \( \rho = 993.4 \text{ kg/m}^3 \), \( \eta = 1.1 \times 10^{-3} \text{ kg/m-sec} \) and \( v_{sound} = 1500 \text{ m/sec} \) for human interstitial fluid at 310 K. Figure 7.1 summarizes Eqn. 7.6 for various parameter choices.

7.2.2.2 Free-Tissue Acoustic Channel Capacity

Figure 7.1 (with blood pressure variations compared; Section 4.9.1,2) and Eqn. 7.7 illustrate the well-known result in acoustics that for a given driving amplitude, micropistons and microspheres are more powerful sound radiators at higher frequencies. That is, input power (driving the radiator) is more efficiently transduced into output power (waves in the medium) both at higher frequencies and at larger radiator sizes. Thus to achieve the highest acoustic channel capacity per unit of input power, the highest practical frequency and the largest possible radiator should be used for nanorobot-to-nanorobot acoustic communications. Of course, at higher frequencies, attenuation becomes more severe and eventually limits the value of even higher frequencies.
7.2.3 Electromagnetic Broadcast Communication

The receipt of externally-generated radiofrequency (rf) power signals up to ~MHz frequencies by nanorobot antennas has already been considered in Section 6.4.2. Such signals are readily adapted to communication and may carry ~MHz signals into the human body, allowing information transfers to in vivo nanorobots of up to ~10^6 bits/sec from an external signal intensity of ~0.1 watts/m^2. Eqn. 6.32 suggests that rf signals pass through soft tissue with negligible absorption, leaving most of the signal energy available for detection by nanorobot receivers. (See also Section 4.7.1.)

While externally-generated rf signals may be detectable by nanorobots, onboard submicron-scale broadcast antennas probably cannot generate rf signals of sufficient power for meaningful device-to-device communications. For example, low-frequency electromagnetic waves of wavelength \( \lambda_{rf} \) and frequency \( f_{rf} \) customarily are generated using a simple electric dipole of length \( d_e \ll \lambda_{rf} / 2 \pi \), or using a magnetic dipole loop antenna of diameter \( d_M \ll \lambda_{rf} / 2 \pi \). If the current carried by either antenna is \( I_{dipole} \), then the average transmitted power for each radiator is given by:

\[
P_{E} = \frac{1}{3} \pi \mu_0 c \left( \frac{d_e}{\lambda_{rf}} \right)^2 \frac{1}{\lambda_{dipole}} \quad \text{(watts)} \tag{Eqn. 7.10}
\]

\[
P_{M} = \frac{1}{12} \pi^5 \mu_0 \left( \frac{d_M}{\lambda_{rf}} \right)^4 \frac{1}{\lambda_{dipole}} \quad \text{(watts)} \tag{Eqn. 7.11}
\]

\[
P_{E} = \frac{4 \lambda_{rf}^2 d_e^2}{\pi^4 d_M^4} \quad \text{(watts)} \tag{Eqn. 7.12}
\]

where \( \mu_0 = 1.26 \times 10^{-6} \) henry/m (permeability constant), and \( c = 3 \times 10^8 \) m/sec (speed of light). For \( f_{rf} = 1 \) kHz and \( d_e = d_M = 1 \) micron, \( P_{E} / P_{M} \approx 10^5 \) and the electric dipole is strongly preferred over the magnetic dipole radiator; at high frequencies, the relative power emission of the electric dipole is even greater.

An electric dipole antenna of length 1 micron and cross-section 1 micron^2 carrying a fairly aggressive current density of ~10^6 amp/m^2 consumes ~200 pW to produce an rf output power of only \( P_{E} = 10^{-11} \) pW. Not only is this extremely energy-inefficient, but the signal is too weak for detection by neighboring nanorobots. The 430-nm electrostatic pendulum antenna described by Eqn. 6.34 has a theoretical minimum incident power detection limit of ~0.01 pW at 0.7 MHz; a limit of \( kT e^{SNR} = 30 \) zJ/cycle for a 10^-11 pW signal implies a maximum operating frequency of ~0.0003 Hz, too low to convey any meaningful information. (Here, signal-to-noise ratio \( SNR = 2 \) (20 dB); error-corrected digital signals can provide virtually perfect communications even below 10 dB.) Furthermore, simple scaling predicts nanoelectric antennas will have nondirectional radiation patterns, since antenna gain \( G \approx A / r^2 \) (Section 6.4.1) at the transmitter surface. Further increases in \( r \) cause \( I \) to decline, because \( e% \) cannot improve beyond a maximum of 100%.

These results imply that nanorobot-to-nanorobot acoustic communications will generally take place at ~10-100 MHz frequencies over 10-100 micron path lengths in vivo. Acoustic messaging over longer path lengths require mobile signal amplifiers such as communicytes (Section 7.2.6), dedicated fixed-position communication organs with repeater protocols, or packet routing networks analogous to the Internet (Section 7.3).
to interrogation) in order to establish spatial distributions of those phenomena—which may include temperature, pressure, biochemical concentrations, bloodflow velocities, detection of specific cell types or metabolic processes. In 1999, the 250-micron microtransponders under development by PharmaSeq were to be expected to be interrogated at a throughput rate of $\simeq 1000/$sec.

Aside from the environmental sources of electrical noise already described in Section 4.7.1, two additional potential sources deserve brief mention. First, radio, television, radar and other broadcast sources in developed countries produce a frequency-dependent broadband background (e.g., a carrier-frequency spike for each source) across a 0.01-1000 MHz spectrum up to $10^{-11}$ watts/m$^2$-Hz which at $\leq$MHz frequencies passes through the human body largely unattenuated (Section 6.4.2). However, a -1 micron$^2$ antenna employing a -1 MHz channel receives at most $10^{-5}$ pW of this broadcast energy, only $0.002$ kT/cycle which is almost certainly undetectable. In other words, micron-scale nanorobots probably cannot directly receive commercial TV, radio, or satellite broadcasts or GPS signals unless mediated by a macroscopic communication organ (Section 7.3.4) or assemblage. Hence such broadcasts should generate no detectable noise at these frequencies.

Second, electrical discharges associated with muscular and neural action throughout the body cause electrical potentials to appear at the skin, which are typically detected in a medical context during electrocardiographs (ECG or EKG, $-10^{-3}$ volt), electroencephalographs (EEG, $-10^{-3}$ volt), and biofeedback monitoring with frequencies ranging from 1-40 Hz. These surface waves also may reach $10^{-11}$ watts/m$^2$-Hz, producing an undetectable $10^{-5}$ pW signal in a -1 micron$^2$ antenna using a 100-Hz channel—although an electric field strength of $10^{-2}$ volts/m is in principle detectable by living cells (Section 4.7.1). Thus, micron-scale nanorobots can sense individual neural discharges nearby but probably cannot directly monitor global brainwave, ECG, gastroelectric or intestinal electrical wave patterns (which require a macroscopic electrosensory organ or, possibly, cellular eavesdropping).

Electromagnetic modulations $\nu_{rf} = 10-100$ MHz (which must be imposed upon microwave, infrared, or optical carrier waves) may be difficult to generate, control and detect using submicron nanorobot components if those components are limited to mechanical $\quad -$GHz motions (Section 10.1.2.2); electronic components may cycle faster. Such carrier waves are $99\%$ absorbed by soft tissue for path lengths ranging from -10 cm for 100 MHz waves, to -1 mm for infrared, or -40 microns for optical photons. Energetic carriers also face a sharply declining count rate at the receiver due to the increasing particle-like nature of the nondirectional transmission carrier at higher frequencies. For example, while micron-scale solid-state lasers are available, from simple geometry, an omnidirectional 1-micron$^2$ optical photon ($10^{-14}$ Hz) emitter with a 1-pW power budget producing $10^7$ photons/sec transmits only $10^3$ photons/sec to a 1-micron$^2$ receiver located 100 microns away (maximum mean free path in tissue; Section 4.9.4). This limits information transfer to $10^3$ bits/sec at an energy cost of $10^{12}$ zJ/bit. Short bursts at much higher bit rates also satisfy the stated power budget; using this strategy, photon intensity must still be held below -100 watts/m$^2$ (Section 6.4.2) at the source, so for a 1 micron$^2$ transmitter the maximum transfer rate is $10^5$ bits/sec in a -0.01 second burst, with bursts repeatable only once per second.

Relatively slow modulations impressed on high-frequency electromagnetic waves might provide another useful communications channel. For example, nanorobots may detect each other’s heat signatures, so modulations of those signatures can be used to transmit messages. Consider a nanorobot immersed in a medium of thermal conductivity $K$, and heat capacity $C_V$, producing $P_n$ watts of thermal power at a distance $R_{detect}$ from a second nanorobot equipped with a thermal sensor of temperature sensitivity $\Delta T_{min}$. The maximum detection range is

$$R_{detect} = \frac{P_n}{4 \pi K \Delta T_{min}}$$  \hspace{1cm} \text{(Eqn 7.13)}$$

For $K = 0.623$ watts/m-K for water at 310 K, $\Delta T_{min} = 10^{-6}$ K (Section 4.6), and $P_n = 10-1000$ pW, then $R_{detect} = 1-100$ microns.

If $P_n$ is pulsed on and off to send a message, the thermal time constant $\tau_{thermal} = 4 \pi R_{detect} C_V / K = 0.0001-1$ sec for $C_V = 4.19 \times 10^4$ watts/m$^3$-K in water at 310 K. Assuming single-channel bandwidth $\nu_{rf} \simeq 1 / \tau_{thermal}$ then detectable signals range from -10 KHz ($-10^4$ bits/sec) at a distance $R_{detect} = 1$ micron from a 10-pW modulated heat signature down to 1 Hz (-1 bit/sec) at a distance $R_{detect} = 100$ microns from a 1000-pW modulated heat signature.

### 7.2.4 Nanomechanical Communication

Internal communications (or interdevice communications through hard junctions between linked devices; Section 5.4.2) are readily achieved via sliding or rotating mechanical rods and couplings. A thin diamondoid rod producing 1-nm displacements and oscillating longitudinally at -1 m/sec may transfer information at -GHz frequencies. Similarly, a diamondoid transmission line carries acoustic compression pulses at $150$ m/sec, providing a 10 GHz bandwidth channel sufficient to support the -GHz clocking speeds typical of sensors and nanomechanical computers (Section 10.1).

A simple communication mode between unlinked nanorobots operating in close quarters is simple physical tapping. During a 1 microsecond contact event, transfer of 1000 bits is allowed at a 1 GHz line frequency assuming single-threaded transmission; arrays of isolated transmission lines can transmit much more information. Cleavage energies for diamond range from $E_{cleave} = 10.6$ J/m$^2$ for the [111] crystal plane to 18.4 J/m$^2$ for the [100] plane. Adopting the more conservative number to estimate the minimum hammer/anvil cleavage velocity $v_{cleave}$, and as a crude approximation, a rodlike diamondoid signal hammer 100 nm in length moving at velocity $v_{hammer}$ with tip area $A_{tip} = 100$ nm$^2$ and mass $m_{hammer} = 3.5 \times 10^{-20}$ kg impacting a diamondoid anvil at speeds of:

$$v_{hammer} < v_{cleave} = \left( \frac{2 E_{cleave} A_{tip}}{m_{hammer}} \right)^{1/2} = 246 \text{ m/sec} \quad \text{(Eqn 7.14)}$$

cannot cause cleavage in an unfractured anvil structure. The smallest detectable impact for this hammer has energy $kT e^{SNR} = (1/2) m_{hammer} v_{hammer}^2$, giving minimum detectable $v_{hammer} = 1.3$ m/sec for $SNR = 2$ (Section 4.3.1) at $T = 310$ K. At this speed, and neglecting anvil mass, a hammer cycled at $v_{hammer} = 1$ GHz ($-10^9$ bits/sec) against a spring-loaded receiver anvil has a displacement stroke of:

$$\Delta x = \frac{(kT e^{SNR})^{1/2}}{2 m_{hammer} v_{hammer}} = 670 \text{ pm} \quad \text{(Eqn 7.15)}$$

well above the minimum detectable displacement of 10 pm (Section 4.3.1), giving a maximum power cost of $kT e^{SNR} v_{hammer}^{-2} = 32$ pW ($-32$ zJ/bit).

Mechanical communication can also be achieved by more passive means, as for example the posting of Braille-like tactile-readable information mechanically on the nanorobot surface which may be scanned by neighboring nanorobots using compliant read tips at the distal ends of flexible manipulator arms (Section 5.3.2.5).
7.2.5 Cable Communication

Signals may be imposed upon any tethered power transmission (Section 6.4.3), allowing the transfer of information as well as energy. This Section examines the ability of various cable carriers to transmit information. The drawbacks of tethered power transmission systems enumerated in Section 6.4.3.6 also apply to in vivo communication cables.

7.2.5.1 Electrical Cables

Uninsulated wires risk electrical leakage and unnecessary heat transfer; unshielded wires produce stray fields that may attract unwanted attention from mobile elements of the immune system, e.g., electrosensitive leukotaxis (Section 4.9.3.1). Thus coaxial cable or “coax” is the proper transmission line for in vivo electrical cable communications. Unlike waveguides, coaxial cable is not a resonant device, hence the traveling wave frequencies it carries may be varied continuously down to DC. The electric fields within a coax line have axial symmetry. However, serious design complications may arise if the transmitted wavelength (λ) is less than the circumference of the annular dielectric space between the conductors, because a second (nonaxial) mode of rf propagation with diametral symmetry can arise if the transmitted wavelength (λ) is less than the circumference of the annular dielectric space between the conductors, because a second (nonaxial) mode of rf propagation with diametral symmetry can then be excited simultaneously. To avoid this complication, a 1-micron diameter coax should be held to a practical upper frequency limit of c / π d coax < 1 x 10¹⁴ Hz for cable lengths > λ.

Intracoaxial scattering and absorption are negligible, so for single-channel bandwidth the maximum coaxial bit rate I max consistent with a nanorobot onboard communication power budget P comm using rf or microwave photons (h ν coax / kT ln(2) << 1) occurs at the full-power transmission frequency

\[ ν_{\text{coax}} (\nu_{\text{max}}) = \frac{P_{\text{comm}}}{kT \ln(2)} \text{ (Hz)} \]  
\[ \text{Eqn. 7.16} \]

Thus the maximum coaxial transmission frequency is 0.1 GHz (10⁸ bits/sec) for P comm = 0.3 pW, 1 GHz at 3 pW, and 100 GHz (10¹¹ bits/sec) for P comm = 300 pW (>3 zJ/bit).

7.2.5.2 Infrared and Optical Cables

Photon energy equals thermal noise (h ν equal / kT ln(2) = 1) in the far infrared, specifically, at ν equal = 6 x 10¹⁵ Hz for T = 310 K. At still higher frequencies, the cost per bit becomes the energy per photon, or h ν optical, hence the maximum transmission frequency becomes

\[ ν_{\text{optical}} = \left( \frac{P_{\text{comm}}}{h} \right)^{1/2} \text{ (Hz)} \]  
\[ \text{Eqn. 7.17} \]

Silica optical fibers are damaged by near-UV photons at frequencies exceeding ~10¹⁵ Hz (λ ~ 300 nm), giving an upper limit on optical information transfer rates. Transmitting 10¹⁵ UV photons per second requires a power budget of P comm = 1 milliwatt; holding intensity to <10⁵ watts/m² for safety reasons (Section 6.4.3.2) mandates a fiber >100 microns in diameter. Cables this large are suitable primarily for relatively short internodal trunk lines (Section 7.3.1). At ~10⁵ watts/m² a 1 micron² fiber can safely continuously transmit ~10⁵ pW, enough to generate 10¹³ photons/sec at 10 THz (~10¹³ bits/sec), or ~10 zJ/bit. Note that below ν equal the energy cost per bit is constant; above ν equal the energy cost per bit rises linearly with frequency. Regenerative systems could in theory recover and reuse some of this energy.

7.2.5.3 Acoustic Cables and Transmission Lines

Consider an acoustic compression signal of frequency ν acoust traveling down a fluid-filled transmission cable of length l cable and radius r cable. In order to be detected at a signal/noise ratio SNR = 2, each pulse cycle must transfer a minimum energy kT e⁻SNR to a receiver of volume l³ at the cable terminus; hence from Eqn. 4.29, the maximum detectable pulse frequency ν max acoust = I peak⁴ l³ / kT e⁻SNR (Hz), where I peak is the acoustic power intensity at the receiver. To obtain this intensity at the receiver, the input signal power at the other end of the cable must be P₀ ≃ π r cable² Itrans (watts) where Itrans is the power intensity of the signal transmitter. Combining these relations with Eqn. 6.44 for acoustic attenuation in cables gives

\[ ν_{\text{acoust}} \leq \left( \frac{L^2}{kT e^{-SNR}} \right) I_{\text{trans}} e^{-2 \alpha_{\text{tube}} l_{\text{acoust}}} \text{ (Hz)} \]  
\[ \text{Eqn. 7.18} \]

where α tube is given by Eqn. 6.45 with a -ν 1/2 acoust dependency. To avoid cavitation in pure water (Section 6.4.3.3) and to ensure safety in the unlikely event of cable detachment in vivo, maximum I peak = 10⁴ watts/m² of acoustic energy (Fig. 6.8). For a cable of diameter 1 micron terminating on a receiver of volume (680 nm)³ sensitive to 10⁶ atm displacements (Section 4.5.1), then ν acoust ~ 1 GHz (10⁹ bits/sec) for a cable of length l cable = 14 microns (~8000 zJ/bit). A 1000-micron long cable can transmit up to ~1 MHz; ν acoust ~ 1 GHz at l cable ~ 5 cm; and ν acoust ~ 1 Hz at l cable ~ 2 meters. For intradevice communications, ν acoust ~ 1 GHz for L = 300 nm, l cable = 1 micron, and r tube = 50 nm, though pure diamond fiber may be more efficient in this case.

A solid diamond acoustic transmission line transfers signals over great distances at GHz frequencies with almost no power losses, in part because of the extreme stiffness of diamond. Consider a rod of volume ν rod at temperature ν rod, made of a material with a thermal coefficient of volume expansion β and heat capacity C V, to which a pressure pulse ΔP is applied at one end that travels at velocity ν sound (~17,300 m/sec for diamond) to the other end. In the worst-case thermodynamic cycle, Drexler ¹⁰ gives the total energy dissipation per pulse W max as

\[ W_{\text{max}} = \frac{T_{\text{rod}} V_{\text{rod}} B^2 \Delta P^2}{C_V} \text{ (joule/cycle)} \]  
\[ \text{Eqn. 7.19} \]

For diamond at T rod = 310 K, β = 3.5 x 10⁻⁶ /K and C V = 1.8 x 10⁶ joule/m³-K, ν sound = 1.8 x 10⁶ m/sec. Thus a 1-atm pulse that is applied to a transmission line consisting of a diamond rod 1 micron in length and (10 nm)² in cross sectional area (ν rod = 10⁻²² m³) requires a pulse input energy of ~10,000 zJ, but suffers an energy loss during transmission of at most W max = 2 x 10⁻⁶ zJ. Under smooth mechanical cycling, nanomechanical systems may approach the isothermal limit and significantly reduce dissipation still further, to ~1% W max at ~1 GHz and ~0.001% W max at ~1 MHz. Thus even at ν rod ~ 1 GHz the losses using 1-atm pulses amount to only ~20 zJ and the transmission of energy is still ~99.8% efficient.

7.2.5.4 Mechanical Cables

Information may be transferred mechanically by modulating the turning frequency of a rotating cable. In this case the mechanical load is applied by the receiver at the cable terminus as it measures frequency changes in the received signal (e.g., FM modulation). Section 6.4.3.4 discusses cable operating modes, frequencies,
dimensions, and power requirements. For example, a rotating cable
2 nm in diameter and ~50 nm long (suitable for internal data
transmissions) operated in AC mode can convey 6 x 10^14 bits/sec at
a power cost of 1 pW (~16,000 zJ/bit).

Power and control signals can also be transmitted through a stiff
reciprocating member, operating in the manner of a bicycle brake
deraillieur cable. Molecular control cables using polyyne rods
encased in single-walled carbon nanotube (Section 2.3.2) sheaths is
one simple example. Polyne is a chain of carbon atoms with
alternating single (0.1377 nm) and triple (0.1192 nm) bonds, hence
the polyyne cable has ~7.8 carbon atoms and a mass of ~1.5 x 10^{-5}
kg per nanometer of length. The stretching stiffness of the single
bond is ~824 N/m, the triple bond ~1560 N/m, and the length of
a bond pair is 0.2569 nm, giving the compliance for a cable of
length L of (7.2 x 10^6) L (m/N), or the stiffness as k_b = (1.4 x 10^{-7})
/ L (N/m). Thus an L = 1 micron polyyne cable stretches by ~1 nm
when a tensile force of 140 pN is applied. Thermal noise produces
an uncertainty in the position of the end of the polyyne cable, which
varies approximately as \Delta x = (kT / k_b)^{1/2} ~ 0.18 nm in the previous
example, at 310 K.

7.2.5.5 Chemomessenger Cables

Unlike the chemical messaging described in Section 7.2.1, which
was dominated by diffusive effects, information transfer using
chemical messenger molecules that are confined within sealed pipes
is controlled by the laminar bulk flow rate of the rapidly-moving
carrier fluid.

Consider a chemomessenger cable of length \( l_{cable} \) and radius \( r_{cable} \)
carrying a fluid of viscosity \( \eta \) at a volumetric flow rate of \( V \)
maintained by a pressure differential of \( p_{cable} \) between the two ends of the cable. Assuming the fluid is a 10% suspension of messenger
molecules of information density \( D_{message} \sim 26 \text{ bits/nm}^3 \) (Section
7.2.1.1) with viscosity similar to human plasma, then net fluid
information density \( D_{fluid} = 2.6 \times 10^{-27} \text{ bits/m}^3 \); from the
Hagen-Poiseuille law (Section 9.2.5) the maximum information
transfer rate is

\[
I_{chemo} = \frac{\pi r_{cable}^4 p_{cable} D_{cable}}{8 \eta l_{cable}} \quad \text{(bits/sec)} \quad \text{(Eqn. 7.20)}
\]

Assuming a safe, ultraconservative \( p_{cable} = 1 \text{ atm} \) and taking \( r_{cable} = 0.5 \text{ micron} \) and \( \eta = 1.1 \times 10^{-3} \text{ kg/m-sec} \), then
\( I_{chemo} = 10^{18} \text{ bits/sec} \) through a cable of length \( l_{cable} = 50 \text{ microns} \), with power con-
sumption \( P_{chemo} = \pi r_{tube}^4 p_{cable}^2 / 8 \eta l_{cable} = 4500 \text{ pW} \) (-0.04 zJ/
bit). For a cable 0.5 meter in length, \( I_{chemo} = 10^{19} \text{ bits/sec} \) and \( P_{chemo} = 0.4 \text{ pW} \) (-0.04 zJ/bit). If \( p_{cable} \) is
more liberally raised to 1000 atm, a cable 0.5 meter in length can transfer
\( I_{chemo} = 10^{20} \text{ bits/sec} \) requiring \( P_{chemo} = 450 \text{ nanowatts} \) (-40 zJ/bit). While these are
phenomenal information transport rates compared to other methods,
two important caveats are in order.

First, such high transfer rates are purchased at the price of
significantly increased receiver complexity and message processing
time, since the message molecules must be captured, oriented,
unspooled, fed past a read head at relatively slow speed, then stored,
recycled, or disposed of properly. Data carrier fluid must be returned
to the transmitter using a second cable; a double-cable pair establishes
a complete fluidic circuit. The additional transmitter complexity
and extra power required for chemical modifications of message
 carriers may be confined to external facilities and hence do not
significantly constrain in vivo operations.

Second, the message travel speed from one end of the cable
to the other is limited to the fluid flow velocity \( v_{fluid} = r_{tube}^2 p_{cable} / 8 \eta l_{cable} \).
Thus a given message requires a travel time \( t_{message} = l_{cable} / v_{fluid} = 8 \eta l_{cable} / r_{tube} \)
\( \text{p}_{cable} ^{0} \text{.001 sec to pass} \) through a 50-micron-long 1-micron-diameter cable; a 10^7 bit message
molecule measuring ~0.4 micron in diameter (Eqn. 7.2) travels at
~5 cm/sec and therefore only transfers the message information at
~10^{-12} bits/sec, or ~5 zJ/bit, near the theoretical minimum. Similarly,
\( t_{message} = 4 \text{ days through a 1-meter-long micron-wide cable at 1 atm}
\) driving pressure (~6 minutes if driving pressure is raised to 1000 atm).

R. Merkle points out that these two restrictions may be ameliorated
by employing a cable transporting monomeric units that can adopt
one of two or more distinct physical conformations. In particular, if
the monomeric units are flat (e.g., small bits of graphite), are held
together by short sections of polyyne (carbyne rods), and if the cable
is formed in a flat housing, then one bit of information can be
transmitted by rotating each monomeric unit into one of two positions
which are separated by 180°. Minimal energy should be required to
rotate a monomer entering the cable input, or subsequently to read
the rotational state of a transported monomer exiting the cable output.

A monomer transport speed of \( v = 1 \text{ m/sec} \) and a 1-nm separation
between successively arriving bit-carrying monomeric units allows
a \( v_{cable} = \text{GHz} \) data transfer rate. From Eqn. 3.19, \( P_{drag} = 10^{-16}
\) watts per monomeric unit giving a total cable power dissipation of
\( P_{cable} = P_{drag} l_{cable} v_{cable} v / V = 10^2 \text{ zJ/bit} \) taking \( l_{cable} = 1
\) meter, \( v_{cable} = 1 \text{ GHz} \) and \( v = 1 \text{ m/sec} \).

7.2.6 Communicytes

A useful supplemental means of information transport through-
out the body is a mobile mass-storage memory device called a
communicyte. Communicytes may serve an analogous function to
postal carriers—messages are delivered to them, passed among them,
and eventually delivered by them to the intended recipients. Mass
mailings to selected recipient subpopulations are also possible. Classes
of communicyte mass memory may include rapidly spooled
hydrofluorocarbon memory tape (Section 7.2.1.1), dense banks of
diamondoid register rods forming a random access memory (RAM)
(Fig. 10.2), or other means (Section 10.2).

In the first case, a 1 micron\(^3\) storage block could contain a 10\(^{10}\)
bit, 0.4-micron\(^3\), 3-meter length of memory tape, with the remaining
0.6-micron\(^3\) volume occupied by read/write heads, spool drivers,
housing and the like. With a read speed\(^{10}\) of ~30 cm/sec (~10\(^{9}\) bits/sec)
the entire data cache may be unloaded in ~10 sec. The energy cost
to rewrite the tape is 1-1000 zJ/bit (Sections 6.5.6 (B) and 7.2.1.8),
therefore the entire tape may be erased and overwritten in 1-1000 sec
assuming a 10 pW tape-handling power budget, giving a write speed
of 10\(^2\)-10\(^3\) bits/sec.

In the second case, register rods of cross-section ~1 nm\(^2\) (described
in Drexler\(^{16}\)) require ~4 nm of length to accommodate a single input/output rod pair and can store at least ~1 bit/40 nm\(^3\). Thus a
0.4 micron\(^3\) block of tightly-packed register-rod RAM may hold
10\(^7\) bits, with the remaining 0.6 micron\(^3\) of the 1 micron\(^3\) storage
block reserved for springs, latches, and other essential mechanisms
(including address-decoding logic) to access the stored data, and
 housings. Read and write speeds are ~10\(^7\) bits/sec; full-cycle energy
losses are ~5 zJ/bit assuming 1 bit/register, \(^{16}\) a ~50 pW power draw
during the ~1 millisec required to read or write the entire onboard
data cache.

Communicytes may also incorporate mechanisms allowing data
transfers among themselves or between themselves and living or
artificial entities (Section 7.4). Communicytes will normally be deployed as components of a communication network of such devices, as described in Section 7.3.2, although they may also be deployed as mobile data repositories or software libraries.

7.3 Communication Networks

To maintain a continuous information flow for sensor data sharing, mission monitoring and wide-area control, it is convenient to establish a communication network in vivo that facilitates point-to-point and broadcast messaging. Two distinct physical network architectures are readily discerned—fiber networks (Section 7.3.1) and mobile networks (Section 7.3.2).

7.3.1 Fiber Networks

Fiber networks may employ fixed-topology data-conducting cables (Section 7.2.5) of virtually any description, including electrical, optical, acoustical, mechanical, or chemical. A complete design analysis of each of these network classes is beyond the scope of this book. The following discussion focuses on high-frequency electrical and optical cables, which most clearly exemplify the salient characteristics of fiber networks.

Suppose that a simple three-dimensional hierarchical Cartesian grid of electromagnetic communication fibers of radius $r_{fiber}$ is embedded in a block of tissue of volume $V_{tiss}$ ($m^3$) with a mean separation between adjacent parallel fibers of $x_{fiber}$ in the rectangular grid. This defines a total of $N_{node} = \frac{V_{tiss}}{x_{fiber}^2} \frac{Xpath}{2}$ intersection nodes and an equal number of cubic voxels, each of volume $x_{fiber}$ ($m^3$) and bounded by $N_{seg} = \frac{3N_{node}}{2}$ fiber segments each of length $x_{fiber}$. Total fiber volume $V_{fiber} = \frac{3\pi}{2} x_{fiber} r_{fiber}^2 V_{tiss} / x_{fiber}^2$ and total fiber length $L_{fiber} = 3 V_{tiss} / x_{fiber}^2$. Allowing a maximum power budget of $P_{net}$ (watts) for the network, then each internodal fiber segment may dissipate up to $P_{net} / N_{seg} x_{fiber}$ (joules/bit) implies a maximum data traffic of $I = P_{net} / E_{bit}$ on each segment between nodes.

For whole-body installation in a human being, $V_{tiss} \approx 0.1 m^3$. For the first subnetwork, $x_{fiber} = 100$ microns, $r_{fiber} = 0.5$ microns, and $P_{net} = 1$ watt, giving total fiber length $L_{fiber} = 3 \times 10^9$ meters, fiber volume $V_{fiber} = 24 cm^3$, and $N_{seg} = 3 \times 10^{11}$ internodal fiber segments defining $N_{node} = 10^{11}$ communications voxels. This network may be emplaced by $10^{11}$ installer nanorobots (Chapter 19) of size $-1000$ microns$^5$ that travel through tissue at 1-100 microns/sec (Section 9.4.4). Each installer unspools -300 microns of internally stored fiber (-240 micron$^3$) in 3-300 sec while consuming $<1$ pW of motive power (Section 9.4.4.2). Hence simple messages may be even faster, so signal passage through at least the local nodes should impose no significant additional delays. Hence simple messages may be passed between any two specific nanorobots in $<0.4$ milliseconds; a complex $10^9$-bit message requires $1000$ sec. By comparison, simple adjacent internodal Internet “pinging” takes $10$ millisecond.

Assigning one nanorobot to each tissue cell ($\sim 10^{13}$ nanorobots) a unique address requires each message packet to contain a log2($10^{10}$) = 44-bit sender identifier plus a 44-bit recipient identifier, mandating a $\sim$10-bit header allowing 12 check bits, the absolute minimum message packet size for a comprehensive cell-addressable network. At a $1$ MHz acoustic nodal access rate, a simple 200-bit message takes $<0.2$ milliseconds to upload or download, plus $<2$ meters/c = $<10^{-8}$ sec to pass through the network at $<0.67c$, where $c$ = $3 \times 10^8$ m/sec (speed of light). Nodal read/write functions should operate at $10^{-9}$ MHz (Section 7.2.6) and optical switching may be even faster, so signal passage through at least the local nodes should impose no significant additional delays. Hence simple messages may be passed between any two specific nanorobots in $<0.4$ milliseconds; a complex $10^9$-bit message requires $1000$ sec. By comparison, simple adjacent internodal Internet “pinging” takes $10$ milliseconds.

Communication with bloodborne nanorobots is most convenient during capillary passage (once every 60 sec), when nanorobots should remain within the operating range of local nodes during the entire transit owing to the narrow width of such vessels (4-15 microns). Capillaries are typically 1 mm long with flow rates of 0.2-1.5 mm/ sec, so the passing nanorobot has 0.7-5 sec to upload or download

* This network design, with each node having 6 connecting fibers, may be unduly redundant, resulting in excess fiber usage. Single nodes may need no more than 4 connecting fibers, forming a tetrahedral grid.
messages. Given the ~0.2 millisecond echo time, a nanorobot entering a capillary can announce its position and begin receiving its email before it has traveled more than ~0.2 micron through the narrowest vessels; messages of up to -5 x 10^6 bits may be sent or received during each capillary transit.

Communication from the arterioles (average ~100 microns diameter) requires contacting the nearest local node which may be 150 microns away, boosting required nanorobot acoustic transmitter power to ~1 600 pW at 100 MHz. Energy costs become prohibitive in the arteries and larger veins and other special locations such as the cardiac chambers, the bursae, and the bladder. Simple 200-bit messages might possibly be transacted by nanorobots immediately adjacent to the 25,000-micron diameter aortal wall; maximum flow velocity of ~1 m/sec carries an unanchored nanorobot beyond the 100-micron operational radius of a medial local node in just ~10^{-4} sec. Under conditions of bradycardia or cardiostasis, communications improve significantly in the vascular regions because bloodborne nanorobots remain longer in the vicinity of nearby embedded nodes.

Subdermal networks may be useful in some specialized applications, but the skin area (~2 m²) is relatively small in comparison with various internal surfaces such as the vascular network (~300 m²).

Fiber nodes may store considerable quantities of useful data. For instance, in the above example there are 10^{11} local network nodes. If each node stores just ~2 600 bits on hydrofluorocarbon memory tape at 26 bits/nm³ (Section 7.2.1.1) accessible in <10^{-5} sec (Section 7.2.6), this tape occupies a volume of ~100 nm³ per node and the entire local nodal network storage capacity is ~3 x 10^{14} bits, or ~one Library of Congress.

7.3.2 Mobile Networks

In a fiberless network, mobile communicytes may be deployed in tissue and blood, and serve as communications nodes. Devices enter a tissue and station themselves for optimum data transfer. All nodal and internodal traffic is acoustic. With a spacing of 100 microns between nodes, continuous internodal message transmission at 100 MHz with f_duty = 1% using an r = 1 micron radiator requires ~600 pW (Eqn. 7.9). From relations given in Section 7.3.1, a whole-body installation with V_tiss ~ 0.1 m³ requires ~10^{11} communicytes uniformly distributed throughout the tissues, a maximum network dissipation of ~60 watts plus another 0.5 watts for node switching losses. Since broadcasts are necessarily omnidirectional for transmitters smaller than ~V_{sound} / ν = 15 microns at 100 MHz, the network is entirely uniform with no backbone and total long-distance message capacity is ~10^6 bits/sec in broadcast mode. Limits on broadcast power can minimize bandwidth overlap and crossstalk between nonadjacent nodes. At 100 MHz and f_duty = 10%, a whole-body network capacity of ~10^7 bits/sec generates an uncomfortable ~600 watts of waste heat; if these communicytes are installed in only <17% of the body volume, then local 10^7 bit/sec networks may be made available at <100 watts. There may be many conceptual similarities to telephonic cellular networks including frequency re-use, handoff algorithms, and the likelihood of large numbers of point-to-point ~10^6 bit/sec communication sessions through different modes without straining the system. Such design details, while interesting, are beyond the scope of this book.

Non-communicyte mobile nanorobots access the network via acoustic channels as in the fiber network. The v_{sound} = 1 MHz/node limit implies that long-distance communications will be sharply curtailed within the tissues. However, local-area communications may have quite satisfactory bit rates. For example, each of 100 nanorobots present in one (100 micron)^3 communications voxel can send continuous ~100 bits/sec messages (using only ~100 Hz of available bandwidth) simultaneously to each of 100 nanorobots located in an adjacent voxel. Thus total monodonal traffic over all 10^{11} nodes can approach the ~10^{18} bits/sec of the optical fiber backbone throughout the entire installation volume. At a 0.01 MHz/channel nodal access rate for each of 100 nearby nanorobots, a simple 200-bit message takes 20 millisecond to upload or download, plus ~200 microns / V_{sound} ~10^{-7} sec to pass between adjacent communications voxels at the speed of sound (V_{sound} ~ 1500 m/sec in soft tissues and whole blood; Table 6.7). Hence simple messages may be passed between any two specific nanorobots in nearby cells in ~40 milliseconds, though a complex 10^5-bit message requires an impractical ~10^7 sec.

Bloodborne communicytes provide a supplementary long-distance messaging capability. Deployment of ~10^{10} 5-micron³ communicytes in a 5.4-liter blood volume (<0.001% concentration by volume) gives a mean interdevice separation of ~80 microns. Communicytes passing through a 1 mm capillary (averaging 10 embedded nodes along its length) in ~1 sec can upload or download ~10^5 bits at each node. Each bloodborne communicyte has a patrol volume of (80 microns)^3 = 500,000 micron³, about ten times the volume of the average capillary, so a fresh communicyte enters each capillary about once every 10 seconds and can receive messages totalling ~10^7 bits during each transit. The fixed patrol volume per bloodborne communicyte implies that these devices will not remain in continuous contact with their bloodborne neighbors while in transit through vessels much smaller than ~80 microns in diameter (e.g., terminal arterioles, metarterioles, capillaries, and the postcapillary and collecting venules; Section 8.2.1). For example, in a 40 micron-diameter terminal arteriole the mean interdevice spacing grows from ~80 microns to ~400 microns, resulting in a temporary communications blackout between bloodborne neighbors during transit through the microvasculature. The blackout may normally last 5-10 seconds (Section 8.2.1.1).

Once received by a bloodborne communicyte, simple long-distance messages are quickly rebroadcast (post-blackout) throughout the bloodstream communicyte fleet (which draws ~6 watts continuous), the signal propagating at near the speed of sound, finally reaching the one communicyte situated nearest to the intended recipient node in at most ~2 meters / V_{sound} ~10^{-3} sec. Assuming a 10 microsec delay between rebroadcasts due to reading a 10-bit routing header at 1 MHz implies an additional rebroadcast delay of at most ~0.2 sec through a chain of ~20,000 communicytes over a ~2 meter path length. Allowing up to ~10 sec to enter and traverse the entire length of the relevant capillary, then long-distance point-to-point messaging requires at most ~11 sec and allows transfer rates of ~10^5 bits/sec or better systemwide. Total long-distance capacity is limited to acoustic bandwidth and the maximum transfer rate of each communicyte, say, at f_duty = 1%, or ~10^6 bits/sec (Section 7.2.6) equivalent to ~0.01% of the entire Internet backbone in 1997. Bloodborne communicytes also provide a useful messaging facility for bloodborne non-communicyte nanorobots.

Besides serving as mobile message repeater stations, bloodborne communicytes may further boost total system capacity by acting as physical message carriers. For example, a single communicyte bearing messages totalling 10^10 bits (Section 7.2.6) circumnavigates the entire vascular circuit once every ~60 seconds. Thus, in theory, point-to-point messages may be carried throughout the entire operational volume at an effective ~10^8 bits/sec rate in this manner, most useful for nanorobots in peripheral tissues attempting to send messages to devices (or postal depots) located in the heart or lungs (organs which the entire blood volume reliably transits once every circuit). (Acoustic downloading at the destination is limited to...
10^9 bits/sec at \( f_{\text{duty}} = 1\% \), but modest locomotion skills and maneuverability would permit physical docking with nodes allowing up to \(-10^{10}\) bit/sec mechanical downloading rates; purely statistical collision-mediated data transfer is not efficient at these low blood-stream concentrations. A total of \(10^9\) (message-carrier) bits may be in transit at any one time in the blood. Except for this message-carrier function and the “blackout” effect in the microvasculature, communicyte network performance is not significantly affected by conditions of bradycardia or cardiostasis.

As in the fiber network, communicyte nodes may store considerable quantities of data. Assuming up to \(-10^{10}\) bits for each of \(10^3\) nodal devices, either network has a maximum total storage capacity of \(-10^{21}\) bits. A single dedicated (4 mm)\(^3\) library nodule (Section 7.3.4) implanted anywhere in the body can also contain \(-10^{21}\) bits.

### 7.3.3 Networks Assessment

Fiber networks excel at high-capacity long-distance communication, with additional capacity readily added to the design by multiplying the number density or bundle density of fibers. Acoustic mobile relay networks have relatively low and fundamentally limited bandwidth. An additional advantage of the fiber network is speed—point-to-point fiber messaging takes \(-10^{-4}\) sec vs. up to \(-10\) sec for the mobile network. Both of these advantages may be important in transdermal communication and whole-body diagnostic monitoring applications.

However, fiber and mobile networks may provide roughly equivalent local transmission speed and capacity if both must employ acoustic nodal access, as this provides the highest endogenously transmitted bit rate per unit power consumption for mobile nanorobots that are arbitrarily positioned network users. Thus in applications requiring predominantly local information transfers, the choice of basic network architecture may be driven by factors other than system total capacity and speed.

Fiber networks also have many significant disadvantages. Both fibers and mobile nanodevices can be made immunocompetent and biocompatible, but fibers are physically vulnerable and can cause serious physical tissue and cellular irritation (Section 6.4.3.6). It is well-known in dermatology that subdermally-placed, but incompletely absorbed, “dissolvable” stitches can work their way back to the surface of the skin; fiber network elements could be similarly ejected from the body. Some of these mechanical difficulties possibly may be overcome using a helical coil or a fiber which is coupled to a biocompatible sheathing using compliant internal springs; large-diameter fibers could be installed within bones where they would experience minimal flexing and immune system exposure, although transmissions across joints or tissues with high flexure or vascularization might still be required to be acoustical.

Networks with mobile components are easier and quicker to install, to reconfigure, to upgrade while in vivo, or to remove in the event of malfunction or change in medical objectives, than fiber networks with immobile components. On the other hand, communication protocols may need to be substantially more complicated to provide the necessary reliability with drifting nodes.

The fixed physical positions of fiber network nodes is also a major disadvantage because tissue movement is ubiquitous inside the human body, causing progressive fiber misalignment. Examples of such movement include untreated metastasizing tumors or benign growing tumors, angiogenesis in injured or heavily exercised tissue, enlargement and shrinkage of tissue channels or prelymphatics after wounding or edema, adipocyte deposition, and the breakdown or buildup of muscle tissue; cell loss and tissue modifications in the brain, sclerotic livers, radiation-damaged bone marrows, chances and blisters, hematomas, and the uterus and related tissues during pregnancy; normal periodic movements in muscle-laden tissues such as the biceps, cardiac tissue, arterial walls, and the thoracoabdominal diaphragm; and other nonpathological but irregular movements such as tissues which may rapidly inflate with fluid, including the sex organs and related tissues, the bladder, and to a lesser extent the lungs. With such movement, nodal positional assignments quickly become obsolete and nodes may be carried out of acoustic range. Fiber spatial distributions may develop undesirable clumping and rarefactions. Fiber-embedded tissues which slough off in the normal course of events (e.g., epidermal, endometrial, placental, gastrointestinal) may carry fibers out into the uterine or intestinal canals, or expose them to air in the case of the epidermis.

The paradoxical conclusion is that overall fiber network reliability may be improved by adding nodal mobility. One crude biological analog is the filopodia of embryonic axons which can drag their fibrous neural cargo through embryonic tissue at \(-0.01\) micron/sec.\(^{279}\) While adult tissue may in theory have a more solidified extracellular matrix (ECM) and greater intercellular connectivity, making dragging more difficult, leukocyte and fibroblast mobility through ECM is typically 0.05-0.70 microns/sec (Section 9.4.4.2).

### 7.3.4 Dedicated Communication Organs

By analogy to implanted energy organs (Section 6.4.4), dedicated macroscopic organs may be installed in the human body to facilitate internal communications. Such organs could serve as high-capacity data storage, computation, and retransmission nodes in internal networks; as storage bays or reprogramming facilities for communicytes; as transdermal links to the outside world to permit rapid inmessaging and outmessaging, offloading of major data processing tasks to ex vivo computational resources or to dedicated computational organs (Section 10.2.5), or integration with external sensory, communications, and navigational facilities; as rf, TV, or satellite broadcast receivers, for example to receive Global Positioning System (GPS) signals or personal email; as radio signal transmitters for outmessaging; as convenient foci for multisystem or autogenous control, coordination, or data distribution; or as library nodules or personal medical history information storage devices.

### 7.4 Communications Tasks

Nanomedical messaging tasks fall into three broad categories:

1. nanorobots communicating with externalities including physicians, laboratory machines, external computers, or even the conscious perception of the patient himself;
2. nanorobots communicating with other nanorobots; and
3. nanorobots communicating with human organs, tissues, or cellular systems.

Communications data flowing into a nanorobot from some extrarobotic source is called inmessaging; outward-bound data transfers are called outmessaging. A comprehensive examination of all these possibilities is beyond the scope of this book. The present discussion is limited to a brief overview of the following elementary situations: inmessaging from external sources (Section 7.4.1), inmessaging from patient or user (Section 7.4.2), intradevice messaging (Section 7.4.3), interdevice messaging (Section 7.4.4), biocellular messaging (Section 7.4.5), outmessaging to the user (Section 7.4.6), outmessaging to an external receiver (Section 7.4.7), and transvenous outmessaging (Section 7.4.8). Some of the examples presented may
appear whimsical from a 20th century perspective, but they illustrate the incredible variety and power inherent in the technology. A discussion of user interfaces is largely deferred to Chapter 12.

7.4.1 Inmessaging from External Sources

Inmessaging occurs when information is conveyed from a source external to the human body, or external to working nanodevices, to a nanorobotic receiver located inside the human body. Security protocols (Chapter 12) are a strong necessity here. The receiver may be associated with medical nanodevices, either individually or collectively, or with specialized communications organs.

Methods of modulating power sources to achieve data transmission have already been examined at length in Section 7.2. To summarize a few of the many possibilities—commands, data, software, processed results of extensive computations, and messages from medical personnel may be conveyed into the human body by:

A. Chemical Inmessaging — communicytes or messenger molecules encased in pills or suspended in ingestible liquids, injectable or IV fluids, mouthwashes, earwashes, eyedrops, inhalants, nasal sprays, dermal patches, suppositories or enemas; or artificially vascularized data secretion glands transdermally linked to external control mechanisms.

B. Acoustic Inmessaging — dermal ultrasound transducers including handheld units, cuffs or bracelets, or transmission through water baths; vibrating treatment tables, beds or chairs; ingestible sonosondes; osteoimplanted acoustic radiators or implanted dental vibrators with rf or direct cable external links; or implanted circumvascular acoustic radiators.

C. Electromagnetic Inmessaging — broadcasts from external rf and microwave sources including epidermally-placed electrodes (Section 4.7.1); optical laser arrays pressed against the skin; ingestible radiosondes and radio telemetry pills; whole-body electrical current flows including transosteal intercellular transmissions; transdermal multi-tesla oscillating magnetic fields; electromagnetic emanations from implanted rf or microwave emitter organs under external control; or direct linkages to internal networks via permanent transdermal access ports.

D. Cable Inmessaging and Message Depots — information may be transported directly into the body through chemical, acoustic, electrical, optical, or mechanical cables or tethers feeding implanted communications organs, message depots, data distribution centers or bulletin boards; such depots may have biocompatible termini in the bloodstream, brain, or intercellular regions to facilitate message exchanges with mobile nanorobots.

E. Eavesdropping — continuous nanorobot monitoring of normal physiological sensory traffic (imported by human sensory organs) with specific sensitivity to defined control or keyed patterns that might have been inserted into the natural data stream, such as modulated visual, auditory, or tactile cues, periodic dermal pain or temperature sensations, or oscillating taste and scent signals.

F. Macrosensing — external modulation of physiological or environmental parameters which can be directly sensed by in vivo nanorobots but controlled by external agencies, including pharmacologically-modulated heartbeat or respiration rate, mental alertness or stage of sleep, atmospheric humidity or pressure, chemical composition of the air, gravity or centrifugal acceleration, particulate radiation flux, or geographical position and altitude (see Section 4.9).

7.4.2 Inmessaging from Patient or User

Autogenous command and control (Chapter 12) gives a patient or user the ability to issue instructions directly to nanodevices present in his or her body. Patient or user inmessaging is an important enabling technology for autogenous control, wherein a human being manipulates a consciously controlled variable, which manipulation is then detectable by in vivo nanodevices. Once the message is received by even a single nanorobot, the information may be rapidly multiplied and passed along to other nanorobots using an internal communications network.

7.4.2.1 Mechanical Inmessaging

Perhaps the simplest example of user inmessaging is finger-drumming on a hard surface such as a tabletop, a method which we will now examine in some detail. A ~15 gm finger may be drummed at a maximum of ~5 Hz over a 2 cm path length at a maximum velocity of ~0.2 m/sec; if the impacting finger is brought to a halt within 1 mm of the -1 cm² fingertip and ~50% of the kinetic energy is converted into acoustic energy, the phalangeal impact produces a maximum compression wave of ~0.3 atm (~300 watts/m²). The lightest audible tapping produces a pressure wave of ~0.01 atm (~0.3 watts/m²). Either pulse is easily detectable by ~10⁻⁶ atm-sensitive nanorobot acoustic sensors (Section 4.5.1) located in the patient’s hand or arm.

Consider five sensor-equipped nanorobots each positioned ~2 mm below the epidermis in the soft tissue just beyond the distal phalanx of five different fingers, and each with access to a local communications network (Section 7.3) spanning the hand. An impact pulse applied to a given finger reaches the sensor stationed in that finger in ~1 microsec by conduction through soft tissue at ~1540 m/sec. The acoustic pulse then travels down the phalanges and metacarpals, crosses the distal row of carpals, then travels back up through the metacarpals and phalanges of the other fingers mostly via bone conduction at ~3550 m/sec (Table 6.7), reaching the adjacent sensors 90-110 microsec after the initial impact. The pulse does not travel exclusively in bone as it must pass through 8-11
Table 7.1. Temporal Signature of a Detectable Mechanical Pulse Conducted Mostly through Bone, which is Applied to One Finger and Received by Nanorobot Sensors Distally Located in Each of the Five Fingers of the Hand

<table>
<thead>
<tr>
<th>Finger #1 (Thumb)</th>
<th>Finger #2 (Index)</th>
<th>Finger #3 (Bird)</th>
<th>Finger #4 (Ring)</th>
<th>Finger #5 (Little)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulse is detected by sensor on finger #:</td>
<td>Arrival Time $X_{path}$ (cm)</td>
<td>Arrival Time $X_{path}$ (µsec)</td>
<td>Arrival Time $X_{path}$ (cm)</td>
<td>Arrival Time $X_{path}$ (µsec)</td>
</tr>
<tr>
<td>1</td>
<td>0.2</td>
<td>1</td>
<td>34.2</td>
<td>97</td>
</tr>
<tr>
<td>2</td>
<td>34.3</td>
<td>97</td>
<td>0.2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>37.5</td>
<td>106</td>
<td>34.8</td>
<td>98</td>
</tr>
<tr>
<td>4</td>
<td>39.5</td>
<td>111</td>
<td>36.5</td>
<td>103</td>
</tr>
<tr>
<td>5</td>
<td>38.7</td>
<td>109</td>
<td>36.8</td>
<td>104</td>
</tr>
</tbody>
</table>

interphalangeal amplitude attenuation provides an additional check on the identity of the tapping finger. An impact deceleration in $-5 \times 10^{-3}$ sec produces a single $-5$ millisec pulse ($-200$ Hz) which from Eqn. 4.52 attenuates from $10,000 \times 10^{-6}$ atm to at most $9997 \times 10^{-6}$ atm over a $20$ cm path length through soft tissue, or to at most $-9900 \times 10^{-6}$ atm over a $35$ cm path length through bone—either of which is detectable by nanorobot acoustic sensors sensitive to changes of $-1 \times 10^{-6}$ atm.

Many conceptually similar methods of mechanical inmessaging may be employed by a patient whose in vivo nanorobots have been appropriately programmed and positioned, including clapping/slapping of hands/feet ($-3$ bits/sec), mandibular clicking ($-2$ bits/sec), tongue-tapping on palate ($-2$ bits/sec), eyelid or eyebrow flapping ($-2$ bits/sec), dental/dermal drumming using a rigid object ($-1$ bit/sec), epiglottal pulsing ($-1$ bit/sec), thoracoabdominal diaphragmatic contractions ($-1$ bit/sec), or toe tapping ($-0.5$ bit/sec).

Other mechanical stimuli at the skin—including pressure, touch, vibration, and tickle (see Schmidt$^{[635]}$)—may also be eavesdropped by medical nanorobots. Interestingly, the frequency response of these receptors may range as high as $800$ Hz for Pacinian corpuscles (sensory organs located at the skeletal joints, tendons, and in the muscles;$^{[634]}$ Table 7.3). Thus an eavesdropping subdermal nanorobot could receive $\geq 100$ Hz modulated signals through this channel that would only be perceived as “flat stimulus” by the conscious human patient.

7.4.2.2 Kinesthetic Inmessaging

With the support of an internal communications and navigational network, proprioceptive nanosensors selected in recent human tissues permit the direct detection of gross limb positions, limb velocities, and body orientation in space with $t_{mean} < 10^{-3}$ sec or better (Sections 4.9.2 and 8.3.3). Thus, for example, data may be rapidly inmessed using a highly literal manual sign language,$^{[297]}$ designed for minimum ambiguity and maximum precision, with a minimum number of discrete symbols to memorize. (Native American sign language has $-400$ root signs and other gestural languages have $-2000$ signs,$^{[731]}$ though a manual alphabet is also used.) Fingers and limbs can be comfortably moved at $0.1-1$ m/sec; if $-1$-cm positional increments can be adequately controlled by a patient and thus can be employed to convey one bit of information, then the maximum channel capacity of signing is $10-100$ bits/sec per finger or $100-1000$ bits/sec for all ten fingers. This estimate compares favorably with the $-100$ bits/sec achieved by hearing-impaired signers translating speech at a conversational rate of $-150$ words/min, and also with the record manual typing speed of $216$ words/minute ($-173$ bits/sec) achieved on an IBM Selectric typewriter.$^{[739]}$ Nanorobots may also eavesdrop on proprioceptive information detected by Pacinian proprioceptors.
Properly monitored voluntary large body motions such as break-dancing or karate programs, or lesser displacements such as head rolling or shrugging movements, leg jiggling or calisthenic exercises, rapid periodic lung inflations, anal sphincter contractions, blinking or eyeball rotations can similarly transfer useful information into the body at ~1 bit/sec. In combination with real-time retinal displays (Section 7.4.6.5), eyeball movement nanosensors can allow patients to use their eyes as an “ocular mouse” to point to an icon or array of alphanumeric characters superimposed on the visual field. This will allow them to initiate a programmed function or to spell out words or numbers. In 1997, an analogous system, known as the Ocular Vergence and Accommodation Sensor (OVAS) manufactured by Applied Modern Technologies Corp. of Garden Grove, California, used 12.5 W/m² IR laser beams reflected from each eye’s retina to provide information about eye movement and other biometric data. The OVAS design included a 12-component optics system and a Pentium processor with algorithms for processing data on the accommodative state, movements, and vergence of the eyes, as well as 10 other ocular functions. 

In 1998, Canon EOS cameras used eye-control focus sensors to detect where in the frame the photographer was looking, and autofocus there, instead of defaulting to the center of the frame. Superimposing retinal displays over normal vision need not prove confusing to the patient. Experiments with head-mounted displays show identical ergonomic defaulting to the center of the frame. Superimposing retinal displays can also be used for high-capacity inmessaging (Section 7.4.6.3), allow patients to use their eyes as an “ocular mouse” to point to an icon or array of alphanumeric characters superimposed on the visual field. This will allow them to initiate a programmed function or to spell out words or numbers. In 1997, an analogous system, known as the Ocular Vergence and Accommodation Sensor (OVAS) manufactured by Applied Modern Technologies Corp. of Garden Grove, California, used 12.5 W/m² IR laser beams reflected from each eye’s retina to provide information about eye movement and other biometric data. The OVAS design included a 12-component optics system and a Pentium processor with algorithms for processing data on the accommodative state, movements, and vergence of the eyes, as well as 10 other ocular functions. 

In 1998, Canon EOS cameras used eye-control focus sensors to detect where in the frame the photographer was looking, and autofocus there, instead of defaulting to the center of the frame. Superimposing retinal displays over normal vision need not prove confusing to the patient. Experiments with head-mounted displays show identical ergonomic efficiency and accuracy for pointing tasks when control characters are projected either on a dark background screen or on a translucent background through which moderately distracting objects and moving traffic are visible.

7.4.2.3 Acoustic Inmessaging

A wide variety of sounds may be generated voluntarily by the patient and be detected by in vivo nanorobots, such as coughing, grunting, humming, various other mechanical body noises (Section 4.9.1.4) and even vocalizations including direct voice commands spoken aloud by the user (Section 4.9.1.5); 1-100 bit/sec inmessaging rates appear feasible by these methods. Intracochlear nanodevices can also be used for high-capacity inmessaging (Section 7.4.6.3).

Soundless subvocalizations may also be detected using kinesthetic techniques (Section 7.4.2.2) to measure ~0.5 mm movements in tongue position and the muscles of the larynx that produce adjacent distinguishable tones in the vocal cords. The triangular waveform harmonics emitted by the larynx are modified during passage through sub- and supraglottal cavities, the shape of which may be voluntarily altered by submillimeter muscular movements — again, well within the range that nanorobots can detect (Section 4.9.2.1). Data sharing by multiple nanorobots using a communication network should permit reconstruction and estimation of the intended vocalization with sufficient reliability to transfer useful noncritical data or commands into the network. In 1998, radar microchips operating at ~2 GHz were already used for soundless speech recognition—the system recognized and interpreted motions of the human glottis, lips, tongue, jaw, and velum in real time.

The field of voice recognition is active, and no systematic review will be attempted here. In 1998, IBM’s speech recognition technology, including the Voice Type Dictation and Personal Dictation systems (IPDS), was well-known and commercially available, as was Kurzweil’s VoiceCommands and VoicePro and NaturallySpeaking from Dragon Systems.

7.4.2.4 Chemical Inmessaging

Ingestion or inhalation of chemical messenger molecules is another way the patient can control the flow of information and commands into his body. The patient could provide the desired data or commands to a desktop manufacturing appliance, which appliance then encodes the information onto billions of copies of a hydrofluorocarbon messenger molecule suspended in a carrier fluid or solid pill, which the patient subsequently inhales or ingests. Messages up to ~10⁷ bits may be distributed throughout the body in ~one circulation time (~60 sec), a data transfer rate up to ~10⁵ bits/sec. Patients without access to personal manufacturing appliances could be given fluid ampules, standardized pill sets or inhalers whose contents are designed to trigger specific predetermined nanorobot behaviors. Such sets must contain personal security codes to prevent their unauthorized use on other patients (see Chapter 12).

A related technique for chemical inmessaging is to allow naturally inhaled or ingested molecules (e.g., gustation or olfaction) to trigger specific nanorobot actions. For example, if continuously circulating detoxification nanorobots are programmed to begin removing alcohol from the bloodstream whenever the presence of garlic compounds is detected, then the patient can voluntarily trigger the detox response by sniffing or chewing a clove of garlic. (A population of 10¹² bloodborne nanodevices with aggregate storage volume ~6 cm³ can reduce serum alcohol from 0.2% to 0.005% in ~1 second by prompt onboard sequestration, followed by metabolism of the entire inventory in ~10 minutes within a systemic caloric budget of ~200 watts, although outflows from ethanol-soaked body tissues into the bloodstream and other factors complicate the process; see Chapters 19 and 25.) Similarly, nanorobots with selected chemical sensors stationed in the vomeronosal organ can vallate papillae or olfactory epithelium could directly detect and respond to specific flavors, perfumes, pheromones or other chemical substances—even molecules which humans are normally incapable of tasting or smelling, including artificial or engineered substances.

Binary message sequences employing the five basic gustatory stimuli (Section 7.4.6.4) must use one of the stimuli as a positional spacer to avoid confusion. For example, a patient sipping the 8-taste sequence “sweet (1), salty (2), sweet (1), salty (6), sour (0), salty (2), sweet (1), salty (7)” has just transferred the binary sequence “1101” into his body. (“Salty” is used as the spacer because salt sensitivity is most widely distributed over the surface of the tongue. “Sour” is used because saltiness partially suppresses perception of bitterness, and because sweet and bitter taste messages are transmitted by the same transducer protein, called α-gustducin.) This inmessaging modality is probably limited to ~0.1 bit/sec but may be boosted considerably by employing complex artificial gustatory stimuli (akin to messenger molecules) which only lingually-embedded nanorobot sensors can detect.

Human olfactory receptors can probably detect ~1000 different smells (Section 7.4.6.4); exposure to a new olfactory stimulus every 1-10 seconds, with each smell worth log₁₀(1000) = 10 bits/symbol, allows information transfer up to ~1-10 bits/sec. Individual chemical assay nanorobots that can detect ~10⁵ different molecules (Section 4.2.4) allow greater diversity and specificity of signal carriers, but at most only ~double the bit rate because log₁₀(10⁵) = 23 bits/symbol. However, the bit rate using long-chain digitally coded artificial messenger molecules is vastly higher, and probably only thermogenically (heat of manufacture and access; Sections 7.2.1.8 and 7.2.6) and biocompatibility (Chapter 15) limited.

7.4.2.5 Electromagnetic and Thermal Inmessaging

The human body does not emit voluntary electrical or electromagnetic radiation except for modest dermal potentials associated with ECGs and EEGs which probably cannot be directly detected by individual in vivo nanorobots (Section 7.2.3). Simple voluntary
actions that might allow electromagnetic inmessaging to subdermally-implanted nanorobots include placing a hand near a heated stove, then pulling it away; or blinking a flashlight beam on and off while the flashlight is pressed against the skin, for nanorobots positioned at most 1-2 cm below the epidermis (Section 4.9.4); or moving a bar magnet across a patch of skin (Section 4.7.2)—all allowing ~0.1-1 bit/sec transfer rates. A thermographic system might also detect the slightly elevated dermal temperature caused by touching a hand to the face or thigh, much like a simple demarcation strategy (Section 8.4.1).

7.4.2 Neural Inmessaging

The most comprehensive inmessaging capability may be afforded by eavesdropping on selected neural sensory traffic (Section 4.9.5). Noninvasive monitoring of neuroelectric impulses (Section 4.8.6) by nanorobots stationed near afferent neurons can detect and interpret any sensory stimulus that can be consciously perceived by the patient. Virtually every inmessaging stimulus described elsewhere in Section 7.4.2—including finger-drumming, hand-signing, vocalized or olfactory signals, dermal heat sensations, scalp/eyebrow myopotentials, and the like—may be indirectly detected by suitably positioned and configured neural monitors. For instance, finger-tapping generates electrical activity in a specific 1 cm² patch of neurons in the primary motor cortex; 1090 although nanorobots suitably positioned in the brachial plexus, or lower in the tree near critical junctions of the median nerve (above the digital, thumb, and palmar cutaneous branches) and the ulnar nerve (above the muscular, dorsal cutaneous, superficial palmar and deep palmar branches), may achieve more rapid and reliable monitoring of motor impulses. Neurons may fire at >100 Hz, so, ignoring significant electrical spike redundancy, inmessaging rates up to ~100 bits/sec per monitored neuron can in theory be achieved with this approach.

Nonsensory volitional neural inmessaging channels are also available to nanorobots monitors stationed in the brain. For example, by consciously recalling a specific series of numbers, letters, names, words, concepts, images, or events, a well-defined constellation of neural impulses of certain frequencies may be caused to pass repeatedly through the same, say, ~200 micron³ blocks of brain tissue for each recalled item. Brain tissue contains numerous “dominance columns” on approximately this scale which appear to represent functionally associated fiber bundles. 1036,1037 For example, 300-500 micron wide columns have been described in the striate and inferotemporal cortex of the macaque monkey; 771 organized patterns of action potential waves have been observed via calcium imaging in rodents, measuring 100-300 microns wide in the retina and ~50 microns wide in the cortex. 771 An experiment using a context-recall memory task in monkeys suggested that an ensemble of as few as 16 motor cortical cells could perfectly classify (i.e., 100% accuracy) all the items in a sequence of five stimuli; 318 ν nanorobotic monitors stationed at those 16 cells could in theory discriminate all five-stimulus sequences.

Thus, without being able to understand any of the higher-order meanings of these signals (e.g., to “read human thoughts”), nanorobots detecting such reproducibly neurographically-localized impulse trains may pool their readings and interpret specific combinations of such localized trains as data or as command strings (a ~0.1-1 bit/sec channel) and then take appropriate action. Biofeedback monitors, 166 ν volitional scalp EEG multiband user interfaces, PET scanners and functional magnetic resonance imaging (fMRI) operate on a crudely analogous principle, but without the benefit of precise impulse localization that is made possible by molecular nanotechnology. In 1998, a glass cone electrode filled with neurotrophic factors (which encouraged neural ingrowth) was implanted in a patient’s motor cortex; the patient learned to mentally operate the output signal as a binary switch. 1074 Another system using cortical potentials achieved transfer rates of ~2 characters/min. 4921

7.4.2.7 Macroscale Inmessaging Transducers

Inmessaging signals from patient to nanorobot may also be mediated using macroscale communications interfaces. These interfaces may be fixed or mobile, and may be either stable or reconfigurable. A crude example of a fixed, stable inmessaging transducer would be a ~1 cm dermally- or cranially-mounted coaxial cable connector which allows data to flow into an implanted dedicated communication organ for subsequent distribution to in vivo nanodevices via an internal communications network. Data passes into the patient through a simple rf or microwave coax cable which is plugged directly into the patient’s socket, readily achieving ~10⁹ bit/sec transfer rates (Section 7.2.5.1) directed from a computer operated by the patient or physician. Up to ~10¹³ bit/sec transfers are available using fiberoptic (Section 7.2.5.2) or chemicable (Section 7.2.5.5) ports. Other fixed transducers may incorporate photoelectric or olfactory sensors, microphones, neural taps, or manipulable tactile elements such as buttons or switches.

At the other extreme of maximum inmessaging versatility, a dermal transducer may be configured upon request and presented to the patient on any part of his body. For example, a tattoo-like ~30 cm² dermal control panel comprised of ~3 billion communication-linked kinesthetically-sensitive chromomorphological nanorobots spaced ~5 microns apart could be assembled under a clear flat patch of epidermis such as the back of the hand. This panel could take the shape of a touch-sensitive alphanumeric keypad resembling a laptop computer keyboard or a touch-tone telephone or home-security keypad, complete with a touch-sensitive grid of visually-readable inscribed characters. (See details in Section 7.4.6.7.) Similarly, a wristwatch-shaped device could accept and interpret verbal commands, then translate them into modulated electromagnetic signals displayed on a 2-D array of photodiodes or microlasers on the back of the watch; the signals could then be received by a nanorobot dermal transducer panel (Section 7.4.6.7) and passed along to the rest of the in vivo network.

7.4.3 Intradevice Messaging

Intradevice messaging has already been briefly described in the context of metamorphic control (Section 5.3.3), lengthy gear trains (Section 6.3.2), sliding and rotating control rods (Section 7.2.4), acoustic transmission lines (Section 7.2.5.3), and mechanical cables (Section 7.2.5.4), and does not appear to present any major fundamental or conceptual difficulties.

To summarize, internal communications within an individual nanorobot may be achieved by impressing modulated low-pressure acoustical spikes on the hydraulic working fluid of the power distribution system, or by mechanically actuating simple networks of rods and couplings, easily achieving transfer rates of >10⁹ bits/sec. 10 Single-channel mechanical, hydraulic or acoustic cables ~10 nm in diameter can be bundled together into a flexible 100-nm diameter 100-channel data bus suitable for controlling flexible manipulators or for maintaining control during metamorphic shape shifts; nested doped buckytube fibers could provide a similar multiplicity of electrical conduction channels. Internal fluid chambers can distribute pressure actuation signals in the manner of the latching ratchet/pawl and vernier ratchet threshold-active pressure actuator mechanisms proposed by Drexler. 10
**7.4.4 Interdevice Messaging**

Physically-linked nanomachines may communicate with each other through their tendons or network connections. Messaging between physically separated in vivo nanodevices will normally take place via acoustic or chemical, and occasionally mechanical or optical, modalities.

Interdevice communications has been extensively discussed in both free-tissue (broadcast) and cable contexts for chemical messaging (Sections 7.2.1 and 7.2.5.5), acoustic messaging (Sections 7.2.2 and 7.2.5.3), electromagnetic messaging (Sections 7.2.3 and 7.2.5.1-2), mechanical tapping and cable messaging (Sections 7.2.4 and 7.2.5.4), metamorphic surface reconstructions (Section 5.3.2.5) and network configurations (Section 7.3). The general conclusion is that untethered free nanorobots will typically have maximum interdevice communication ranges on the order of -100-200 microns; interdevice messaging much beyond that distance probably must be handled by a local or wide-area communications network.

**7.4.5 Biocellular Messaging**

Nanodevices can also directly communicate with cells by intercepting, modulating, or initiating the flow of natural chemical messenger molecules. Because cells are the fundamental unit of biological organization, it is useful first to examine the various modalities of cellular chemical communication in general terms (Section 7.4.5.1). Properly designed and positioned nanodevices can receive chemical messages from cells (Section 7.4.5.2), send chemical messages into cells (Section 7.4.5.3), modulate natural cellular message traffic (Section 7.4.5.4), or communicate directly with neurons (Sections 7.4.5.5 and 7.4.5.6). Information can also be transferred mechanically into the cell via links between the extracellular matrix (ECM) and the cytoskeleton in ways that have yet to be fully elucidated.

**7.4.5.1 Natural Cellular Communications**

Cells communicate chemically with each other in two ways—contact signaling and secretory signaling. In the first category, chemical messages may be passed between cells in direct contact through gap junctions that directly join the cytoplasm of the interacting cells (Section 5.4.2). As in the case of foreign or self-antigen presentation (Sections 5.3.6 and 8.5.2.1), cells may also display plasma-membrane-bound molecules on their surfaces which may be "read" by other cells that come into physical contact with them. Signals may also be transduced into cells directly from the ECM.

In the second category—the principal focus of this Section—there are four basic classes of secretory signaling molecules in the human body, as described briefly below and incompletely listed in Table 7.2. Normal human bloodstream concentrations of some of these molecules are given in Appendix B. At least several hundred different chemical messengers are known to be used in the human body, and many thousands may exist. Biologist Dennis Bray of Cambridge University believes that up to ~50% of the human genome may code for proteins involved in cell signaling, but 10%-20% for "classical" communication pathways is probably more reasonable. For example, the now-sequenced 97-megabase genome of the nematode Caenorhabditis elegans has ~19,099 protein-coding genes, of which ~11% are involved in signal transduction and another ~11% in transport and secretion.

The first class of secretory signaling molecules is comprised of the endocrine hormones, which are usually secreted by exocytosis and travel through the bloodstream to distant target cells (e.g., "volume transmission") where they act as direct effectors on cell activity. For example, a steroid hormone, upon entering a target cell, binds to a complementary receptor protein. Binding induces conformational changes in the receptor and activates it, increasing the receptor's affinity for DNA and allowing it in turn to bind to specific genes in the nucleus, thus directly regulating their transcription. This process is extremely wasteful because many activated receptors bind to DNA at sites where their presence has no effect; notes that activated receptors have a tenfold increased affinity for nonspecific DNA.

The second class consists of local chemical mediators or paracrine, which travel only a short distance (up to ~1 mm) to neighboring cellular targets. This group also includes the autocrines which act upon a cell's own receptors, such as the prostaglandins which are made by cells in all tissues. The single largest group in this class is the cytokines. Cytokines are soluble proteins or glycoproteins, produced by leukocytes and other cell types, which act as chemical communicators between cells but not as effector molecules in their own right. Cytokines include secretory peptides formerly classified as lymphokines, monokines, interferons, colony-stimulating factors, chemotactic factors and growth factors; over 150 different cytokines had been cloned by 1998.

The third class of chemical signaling molecules is the neurotransmitters, including the 60+ known neuropeptides (typically up to 30-40 residues long), which mediate communications between nerve cells and other cells. Interestingly, a single neurotransmitter signaling molecule may have different effects on different target cells. For instance, acetylcholine stimulates the contraction of skeletal muscle cells but decreases the contraction rate and force in heart muscle cells. The immune and neuroendocrine systems produce and respond to similar signal molecules (e.g., neuropeptides and cytokines). Neuropeptides have been shown to effect immune responses through their influence on cytokine production and action, and cytokines are known to induce or influence the production of peptidergic messenger substances.

The fourth class is the intracellular messengers, also known as intracellular mediators or second messenger substances. While small hydrophobic signaling molecules like the steroids and thyroid hormones pass through the cell membrane and activate receptor proteins inside the cell, hydrophilic signaling molecules such as neurotransmitters, most hormones and local chemical mediators activate receptor proteins on the surface of the target cell. Cell surface receptors for these signaling molecules may be either diffusely distributed or localized to specific regions of the membrane, and their number can vary from 500 to more than 100,000 per cell for a specific ligand.) Because cell surface receptors cannot regulate gene expression directly, they must emit a second messenger molecule into the cytosol which can then trigger the desired regulatory action. Cyclic AMP (cAMP), an intracellular messenger that tells certain types of cells to break down their glycogen stores in response to receipt of epinephrine at the cell surface, is perhaps the best-known example (Section 7.4.5.4). Another example is the voltage-gated Ca++ ion channel (Sections 3.3.3 and 7.4.5.3), which participates in nerve pulse events but also in signal transduction via the inositol-phospholipid pathway for more than 25 different types of cell surface receptors of hormonal signals. For instance, excitation-contraction coupling in skeletal muscle cells requires the release of Ca++ ions through ryanodine receptor channels in the SR (sarcoplasmic reticulum; Section 8.5.3.5), each of which opens to provide 4-8 picocamp Ca++ ion currents in 1-2 millisecond bursts, through the SR membrane.
7.4.5.3 Outmessaging to Cells

Nanodevices may receive the natural chemical messages transmitted from and between cells simply by eavesdropping on the natural molecular message traffic. (Of course, mechanical (e.g., cytoskeletal) and electrical (e.g., ionic) cellular emanations may also be detectable by nanodevices.)

For instance, consider a -0.1 micromole puff of epinephrine (aka adrenalin, C9H13NO3, MW = 183 daltons) released into the human bloodstream from cells in the adrenal medulla located above the kidneys. Epinephrine is an emergency hormone emitted in response to stressful conditions which acts to increase heart rate, decrease the blood flow to the gut, increase blood flow to skeletal muscles, increase circulation of oxygen and nutrients to other organs, and mobilize the production of metabolic fuels such as fatty acids and glucose (by causing liver and muscle cells to break down glycogen)—establishing the primitive mammalian “fight or flight” response.

Bloodstream concentrations rise over 1-10 sec to a peak of \(e_{\text{ligand}} \sim 10^{-8}\) molecules/nm\(^3\) (Appendix B), then from Eqn. 4.3 a bloodborne nanorobot with a single epinephrine chemical sensor having just -0.4 nm\(^2\) of active surface area can detect the peak concentration in \(t_{EQ} \sim 0.3\) sec. A patch of 260 such receptors measuring \(-500\) nm\(^2\) in total physical surface area can detect the peak concentration within \(-1\) milisecond of its occurrence, or can detect 10% of peak concentration within 10 miliseconds from the start of the ramp-up. Epinephrine and most other signaling molecules cannot be directly measured by nanorobots stationed within the cytosol because the molecule is hydrophilic and cannot pass through the cellular membrane (unlike the steroids). For the most rapid possible response to changing hormone levels, chemoensor-equipped nanorobots should be positioned on or inside the cells and organs which produce specific signals of interest, with the results of positive detections of increased activity quickly passed along through the in vivo communications network (Section 7.3).

For the rarest cytokines or steroid hormones present at \(e_{\text{ligand}} \sim 10^{-12}\) molecules/nm\(^3\) in the bloodstream or cytosol (equivalent to \(-8\) molecules inside one (20 micron\(^3\)) tissue cell or \(-1\) molecule inside a single leukocyte), a \(-5000\) nm\(^2\) sensor patch containing \(-100\) individual receptors (each having \(-6\) nm\(^2\) of active area) can register a detection in \(-10\) sec for a single cytosol-based nanorobot. However, cellular signaling molecules are more commonly present in the cytosol at \(-10^{-7}\) molecule/nm\(^3\) levels,\(^{531}\) enabling \(-1\) millisecond detections using this sensor configuration. By comparison, the typical target tissue cell has \(-10,000\) steroid receptors (each binding one steroid molecule) floating in the cytosol.

Nanodevices may also eavesdrop on morphogenic signals. For example, simple changes in the extracellular protease/antiprotease concentration ratio (which causes endothelial cells to detach from their parent vessel and invade their underlying stroma during angiogenesis) are readily detected using chemical nanosensors. Continuous monitoring of the relevant natural cytokine traffic can help nanorobots maintain a complete picture of current cellular activities and states. Indeed, the principal design challenge is likely to be sensory traffic data management rather than sensor speed or specificity. Living cells are awash in a sea of messages from hormones, cytokines, growth factors, and a host of other message-carrying molecules, all of which must be transmitted through complex chemical networks of intermediates and multiple signal pathways into the cell’s interior.

7.4.5.4 Cell Message Modification

Individual micron-scale nanorobots have insufficient chemical release volume or Manufacturing capacity to transmit endocrine-like chemical signals throughout the entire bloodstream (Section 7.2.1.8), although simultaneous coordinated chemical releases may be made from large numbers of in vivo nanomachines for special purposes. Nanodevice outmessaging to cells thus will normally be limited to localized paracrine-like, neurotransmitter, or intracellular messaging, and of course nuclei molecules which may have a direct control function. Indeed, the most efficient outmessaging to cells may involve the assembly of customized RNA molecules or their antisense variants (using locally available raw materials such as RNA bases and amino acids) to up- or down-regulate specific gene activities (Chapter 12) thus taking advantage of the cell’s inherent amplification by the ribosomes.

As a simple example, Ca\(^{++}\) serves as an intracellular mediator in a wide variety of cell responses including secretion, cell proliferation, neurotransmission, cellular metabolism (when complexed to calmodulin), and signal cascade events that are regulated by calcium-calmodulin-dependent protein kinases and adenyly cyclases. The concentration of free Ca\(^{++}\) in the extracellular fluid or in the cell’s internal calcium sequestering compartment (which is loaded with a binding protein called calsequestrin) is \(-10^{3}\) ions/nm\(^3\).\(^{3}\) However, in the cytosol, free Ca\(^{++}\) concentration varies from 6 \(-10^{6}\) ions/nm\(^3\) for a resting cell up to 3 \(-10^{6}\) ions/nm\(^3\) when the cell is activated by an extracellular signal; cytosolic levels \(-10^{6}\) ions/nm\(^3\) may be toxic,\(^{531}\) e.g., via apoptosis (Sections 10.4.1.1 and 10.4.2.1).

To transmit an artificial Ca\(^{++}\) activation signal to a typical tissue cell in \(-1\) millisecond, a single nanorobot stationed in the cytoplasm must promptly raise the cytosolic ion count from 480,000 Ca\(^{++}\) ions to 24 million Ca\(^{++}\) ions, a transfer rate of \(-2.4 \times 10^{15}\) ions/sec which may be accomplished using \(-24,000\) molecular sorting rotors (Section 3.4.2) operated in reverse, requiring a total nanorobot emission surface area of \(-2.4\) micron\(^2\). Or, more compactly, pressurized venting or an ion nozzle may be employed (Section 9.2.7). Onboard storage volume of \(-0.1\) micron\(^3\) can hold \(-2\) billion calcium atoms, enough to transmit \(-100\) artificial Ca\(^{++}\) signals into the cell after ionization (2877 zl/ion double-ionization energy)\(^{763}\) even assuming no recycling.

In addition to the amplitude modulation (AM) of Ca\(^{++}\) signals noted above, De Koninck and Schulman\(^{1121}\) have discovered a mechanism (CaM kinase II) that transduces frequency-modulated (FM) Ca\(^{++}\) intracellular signals in the range of 0.1-10 Hz. Fine tuning of the kinase’s activity by both AM and FM signals (either of which is readily detected or generated by in cyto nanorobots) may occur as the molecule participates in the control of diverse cellular activities.

Nanorobots may also mechanically outmessage to cells. The cytoskeleton of living cells and nuclei are hard-wired such that a mechanical tug on cell surface receptors can immediately change the organization of molecular structures inside the nucleus and the cytoplasm,\(^{942}\) and can dramatically alter cellular electrical conductivity.\(^{1100}\) The intermediate fiber network (Section 8.5.3.11) alone is sufficient to transmit mechanical stress to the nucleus.\(^{942}\) The minimum force required to initiate signal transduction may be as low as \(-10\) pN (Section 9.4.3.2.1), well within the \(-100\) nanonewton capacity of the nanomanipulator arm described in Section 9.3.1.4.

7.4.5.4 Cell Message Modification

Properly configured in cyto nanorobots can modify natural intracellular message traffic according to preprogrammed rules or by following external commands issued by the attending physician. In the case of steroids and thyroid hormones, this may involve the direct manipulation of the signaling molecules themselves (after they
Table 7.2. The Four Major Classes of Human Cell Secretory Signaling Molecules\textsuperscript{531,755,767,769}

<table>
<thead>
<tr>
<th>I. Endocrine Hormones</th>
<th>II. Local Chemical Mediators</th>
<th>III. Neurotransmitters</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>STEROIDS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldosterone</td>
<td>Leukotriene A(_2) (LTA(_2))</td>
<td>Acetylcholine</td>
</tr>
<tr>
<td>Catecholestrogens</td>
<td>Leukotriene A(_3) (LTA(_3))</td>
<td>Acetylserotonin</td>
</tr>
<tr>
<td>Cholecalciferol (vitamin D(_3))</td>
<td>Leukotriene B (LTB)</td>
<td>(\gamma)-Aminobutyric acid (GABA)</td>
</tr>
<tr>
<td>Cortisol (hydrocortisone)</td>
<td>Leukotriene C (LTC)</td>
<td>Dopamine</td>
</tr>
<tr>
<td>Dehydroepiandrostosterone (DHEA)</td>
<td>Leukotriene C(_3) (LTC(_3))</td>
<td>Glutamic acid</td>
</tr>
<tr>
<td>11-Deoxycorticosterone</td>
<td>Prostaglandin A(_2) (PGA(_2))</td>
<td>Glycine</td>
</tr>
<tr>
<td>Dihydrotestosterone</td>
<td>Prostaglandin E(_1) (PGE(_1))</td>
<td>Histamine</td>
</tr>
<tr>
<td>Estradiol &amp; Estriol</td>
<td>Prostaglandin E(_2) (PGE(_2))</td>
<td>Melatonin</td>
</tr>
<tr>
<td>Estrone</td>
<td>Prostaglandin F(_1) (PGF(_1))</td>
<td>Nitric Oxide</td>
</tr>
<tr>
<td>Progesterone</td>
<td>Prostaglandin F(_2) (PGF(_2))</td>
<td>Norepinephrine (noradrenaline) (NEP)</td>
</tr>
<tr>
<td>Testosterone</td>
<td>Thromboxane A(_2) (TXA(_2))</td>
<td>Octopamine</td>
</tr>
<tr>
<td>AMINO ACID DERIVATIVES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epinephrine (adrenalin)</td>
<td>Histamine</td>
<td>Serotonin</td>
</tr>
<tr>
<td>Thyroid Hormone (thyroxine)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEPTIDES AND PROTEINS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adrenocorticotropic hormone (ACTH)</td>
<td>Hematopoietin Family: Erythropoietin, IL-2 (Interleukin), IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-13, G-CSF, GM-CSF (Granulocyte-Macrophage Colony Stimulating Factor), M-CSF, CNTF, OSM, LIF (Leukemia Inhibitory Factor), IFN(<em>\alpha) (Interferon), IFN(</em>\beta), IFN(_\gamma)</td>
<td></td>
</tr>
<tr>
<td>Antimullerian hormone</td>
<td>EGF Family: Epidermal Growth Factor (EGF, or urogastrone), TGF(_\beta) (Transforming Growth Factor)</td>
<td></td>
</tr>
<tr>
<td>Arg-Vasopressin (AVP)</td>
<td>G(<em>\beta)-Trefoil Family: FGF(</em>\beta) (Fibroblast Growth Factor), FGF(_\beta), IL-1(\alpha), IL-1(\beta), IL-1(R)</td>
<td></td>
</tr>
<tr>
<td>Arg-Vasotocin</td>
<td>TNF Family: TNF(<em>\alpha) (Tumor Necrosis Factor), TNF(</em>\beta), LT(\beta)</td>
<td></td>
</tr>
<tr>
<td>Atrial natriuretic factor (ANF)</td>
<td>Cysteine Knot Family: NGF (Nerve Growth Factor), TGF(<em>\beta)1, TGF(</em>\beta)2, TGF(_\beta)3, PDGF (Platelet Derived Growth Factor), VEGF (Vascular Endothelial Growth Factor)</td>
<td></td>
</tr>
<tr>
<td>Autoimmune antiinsulin</td>
<td>Cheomkin Family: IL-8, MIP-1(\alpha) (Macrophage Inflammatory Protein), MIP-1(\beta), MIP-2, PF-4, PBP, I-309/TCA-3, MCP-1, MCP-2, MCP-3, (\gamma)IP-10</td>
<td></td>
</tr>
<tr>
<td>Cacitonin (CT)</td>
<td>Other Cytokines: Angiopoietin, Eosinophil Chemotactic Factor, Eosinophil Stimulator Promoter, Fibropeptide B, Insulinlike Growth Factors (IGF), Kallikrein, Leukocyte Inhibitory Factor, Lymphocyte Stimulating Factor, Macrophage Activating Factor (MAF), Macrophage Growth Factor (MGF), Macrophage Stimulatory Protein (MSP), Neutrophil Chemotactic Factor, Platelet Activating Factor, Proliferation Inhibitory Factor, Thromboxipentin (TPD), Thymopoietin</td>
<td></td>
</tr>
<tr>
<td>Enteroglucagon (GLI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follicle-Stimulating hormone (FSH)</td>
<td>ACTH</td>
<td></td>
</tr>
<tr>
<td>Gastric Inhibitory Polypeptide (GIP)</td>
<td>Angiotensin II</td>
<td></td>
</tr>
<tr>
<td>Gastrin</td>
<td>Bombsin</td>
<td></td>
</tr>
<tr>
<td>GH Releasing Factor (GRF)</td>
<td>Bradykinin</td>
<td></td>
</tr>
<tr>
<td>Growth Hormone (GH, somatotropin)</td>
<td>Carnotin</td>
<td></td>
</tr>
<tr>
<td>Glicentin</td>
<td>Cholecystokinin (CCK, pancreaticin)</td>
<td></td>
</tr>
<tr>
<td>Human Chorionic Gonadotropin (hCG)</td>
<td>Corticotropin Releasing Factor (CRF)</td>
<td>Dendorphin</td>
</tr>
<tr>
<td>Human Placental Lactogen (hPL)</td>
<td>(\beta)-Endorphin</td>
<td>Enkephalin</td>
</tr>
<tr>
<td>Inhibin</td>
<td>Galanin</td>
<td></td>
</tr>
<tr>
<td>Insulin</td>
<td>Gastrin &amp; Gastrin Releasing Peptide</td>
<td></td>
</tr>
<tr>
<td>Leptin</td>
<td>GH Releasing Hormone (GHRH)</td>
<td></td>
</tr>
<tr>
<td>Lipotropin (LPH)</td>
<td>Glucagon &amp; Insulin</td>
<td></td>
</tr>
<tr>
<td>Long-Acting Thyroid Stimulator (LATS)</td>
<td>Kyotrophin</td>
<td></td>
</tr>
<tr>
<td>Melanocyte-Stimulating Hormone (MSH)</td>
<td>LH Releasing Hormone (LHRH)</td>
<td></td>
</tr>
<tr>
<td>Motilin</td>
<td>Lipotropin &amp; Motilin</td>
<td></td>
</tr>
<tr>
<td>Osteoclast-Activating Factor (OAF)</td>
<td>Neurokinin A &amp; B</td>
<td></td>
</tr>
<tr>
<td>Ovarian Growth Factor</td>
<td>Neuropeptide Y &amp; P</td>
<td></td>
</tr>
<tr>
<td>Pancreatic Glucagon</td>
<td>Neurotensin</td>
<td></td>
</tr>
<tr>
<td>Pancreatic Polypeptide (PP)</td>
<td>Oxytocin</td>
<td></td>
</tr>
<tr>
<td>Parathyroid Hormone (PTH)</td>
<td>Proctolin</td>
<td></td>
</tr>
<tr>
<td>Platelet Growth Factor</td>
<td>Prolactin (PRL)</td>
<td></td>
</tr>
<tr>
<td>Proinsulin</td>
<td>Protein A &amp; B</td>
<td></td>
</tr>
<tr>
<td>Prolactin-Releasing Factor (PRH)</td>
<td>Neuropeptide A &amp; B</td>
<td></td>
</tr>
<tr>
<td>Relaxin</td>
<td>Secretin</td>
<td></td>
</tr>
<tr>
<td>Secretin</td>
<td>Somatostatin</td>
<td></td>
</tr>
<tr>
<td>Somatomedins</td>
<td>Substance P</td>
<td></td>
</tr>
<tr>
<td>Thymic Humoral Factor (THF)</td>
<td>Tachykinins</td>
<td></td>
</tr>
<tr>
<td>Thyroid-Stimulating Hormone (TSH)</td>
<td>Vasoactive Intestinal Peptide (VIP)</td>
<td></td>
</tr>
<tr>
<td>Thyrotropic-Releasing Hormone (TRH)</td>
<td>Vasopressin</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IV. Intracellular Messengers</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arachidonic acid</td>
<td>Guanine Nucleotide Exchange Factors (GEFs)</td>
</tr>
<tr>
<td>Ca(++) (w/o or w/o calmodulin)</td>
<td>Inositol triphosphate</td>
</tr>
<tr>
<td>Cyclic AMP</td>
<td></td>
</tr>
<tr>
<td>Cyclic GMP</td>
<td></td>
</tr>
<tr>
<td>Diacylglycerol</td>
<td></td>
</tr>
<tr>
<td>Diadenosine tetraphosphate</td>
<td></td>
</tr>
<tr>
<td>GTPase-Activating Proteins (GAPs)</td>
<td></td>
</tr>
<tr>
<td>Guanin Nucleotide Exchange Factors (GEFs)</td>
<td></td>
</tr>
<tr>
<td>Inositol triphosphate</td>
<td></td>
</tr>
</tbody>
</table>
have passed through the cell membrane) or their bound receptor complexes. However, most signaling molecules are absorbed at the cell surface, initiating a signal cascade which must be modulated by manipulating the second-messenger molecules or other components of the signal cascade. A complete analysis of the modifications that nanorobots may impose upon normal intracellular message traffic is beyond the scope of this book. However, a few basic examples of modifying action may be illustrated using the familiar cyclic AMP (cAMP; \( \text{C}_10\text{H}_10\text{N}_3\text{O}_1\text{P} \), MW = 328 daltons) messaging system.

A. Amplification — A single epinephrine molecule received by a β-adrenergic receptor at a cell surface transduces the activation of dozens of G-protein α subunits, each of which in turn activates a single adenylate cyclase enzyme which cyclizes hundreds of ATP molecules into cAMP molecules. The intracellular population of cAMP (in muscle or liver target cells) is normally \( \leq 10^{-6} \text{M} \) or -5 million molecules for a typical (20 micron)\(^3\) tissue cell. Stimulation by epinephrine raises the cAMP population to ~25 million molecules in a few seconds. Upon detecting this rising tide of cAMP in a few milliusec, a single in cyto nanorobot could amplify this existing chemical signal by instantly releasing 20 million cAMP molecules (occupying -0.01 micron\(^3\)) from onboard inventories—decreasing cellular response time by several orders of magnitude.*

B. Suppression — Similarly, upon detection of rising cAMP levels in target cells, resident nanorobots could use molecular rotors to rapidly remove cAMP from the cytosol as quickly as it is formed, even under maximum adrenal stimulation. From Eqn. 3.4 the diffusion-limited intake current at the basal concentration (\(-6 \times 10^{-7}\) molecules/nm\(^3\)) for a cAMP-absorbing spherical nanodevice 1 micron in radius is -4 million molecules/sec, so a single such device could probably keep up with natural cAMP production rates and thus completely extinguish the response by preserving a flat basal concentration. (As a practical matter, it may be more efficient to simply control epinephrine generation at its glandular source unless it is desired to interface with just a single tissue type.) Simultaneously, the cAMP-absorbing nanorobot may hydrolyze the stored cAMP (hydrolisis energy is -11.1 Kcal/mole or -77.1 zJ/molecule)\(^{329}\) in the manner of the cAMP phosphodiesterases (c.f. nanofactories, Chapter 19), then excrete these deactivated AMP messenger molecules back into the cytosol. Similar methods might be useful in ligand-gated ion channel desensitization or in disease symptom suppression (Chapter 24)—as for example, in suppressing the prolonged elevation of cAMP in intestinal epithelial cells associated with the cholera toxin, that produces severe diarrhea by causing a large influx of water into the gut.

C. Replacement — Combining suppression and amplification, an existing chemical signal may be eliminated and replaced by a different—even an opposite—message pathway using nanorobot mediators. Alternative pathways may be natural or wholly synthetic. Novel responses to existing signals may be established within the cell to enhance functionality or to improve stability or controllability. For example, detection of one species of cytokine by a nanorobot could trigger rapid specific absorption of that cytokine and a simultaneous fast release of another (different) species of cytokine in its place. Such procedures must take into account the redundant signaling pathways and backup systems (e.g., developmental signals, immune system, blood clotting). Medical nanorobotics will allow the replacement of many redundant pathways with more refined and specific responses.

D. Linkage — Previously unlinked signal cascades may be artificially linked using in cyto nanodevices. As a fanciful example, the receipt of epinephrine by nanorobots located in the capillaries of the brain could trigger these devices to suppress the adrenalin response while simultaneously releasing chemical messengers producing message cascades that stimulate production of enkephalins or other opioids, thus encouraging a state of psychological relaxation rather than the “fight or flight” response to certain stressful conditions.

7.4.5.5 Inmessaging from Neurons

Inmessaging from neurons to nanodevices has already been considered in Sections 4.8.6, 4.9.5 and 7.4.2.6. In these discussions, it was concluded that at least half a dozen different methods should permit nanorobot sensors to noninvasively detect, measure, and count the passage of nerve impulses. Neurographics, neuroinformatics,\(^{1298}\) connectivity mapping and individual neural targeting are described in Chapter 25.

7.4.5.6 Outmessaging to Neurons

In vivo nanorobots can outmessaging to neurons as well. The most direct method is synaptic stimulation. In excitatory cholinergic synapses, perhaps the most common type, the arrival of a presynaptic nerve impulse stimulates the release of up to -10\(^5\) molecules of acetylcholine (\( \text{C}_2\text{H}_4\text{NO}_2 \), MW = 146 daltons) from synaptic vesicles into the synaptic cleft in ~1 millisecond, producing a local concentration of ~3 x 10\(^4\) molecules/nm\(^3\) (~0.0005 M).\(^{531}\) The acetylcholine molecules diffuse across the cleft to the postsynaptic membrane (Section 4.8.6.4), where they bind to specific receptor proteins (Section 3.3.3), opening ion channels within 0.1 millisecond and partially depolarizing the cell, increasing local Na\(^+\) permeability. When depolarization reaches ~15% of full voltage range (e.g., from ~60 mV down to about ~40 mV), voltage-activated sodium channels open, making the membrane highly permeable to Na\(^+\) ions. These ions rapidly flow into the cell and initiate the ~100 nanoamp action potential spike. These channels close spontaneously after a short interval to reduce permeability, thus allowing the membrane voltage to be restored to its normal resting potential (Section 3.3.3).

A sorting-rotor-tipped manipulator (Section 3.4.2) or pressurized nanoinjector (Section 9.2.7.1) placed in the vicinity of the synaptic cleft can emit a puff of -10\(^5\) acetylcholine molecules (~20,000 nm\(^3\)), producing partial depolarization and initiating nerve impulses at will. (This is preferable to the use of unnatural depolarizing agents such as veratridine.) A 1-micron\(^3\) storage volume contains ~5 billion molecules, sufficient to induce ~50,000 discharges or ~1 hour of continuous firing at 15 Hz. The initiating mechanism must provoke a discharge in ~1 millisecond, a throughput rate of 10\(^8\) molecules/sec requiring, say, ~100 sorting rotors with total manipulator tool tip area ~10\(^4\) nm\(^2\). At 15 Hz continuous firing, ~10\(^6\) molecules/sec are

* From a practical standpoint, R. Bradbury notes that designers should have knowledge of the biochemical pathways that are most critical in the response to a signal, hence nanorobots may be able to skip the intermediate steps required in the biological system. In the present example, the cascade is typically intended to increase the glucose available for muscles, so it might make more sense to skip the cAMP step and simply commence disassembly of glycogen (in the liver and muscles) and accelerate the pumping of glucose into the bloodstream (from the liver).

** Under controlled mechanochemical conditions (Chapter 19) where the effective heat of protonation of the buffering system can be ignored, the net enthalpy of hydrolysis at 298 K and pH 7 may fall as low as ~1.95 zJ/molecule, or ~0.28 Kcal/mole.\(^{1142}\)
consumed. Sorting rotors placed near the cleft can also absorb the two breakdown products (acetate and choline) of acetylcholinesterase activity (~150 microsec turnover time\(^\text{3142}\)) and recycle them into new molecules of acetylcholine. Taking the enthalpy of hydrolysis under physiological conditions\(^\text{3142}\) as ~70 zJ/molecule (~10 Kcal/mole) for the acetylcholinesterase-mediated hydrolysis reaction,\(^\text{3142}\) the power requirement for continuously recycling 10\(^6\) molecules/sec is ~0.07 pW. A similar rotor transport mechanism may also be used to rapidly extract acetylcholine molecules from the synaptic cleft, thus extinguishing a passing nerve impulse. Note that to exert control over nerve impulses, neurotransmitter injectors must be placed very close to the post-synaptic surface because the effective diffusion radius of acetylcholine is only a few microns, due to the high efficiency of local acetylcholinesterase.\(^\text{803}\)

Many compounds other than acetylcholine may serve as neurotransmitters (Table 7.2), including nitric oxide and possibly carbon monoxide.\(^\text{1125,1129}\) Catecholamines such as dopamine, norepinephrine and epinephrine, synthesized in the adrenal gland and elsewhere, act as neurotransmitters at adrenergic synapses which are found at the junctions between nerves and smooth muscles in internal or-.

The junctions between nerves and smooth muscles in internal or-

and epinephrine, synthesized in the adrenal gland and elsewhere, act as neurotransmitters at adrenergic synapses which are found at the junctions between nerves and smooth muscles in internal organs such as the intestine and in nerve-nerve junctions in the brain. There are also histaminergic neurons found exclusively in the hippocampus.\(^\text{1123}\) Most tissues in the human body are innervated by several different types of nerve cells, each using a different neurotransmitter, allowing a great diversity of signals and responses. But synaptic-resident nanorobots should be able to monitor, stimulate, or extinguish all of these signals, even given the confusing geometries entailed by synaptic clusters.

While some synapses use excitatory neurotransmitters that elicit an electrical signal, others use different neurotransmitters such as GABA, glycine, or the enkephalins that deliver an inhibitory signal, suppressing electrical response. Some molecules work both ways. For example, acetylcholine is excitatory at neuromuscular junctions but may be excitatory or inhibitory in the central and peripheral nervous system. A nerve cell may receive thousands of these excitatory and inhibitory chemical inputs, and their relative number determines the probability of axonal firing. Again, selections from the complete known library of neurotransmitters of all types may be made available for storage, synthesis, absorption or release by properly configured nanorobots.

Neuropeptides, which can act as highly specific triggers for complex patterns of activity in the nervous system (including memory, learning, perception, mood, and behavior) and allow long-term chemical modulation of synaptic sensitivity, add yet another complication to neural outmessaging by nanorobots. Neuropeptides are manufactured by ribosomes on rough endoplasmic reticulum within the neuron cell body and must be transported to axon terminals for release by “fast axonal transport”—a journey that may take a day or more for a long axon. They are stored in intracellular vesicles from which they are released to the extracellular space of peptidergic neurons. Neuropeptides released into the synaptic cleft serve as neurotransmitters (to inhibit the action of excitatory neurotransmitters) or neuromediators (to prolong the action of neurotransmitters), both functions easily duplicated by nanorobots. However, neuropeptides released into the bloodstream act as long-range neurohormones, and neuropeptides released into the extracellular space act as paracines and enter both pharmacodynamic and pharmacokinetic cascades.\(^\text{770}\) Nanorobot management of these more diffusely distributed tissue neuropeptide concentrations (e.g., by mimicking the action of natural peptidases of broad substrate specificity) is a more difficult, though not intractable, task.

There are many other methods by which in vivo nanorobots can outmessage to neurons:

1. Periodic manipulation of electric (Section 4.8.6.1) and magnetic (Section 4.8.6.2) fields may trigger neural impulses.\(^\text{3328}\)

2. An excess population of Na\(^+\) ions outside the myelinated nerve fiber establishes an intra-axonal resting potential of ~60 mV. The net movement of Na\(^+\) ions associated with a single neural impulse is ~0.3 ion/micron\(^2\) in ~0.5 milliseconds,\(^\text{326}\) so the direct injection of a few electric charges should produce local depolarization, initiating a neural impulse. Note that the voltage balance is maintained by differing concentrations of several ions, not just Na\(^+\), although the inrush of Na\(^+\) is the primary cause of membrane voltage change during the discharge spike.

3. Membrane ionic currents may be pharmacologically manipulated by nanodevices permanently stationed within the cytosol. For example, sodium pumps are completely inhibited by the injection of 10-100 nM of cardioactive steroids such as ouabain and strophanthidin.\(^\text{805}\) Injection of ~300 nM tetrodotoxin inhibits sodium currents while leaving potassium currents unchanged; inserting tetraethylammonium blocks only the potassium currents, while leaving sodium currents unchanged.\(^\text{804}\)

4. Establishing a localized ~100 mV circumaxonal (external) superexcess of positive charge may present an electrical barrier steep enough to quench a depolarization wave, thus extinguishing a passing nerve impulse. Techniques of electronic anesthesia, such as transcutaneous electronic nerve stimulation (TENS) and cell demodulated electronic anesthesia (CEDETA), involve crude applications of similar principles.\(^\text{3299,3300}\)

5. Long-term electrical stimulation of neurons using electrodes to apply 0.1-Hz pulses for ~10\(^5\) sec reduced the expression of neural cell adhesion molecule L1 (NCAM) by a factor of ~13.\(^\text{1063}\)

6. Direct ultrasonic stimulation of neurons.\(^\text{3535}\)

7.4.6 Outmessaging to Patient or User

In many applications, in vivo medical nanodevices may need to communicate information directly to the user or patient. This capability is crucial in providing feedback to establish stable and reliable autogenous command and control systems (Chapter 12). Outmessaging from nanorobot to the patient or user requires the nanodevice to manipulate a sensory channel that is consciously available to human perception, which manipulation can then be properly interpreted by the patient as a message.

Sensory channels available for such communication include sight, audition, gustation and olfaction, kinesthesia, and somesthetic sensory channels such as pressure, pain, and temperature. Nanodevice-originated data may be superimposed upon the natural sensory traffic by:

1. generating an artificial sensory stimulus,

2. direct stimulation of the receptor in the absence of actual sensory stimulus, or

3. triggering artificial action potentials in the afferent nerves that carry information from the sensor to the CNS.

7.4.6.1 Somesthetic Outmessaging

Human skin contains about 700,000 pressure or touch receptors, mostly concentrated on the fingers (~17,000 receptors on all fingertips) and on the face—especially the lips and the tip of the tongue.
In addition to information about the presence of tactile stimuli, touch receptors can also discriminate both intensity and spatial direction of the stimulus (Table 7.3). Meissner’s corpuscle, located in the papillae of the corium just beneath the epidermis, is a receptor for light touch. The Pacinian corpuscle is a deep touch or pressure receptor located in the deep parts of the dermis, in the connective tissue around muscles, tendons, and joints, and in the mesenteries supporting the visceral organs.

A networked population of outmessaging nanorobots, with each nanorobot positioned near a Meissner’s or Pacinian corpuscle and capable of triggering artificial nerve impulses (Section 7.4.5.6), could create coordinated spatial and temporal patterns of pressure sensations that could be interpreted by the patient as specific messages. For example, nanorobots neurally emulating the sensation of pressure in the fingers in the sequence Thumb-Thumb-Ring-Index-Thumb are communicating the message “11421” to the user at a transfer rate of 1-10 bits/sec. Similarly, stimulating the arrector muscles artificially triggers the nerve plexus surrounding epidermal hair follicles (which register hair motion), creating well-defined and potentially communicative patterns of “gooseflesh” on the skin.

Mild stimulation of pain receptors (nociceptors), consisting of unmyelinated free sensory nerve endings with pain sensation mediated by the peptide nociceptin, could also produce interpretable messages. For instance, some C fibers when sufficiently stimulated produce burning pain, while small D fibers signal a prickling pain. (There are ~228 pain receptors/cm² in the neck region, ~188/cm² on the back of the hand, ~95/cm² on the radal surface of the middle finger, and ~44/cm² on the tip of the nose.) Heat receptors (brushes of Ruffini) and cold receptors (end bulbs of Krause) also may be employed for outmessaging. Cold spots outnumber warm spots by a factor of 4-10, with temperature receptors concentrated on the face and hands and specific regions showing marked differences in the mix of sensor types (e.g., the forehead contains 0.6 warm receptors/cm² but 8.0 cold receptors/cm²). Nanorobots could employ these mechanoreceptive afferents; G = glabrous (hairless) skin; H = hairy skin.

Mild stimulation of pain receptors (nociceptors), consisting of unmyelinated free sensory nerve endings with pain sensation mediated by the peptide nociceptin, could also produce interpretable messages. For instance, some C fibers when sufficiently stimulated produce burning pain, while small D fibers signal a prickling pain. (There are ~228 pain receptors/cm² in the neck region, ~188/cm² on the back of the hand, ~95/cm² on the radal surface of the middle finger, and ~44/cm² on the tip of the nose.) Heat receptors (brushes of Ruffini) and cold receptors (end bulbs of Krause) also may be employed for outmessaging. Cold spots outnumber warm spots by a factor of 4-10, with temperature receptors concentrated on the face and hands and specific regions showing marked differences in the mix of sensor types (e.g., the forehead contains 0.6 warm receptors/cm² but 8.0 cold receptors/cm²). Nanorobots could employ these mechanoreceptive afferents; G = glabrous (hairless) skin; H = hairy skin.

Table 7.3. Tactile Receptors in the Human Skin<sup>854,869,1633-1637</sup>

<table>
<thead>
<tr>
<th>Skin Receptor Type</th>
<th>Receptor Class</th>
<th>Skin Type</th>
<th>Probable Sensory Correlation</th>
<th>Receptive Field, Range (and Median)</th>
<th>Frequency Range (and Most Sensitive)</th>
<th># of Fingertip Receptors (and on Palm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacinian corpuscles</td>
<td>PC</td>
<td>G, H</td>
<td>Vibration, tickle</td>
<td>10-1000 mm² (100 mm²)</td>
<td>40-800 Hz (200-300 Hz)</td>
<td>21/cm² (9/cm²)</td>
</tr>
<tr>
<td>Meissner’s corpuscles</td>
<td>RA</td>
<td>G</td>
<td>Touch, tickle, motion, vibration</td>
<td>1-100 mm² (13 mm²)</td>
<td>10-200 Hz (20-40 Hz)</td>
<td>140/cm² (25/cm²)</td>
</tr>
<tr>
<td>Hair follicle nerve</td>
<td>RA</td>
<td>H</td>
<td>Touch, vibration</td>
<td>~0.01 mm²</td>
<td>10-100 Hz (?)</td>
<td>10/cm² (200/cm² (scalp))</td>
</tr>
<tr>
<td>Ruffini ending</td>
<td>SA II</td>
<td>G, H</td>
<td>Stretch, shear, tension (?)</td>
<td>10-500 mm² (60 mm²)</td>
<td>7 Hz</td>
<td>49/cm² (16/cm²)</td>
</tr>
<tr>
<td>Merkel’s cells</td>
<td>SA I</td>
<td>G</td>
<td>Pressure, edge (?)</td>
<td>2-100 mm² (11 mm²)</td>
<td>0.4-100 Hz (7 Hz)</td>
<td>70/cm² (8/cm²)</td>
</tr>
<tr>
<td>Tactile disks</td>
<td>SA I</td>
<td>H</td>
<td>Pressure, edge (?)</td>
<td>3-50 mm²</td>
<td>1-100 Hz (?) (10 Hz ?)</td>
<td>70/cm² (8/cm²)</td>
</tr>
</tbody>
</table>

PC = Pacinian afferents; RA = rapidly-adapting afferents; SA II = slowly-adapting, large receptive field afferents; SA I = slowly-adapting, small receptive field mechanoreceptive afferents; G = glabrous (hairless) skin; H = hairy skin.

The Optacon (Optical-to-Tactile Converter), originally developed by NASA, scanned normal inkprint and converted the characters into a tactile format on a 24 x 6 vibrating-pin array, enabling the blind to feel-read without Braille. Average users achieved ~30 words/min, or up to ~100 words/min in exceptional cases. The Optacon was first sold in 1970 by Telesensory Inc. but is no longer manufactured. Blazie Engineering, Inc., which acquired all rights to the Optacon in 1998, also sells Power Braille, an 81-character 8-dot refreshable Braille display for desktop computer applications.

7.4.6.2 Kinesthetic Outmessaging

Direct stimulation of kinesthetic sensory neural channels for outmessaging may prove disorienting, disturbing, confusing, or even nauseogenic to the user, as for example a periodic manipulation of the vestibular ganglion (creating synthetic vertigo), muscular stretch receptors (creating “phantom limb” effects), or reflex arcs (e.g., sneezing, knee jerks, lactation, pupillary dilation, or orgasm). Patients might be slightly more comfortable if nanorobots generate a genuine sensory stimulus which can then be detected as a true sensory reading. Examples might include the creation of patterned nervous tics by triggering small-muscle spasms in the face or limbs (switches), temporally-structured subvocal hiccups by stimulating and inhibiting the phrenic nerve causing tiny spasms in the diaphragm muscle, or periodic blepharospasms (winking) by stimulating the orbicularis palpebrarum muscle (ring muscle) of the eye.

Selective activation of motor neurons controlling the ~400 skeletal muscles in the human body can produce involuntary macroscopic limb motion. To take a familiar example, properly positioned neurostimulatory nanorobots may trigger in sequence the extensor indicis and the flexor indicis muscles, causing the second and third phalanges of the index finger first to extend, then to retract; or trigger
the extensor minimi digiti and the flexor brevis minimi digiti muscles, causing the second and third phalanges of the little finger to extend, then the first phalanx of the little finger to flex; or trigger the extensor pollicis brevis and the flexor pollicis brevis muscles, causing the base of the first phalanx of the thumb to extend, then retract. These artificially induced finger movements may be directly interpreted as messages by a trained user at a 1-10 bit/sec transfer rate, or else the patient may grasp a handheld device that translates the controlled sequence of forced gripping motions into a correct alphanumeric or voiced message.

7.4.6.3 Auditory Outmessaging

The human ear consists of an air-filled external auditory canal terminating on the ~1 cm² tympanic membrane (~100 microns thick), whose vibrations are transmitted through the bony ossicles of the ~2 cm³ middle ear to the membranous oval window of the inner ear (Fig. 7.3). From there, sound energy enters the cochlea as compressional waves in the perilymph, an incompressible watery fluid that fills most of the ~100 mm³ spiral cochlear tube. This ~35-mm long tube resembles a snail shell, coiled two and a half times around the modiolus, itself the center axis along which pass the cochlear nerves and blood vessels. About 24,000 hair cells (multifiber stereocilia) on the basilar membrane under the tectorial membrane (Fig. 7.4) inside the cochlea resonate in response to different acoustic frequencies. These cells make complex synaptic contacts with neurons of the spiral ganglion, which passes this aural information into the auditory (8th cranial) nerve and thence to the brain for decoding.

A large nanodevice (~20-50 microns) stationed in the perilymph of the cochlea can “speak” directly into the human ear, loud enough to be heard. Total acoustic power delivered to the tympanic membrane during normal audition is ~10⁻⁴ pW (0 dB) at the threshold of hearing, ~0.1 pW for whispering (30 dB), and ~100 pW for normal conversation (60 dB). At the low frequencies (30-10,000 Hz) common to human speech, micron-scale acoustic radiators are very inefficient (Section 7.2.2.1) although there are virtually no attenuation losses (Eqn. 4.52). Even so, from Eqn. 7.7 a cochlear nanodevice able to apply an input power of 2000 pW to an acoustic omnidirectional piston radiator 20 microns in diameter can generate 10 KHz sound waves at an output power ~ 0.1 pW, approximately the intensity level of whispering. At a more comfortable 3 KHz (the peak sensitivity of human hearing), P_in must rise to ~22,000 pW or piston diameter must increase to 45 microns to maintain P_out ~ 0.1 pW, or a duty cycle substantially less than 100% must be employed. If linked to a communication network, the cochlear nanodevice can receive information to be audibilized from the network. This device can also be used for acoustic inmessaging by allowing it to intercept sound waves before they reach the basilar membrane, then passing this information immediately to the network (which thus receives the signal sooner than the brain). Perilymph contains glucose for power at roughly serum concentrations; alternatively, the cochlear device can absorb and store the natural ambient acoustic energy present in the auditory canal (Section 7.4.8).

Similarly, a nanorobot located in the endolymph-filled scala media, the innermost of the three parallel chambers of the cochlea, can attach itself to the reticular lamina of the organ of Corti and physically manipulate the stereocilia to produce the desired audible stimulus. Stereocilia are spaced over the organ of Corti at a linear density of ~1 cell/micron along the 30 mm length of the cochlear spiral tube. Continuously stimulating all ~100 cells within reach of a 20-micron diameter nanorobot with bidirectional 40-micron long extensible manipulators should produce a narrow-frequency tone up to ~0.4% of maximum pressure amplitude, close to normal conversational levels. Hair cells have a membrane time constant of ~0.5 millisecond in theory permitting a transfer rate of up to ~2000 bits/sec if all stereocilia are manipulated coherently as a single channel. Careful design should prevent unwanted stimulation of tinnitus.
Direct neural stimulation of some of the ~31,000 spiral ganglion cells allows transduction of information into the human auditory system. Each ganglial cell carries temporal and amplitude information on a narrow band of acoustic frequencies received by the ear, essentially a crude Fourier transform of the original audio signal. Properly positioned nanorobots can superimpose artificial signals on this natural traffic which may be perceived by the patient as a frequency-reconstructed voice of limited tonal range speaking words or numbers. Nanorobots stationed at the saccule, an otolithic organ in the vestibule of the inner ear, can stimulate direct ultrasonic hearing in humans up to ~108 KHz.1372 In 1998, cochlear implant devices were in wide use, with audible sounds transmitted by FM radio waves to an electrode array in the inner ear, providing direct electrical stimulation of the auditory nerves.1891-1893

A review of current technology in artificial speech synthesis1653,1654 is beyond the scope of this text.

7.4.6.4 Gustatory and Olfactory Outmessaging

In theory it would be possible to insert artificial flavors onto the lingual surface, or to release artificial scents into the nasal epithelium. However, human taste buds have a lifetime of ~10 days, requiring constant repositioning of resident chemical emission devices. Human olfactory receptors are located at the back of the nose (Fig. 8.11), covered by a sheet of mucus, adding signal time delays and dispersion (due to finite and differential molecular diffusion rates) to high-frequency olfactory messages. Additionally, frequent large-volume chemical emissions from individual nanorobots is not feasible due to scaling factors (Section 7.2.1.8).

In this case it appears more efficient to stimulate the gustatory and olfactory nerves directly. This minimizes power consumption, reduces the required volume of chemical releases by many orders of magnitude, and permits rapid multiplexing of numerous distinct synthetic sensory signals.

A. Gustatory Outmessaging — There are ~12,000 taste buds innervated by nerve fibers that lose their myelin sheaths as they pass through the basement membrane. Large fibers end on two or more bud cells, with the endings of small fibers invaginating the receptor cell membrane. The large taste fibers from the anterior two-thirds of the tongue branch from the lingual nerve to form a slender nerve, the chorda tympani, which traverses the eardrum as part of the facial (7th cranial) nerve en route to the medulla. The afferent fibers from the posterior third of the tongue collect in the lingual branch of the glossopharyngeal (9th cranial) nerve, through the petrosal ganglia and into the medulla. Since taste buds have a response repertoire of only five known distinct sensations (sweet, sour, bitter, salt, and umami), nanorobots seeking to artificially provoke these sensations may proceed to the chorda tympani or petrosal ganglia with the objective of triggering a higher-level confluent signal on each of a handful of distinct ganglial positions corresponding to the five primary stimuli plus a small set of mixed-sensitivity combinations. Proper identification of the nerve bundles carrying information on each of the five primary stimuli may require a brief training session or prior connectivity mapping (Chapter 25). Note also that natural taste response time is diffusion-limited to ~1 sec (~1 Hz, ~1 bit/sec), and it is not yet known whether the natural

*Umami, a complex “meaty” or “savory” tastiness factor, was first identified by a Japanese scientist, Professor Ikeda, at the University of Tokyo.3030-3034 Umami is specifically associated with monosodium glutamate (MSG)3035, sodium inosinate, and sodium guanylate. Specific taste bud receptors (e.g., mGluR4) for umami have been discovered.2031,2034

Fig. 7.4. Cross-section of the cochlear duct (redrawn from Wilson526).
conscious brain can be educated to process multiplexed high-frequency (up to 10-100 Hz, or ~10-100 bits/sec) artificially-stimulated taste signals.

B. Olfactory Outmessaging — Humans have up to 50 million olfactory sensors spaced an average ~3 microns apart in the olfactory epithelium (~2.5 cm²/nostril) located in the upper part of the nasal cavities on the surface of the superior conchae and upper part of the septum (see Figure 8.11). Each sensor protrudes a long and extremely thin olfactory nerve fiber. These fibers are grouped into ~20 filaments in each nasal cavity, and the filaments pass through the cribiform plate of the ethmoid bone into the olfactory bulb (the terminus of the olfactory (1st cranial) nerve) in the forebrain in the cranial cavity. The nasal epithelium also contains a few bare nerve endings from fibers of the trigeminal (5th cranial) nerve. The responses of olfactory receptor neurons are rarely specific to only one odor; the majority of cells respond to a broad range of substances.3433

The olfactory nerve fibers enter the olfactory bulb to end in a series of intricate basketlike synaptic terminations or antenna-shaped dendritic clusters called glomeruli. Each olfactory glomerulus receives impulses from about 26,000 receptors and sends them on through 24 mitral cells and 68 tufted cells, thus showing a high degree of neural convergence.

The number of such glomeruli (~1000) may approximate the number of nonmultiplexed distinct olfactory sensations available to man. Humans can identify over 10,000 structurally distinct odorant ligands807—some have claimed as many as 400,000 discernible odors1128—but this does not imply an equally large repertoire of odorant receptors, because structurally related odorants may bind to the same receptor molecules.1128 Indeed, individual neurons respond strongly to closely related groups of odorants and weakly or not at all to many others.807 Sensory neurons with similar ligand specificities project to common glomeruli. The family of 7TD membrane proteins, which is localized to the olfactory cilia and shares homology with the superfamily of neurotransmitters and neuropeptide 7TD receptors,808 constitutes an extremely large multigene family likely to encode odorant receptors that transduce intracellular signals by interacting with cellular transmembrane G proteins which then activate second messenger systems inside the cell. This family of putative odorant receptors could constitute one of the largest gene families in the genome, with perhaps as many as 500-1000 genes,809 approximately one receptor type for each glomerulus.

If each glomerulus may be treated as the locus for a single unique olfactory sensation, this may provide an essentially clear output channel for a neurostimulatory nanodevice. Such a device could trigger action potentials in a single glomerulus producing a perception of a rapidly pulsing (possibly up to ~10-100 Hz, or ~10-100 bits/sec) rare scent. Such signals would no longer be diffusion limited (to ~1 Hz, ~1 bit/sec) as is the case with natural olfaction, and rapid adaptation caused by receptor saturation would also be eliminated for these signals since the signals are inserted at the glomeruli and don’t pass through the sensory receptors at all. A thousand coordinated nanorobots positioned one at each glomerulus could provide complete olfactory sensation control and the potential for massive multiplexing.

As with taste messaging, it is unknown whether the natural conscious brain can be trained to process such multiplexed high-frequency smell signals. In 1998, studies of the human olfactory (combinatorial) code were in progress.3207

7.4.6.5 Ocular Outmessaging

Nanorobots may send messages to the patient by emitting photons directly into the eye (Fig. 7.5), generating an artificial visual stimulus. Retinal displays could be broadcast by photoemissive nanodevices located:
1. in the palpebral conjunctiva (inner mucosal surface) of the eyelid, where access to nutrients from the capillary bloodflow is still available in muscular tissue below the fibrous plates;

2. on the exterior surface of the cornea, where nanorobots may obtain power from transcorneal-diffused glucose and other sugars in the lacrimal fluids at $c_{\text{ligand}} = 2.9 \times 10^{-4}$ molecules/nm$^3$ (~10% blood plasma concentration)$^{585}$.

3. on the interior corneal or exterior lens surface ($c_{\text{ligand}} = 1.6-3.7 \times 10^{-3}$ molecules/nm$^3$ for glucose in the aqueous humor, ~plasma concentration)$^{585,3284}$ or in the anterior chamber in front of the lens (implanting an artificial lens in the anterior chamber was considered an experimental vision correction procedure in 1998);

4. inside the lens which is 68% water ($c_{\text{ligand}} = 0.8-2.0 \times 10^{-3}$ molecules/nm$^3$ for glucose)$^{585}$ -50% plasma concentration;

5. the interior lens or retinal surface with access to glucose in the vitreous humor (same concentration as aqueous humor); or

6. within the individual rod (dim light, monochromatic) and cone (bright light, color-sensitive) cells — retinal glucose is stored in glial Muller cells and is supplied upon demand$^{3260}$

Oxygen is available at ~5% of normal plasma concentration via solvation in the lacrimal fluids and diffusion into the aqueous humor; CO$_2$ and lactate exit by the same route. Formation of potentially signal-blocking scar tissue is avoided by using mobile nanorobots with active, biocompatible exterior surfaces$^{3234}$ (Chapter 15).

We shall now consider four different methods for ocular outmessaging:

A. Extraretinal Projection — Photomissive nanodevices must generate a photon flux intensity sufficient to equal or exceed ambient background illumination levels. With the eyelid lightly closed, normal indoor illumination transmitted through this thin muscular fibrous tissue produces a background retinal flux of at least $I_{\text{min}} = 10^{-2}$ watts/m$^2$. With eyes open in a normally lighted room, the background flux may be ~1 watt/m$^2$ or even higher (Section 4.9.4).

Normally the lens of the eye can accommodate viewing distances no closer than $X_{\text{min}} = 10$ cm (measured from posterior lens surface) in adults. Photons emitted from nanodevices located in the eyelid ($X_{\text{object}} = 8.6$ mm), exterior corneal surface ($X_{\text{object}} = 7.3$ mm), interior corneal surface ($X_{\text{object}} = 6.2$ mm), or exterior lens surface ($X_{\text{object}} = 3.8$ mm) cannot be brought to focus inside the eye unless they are tightly collimated. A cylindrical emitter of length $h_e = 25$ microns and radius $r_e = 0.5$ micron can produce a collimated beam with an axial divergence angle $\phi_e = 0.5$ tan$^{-1}(2 r_e / h_e) = 1^\circ$ (2° full-width divergence), which is geometrically equivalent to rays passing through a pupil aperture $d_{\text{pupil}} = 5.6$ mm after emanating from an object located at a distance $X_{\text{object}} = d_{\text{pupil}} / 2$ tan$^{-1}(\phi_e) = 12-15$ mm which is~ $X_{\text{min}}$, hence is focussable. One major difficulty with eyelid projection is that tight flexing of the corrugator supercilii muscle (which draws the eyebrow downward and inward) causes the eyelid tissue to ripple and fold, potentially producing significant optical distortions. A large projector suspended in the anterior chamber might offer greater spatial stability. These devices are probably small enough to avoid stimulating lid edema or blepharitis.

Present-day micron-scale LEDs emit up to ~1 watt/m$^2$ (Section 5.3.7), so extraretinal displays from such emitters are comfortably read by patients. If the emitters overwrite 10% of the visual field, then total optical power input through the lens is $P_{\text{out}} = 0.03-3.0$ microwatts. Assuming a very conservative energy conversion efficiency of e$\% = 0.01(1\%)$ for LEDs (Section 5.3.7), the emitters consume a total input power (and radiate waste heat) of $P_{\text{in}} = P_{\text{out}} / e\% = 3-300$ microwatts, well below the conservative 2000 microwatt maximum for 0.1 cm$^3$ of eyelid tissue (Eqn. 6.54). Emitter power density ranges from ~0.4 megawatts/m$^2$ at 10$^2$ watts/m$^2$ intensity up to ~40 megawatts/m$^2$ at 1 watt/m$^2$ intensity.

B. Foveal Projection — The fovea is a small pit in the macula lutea, opposite the visual axis, which is the spot of most distinct vision (Fig. 7.5). There are no rod cells in the fovea and relatively few in the near-foveal region. Foveal cone cells have diameters of 1-5 microns$^{585}$ (average $d_{\text{fov}} = 3$ microns) with a minimum center-to-center separation of ~2.6 microns (average $x_{\text{fov}} = 5$ microns) across the foveal surface of the retina. The retina is arranged “inside out” so that light must pass through nine layers of connecting nerve cells and other tissue to reach the receptors which lie at the deepest level (Fig. 7.6, enlargement of box in Figure 7.5). This covering tissue is $x_{\text{tiss}} = 130$ microns deep across the fovea (the ganglion and bipolar layers are pushed aside, directly above the fovea) and $x_{\text{c}} = 300$ microns thick elsewhere on the retina. The foveal and near-foveal region extends ~2° (~0.55 mm$^2$) around the visual axis, comprising ~20,000 cone cells and providing a resolvable visual angle of 0.5-2.3 minutes of arc (sufficient for reading printed text) and creating a 140 x 140 pixel readable billboard.

Consider a cylindrical emitter of diameter ~1 micron affixed to the retinal surface over the fovea. This emitter produces a collimated beam with an axial divergence angle $\phi_e = 1^\circ$, which diverges to a width $w_{\text{beam}} = x_{\text{tiss}} \tan(2 \phi_e) = 5.2$ microns after traveling a distance $x_{\text{tiss}} = 130$ microns into the fovea. Each cylindrical emitter can target one foveal cone cell with no overlap, so ~20,000 photomissive nanorobots are required for complete control of all foveal cone receptors. A population of 20,000 photomissive nanodevices ~1 micron in diameter attached to the inner retinal membrane above the foveal surface blocks only ~3% of photons entering the fovea via the lens, an unnoticeable diminution of the natural incident intensity that is visible to the patient.*

Given that $d_{\text{fov}} < w_{\text{beam}}$ and $\pi r_e^2 < 1$ microns$^2$, achieving 0.01-1 watts/m$^2$ intensity at the receptor cells to overcome the natural illumination background requires an emitter photonic intensity of 0.04-4 watts/m$^2$, already within state-of-the-art in 1998. At a very conservative $e\% = 0.01(1\%)$ ergophotonic efficiency, $P_{\text{in}} = 3-300$ pW per nanodevice or 0.06-6 microwatts for the entire population, well below the ~0.1 watt recommended thermogenic limit for the eyeball from Eqn. 6.54. If the nanorobot employs an oxyglucose power supply, continuous foveal control consumes ~$10^{10}-10^{12}$ glucose molecules/sec which would exhaust the ~$10^{19}$ glucose molecules present in the vitreous humor in ~0.3-30 years but would more quickly exhaust the oxygen supply (Section 6.5.3) in ~0.5-50 days even if there were no diffusive resupply from the surrounding tissues; fortunately, such resupply appears highly likely.

If the photomissive foveal nanorobot population is extended to cover the rest of the retinal surface (with most devices not emitting photons most of the time), then real-time ocular and cranial positional

* J. Logajon points out that if the cylindrical emitter has a photodetector on the end facing the lens, then even the ~3% loss could be largely retrieved by amplification; the emitters could also serve as dark-area amplifiers, providing improved night vision or infrared-to-visible light conversion (Chapter 30).
information provided by kinesthetic macrosensors (Section 4.9.2.1) via the communication network can be used to shift displayed images in synchrony with eye and head movements, allowing displays to appear fixed in the field of view as in normal viewing of external objects. (Information gleaned from the oculomotor, trochlear, trigeminal, abducens, and spinal accessory nerves may assist in this control process.) Accommodation sensors (Section 7.4.2.2) assist in maintaining proper focus. Foveal nanorobots should be designed to avoid stimulation of uveitis or retinitis.

C. Ganglionic Stimulation — Because of the inverted structure of the retina, neurostimulatory nanorobots attached to the inner retinal membrane can trigger artificial action potentials directly in the axons of afferent ganglia carrying information from rods and cones to the optic (2nd cranial) nerve. It is useful to briefly review the basic structure of the human retina (Fig. 7.6).

Starting with the outermost layer and moving inward toward the vitreous humor that fills the eyeball, the choroid layer (Fig. 7.5) is covered with a very sensitive optically absorbing pigmented epithelium which bleeds and detaches easily from the retina if physically disturbed by overpressures of ~10 mmHg. Above and covering the pigment layer lie about 250 million rod cells and 6 million cone cells. Rod and cone cells communicate by electronic conduction, not by action potentials. Rod and cone receptor cells converge in a complex synaptic network on bipolar and horizontal cells. Bipolar cells are depolarizing (inhibited by rod/cone neurotransmitters) or hyperpolarizing (excited by rod/cone neurotransmitters), allowing the transmission of positive and negative signals to amacrine and ganglion cells. Horizontal cells laterally connect rods and cones to bipolar cells, allowing lateral inhibition in the retina. Amacrine cells connect bipolar cells to ganglion cells. Numerous types of ganglion cells react to contrast borders, intensity changes and color contrasts. Preprocessed visual information flows into ~1 million ganglion cells whose fibers collect into a giant bundle and exit the eyeball under the ~1.5 mm-diameter optic disk or “blind spot”, forming the optic nerve. Convergence is maximal at the periphery of the retina—rods outnumber cones by more than 10:1 in the periphery, and up to $10^3$-$10^4$ rods may report to a single ganglion cell. Convergence is minimal in the fovea, where one cone cell may synapse with a single ganglion cell through a single bipolar cell.

Low-level signal processing aided by amacrine and horizontal cells in the bipolar layer thus reduces the optic data traffic from 256 million to only 1 million channels at the ganglionic axons which lie closest to the nanorobots attached to the inner retinal membrane. The ganglia are spaced from ~3 microns apart bordering the fovea to ~100 microns apart at the farthest periphery of the retina, averaging ~30 microns. For simplicity, if we regard each ganglion as carrying the equivalent of a single pixel of data in a 1000 x 1000 pixel visual field, then a comprehensive retinal ganglionic management system comprising ~$10^6$ neurostimulatory devices installed within each eyeball allows complete control of the entire human visual field with a refresh rate close to the flicker frequency for rods of ~15 Hz, with ocular data transfer near ~$10^7$ bits/sec which is near the practical upper limit for in vivo mobile communication networks (Section 7.3.2). Assuming up to $P_{in} \sim 30$ pW for each neurostimulatory nanorobot (Section 7.4.5.6) a continuously operated retinal ganglionic control system consumes at most ~30 microwatts, comparable to ocular projection systems described in (A) and (B), so energy supply and heat generation are not significant design limitations.

As of 1998, prototype retinal implant components had been installed and tested in rabbits. These test devices delivered ~10 microamps per electrode directly to ganglion cells and produced measurable activity in the visual cortex of the animals’ brains.1044

---

Fig. 7.6. Above, the human retina (redrawn from Vander et al.866 and Feynman et al.838).
Other retinal implant devices were under investigation by Rolf Eckmiller at Universität Bonn, by a Harvard Medical School/MIT collaboration, and by a Japanese group at Nagoya University. Eckmiller expects the first implant in a human volunteer by 2001. Direct electrical stimulation of the human visual cortex was first attempted in 1968 using 100 Hz pulses applied in 200 microsec bursts via an array of 81 electrodes implanted in a blind person’s brain, with poor results.

D. Direct Photoreceptor Stimulation — A single photoemissive nanorobot may be carefully stationed in each of the 256 million rod and cone cells of the human eye. If the nanorobot is positioned in the pigment-rich apical region of the cell, e.g., the ~9-nm thick double-membranous lamellae, or planar structures, of the outer segment of the rod cell; Figure 7.6) its emitted photon is likely to be detected even without collimation. Assuming a very conservative 1% energy conversion efficiency to generate optical photons, 10% absorption efficiency of optical photons within the cell, and a 25 Hz image refresh rate, power consumption is only ~0.01 pW per nanorobot for continuous transmission and ~2.5 microwatts for complete control of the human visual apparatus. Outmessaging rates are limited by the peak capacity of the external communications link, ~10^7 bits/sec for a mobile network using ν = 100 MHz and f_{data} = 10% (Section 7.3.2).

Chemical or mechachemical stimulation may also be used in place of photons. For example, such methods could induce the 11-cis to 11-trans torsional isomerization of the retinal chromophore in rhodopsin, the first step in photoreceptor stimulation, which occurs naturally in ~0.2 picosec. This conformational change opens a calcium channel in the rod cell membrane; the rapid calcium ion influx triggers a nerve impulse, and light is perceived by the brain. (Each rod rises 1-2 microkelvins in temperature during the heat burst following a light flash causing 180-1800 rhodopsin photoisomerizations per rod.)

Direct nanorobot modulation of intracellular concentrations of α-transducin, the G-protein that transduces the light signal in retinal cells, or directly opening an artificial transmembrane Ca++ channel allowing a Ca++ influx (mediated by a membrane-spanning nanorobot) may be even more efficient.

7.4.6.6 Artificial Symptoms

The triggering of artificial symptoms in the human body offers at best a primitive, noisy, misinterpretable, and low bit rate outmessaging channel. Outmessaging protocols may allow a population of in vivo nanorobots to communicate systemswide status (e.g., “low serum oxygen,” “low serum glucose,” “immune system under attack,”) or specific conditions (e.g., “cancer tumor detected in breast,” “breach of intestinal wall detected”) directly with the patient by inducing recognizable physiological cues (e.g., fever, nausea, shivering, tingling, gasping, tinnitus, hypothalamic “reward” center stimulation). But message noise level is high because these signals are readily confused with natural symptoms that may normally communicate the presence of entirely unrelated conditions. One suggestion to improve the signal/noise ratio is to insert artificial genes that code for enzymes producing a harmless but chromatic signaling compound that colors the urine when excreted, or which adds an unusual color to hair or fingernails as a diagnostic telltale. Nanodevices also could “manually” trigger such a gene, or manufacture the colorant themselves. Alternatively, coordinated nanorobot-induced neural firings might produce externally-measurable changes in the cranial radiative electric field. All these sorts of information transfer are very energy inefficient and are about as informative as “Indian smoke signals.” Artificial symptoms have only limited utility as a primitive channel for nanodevice outmessaging.

7.4.6.7 Macroscale Outmessaging Transducers

Outmessaging transducers (Section 7.4.2.7) may also be employed as outmessaging interfaces between in vivo nanorobots and patients. For example, a simple transduction organ (e.g., macroscale transdermal needle) inserted into the bloodstream could chemically or acoustically interrogate passing nanodevices and obtain information from them as they travel by, without the need for nanodevice removal or locomotion. This information may then be transferred out through the skin to a small external receiving device accessible to the patient.

In vivo nanodevices can also join together to form coordinated aggregates which may be large enough to emulate the action of a macroscale communications device. In some instances, this approach will produce only energy-inefficient marginally-feasible results. For example, consider the concept of the subdermal nanospeaker or “talking tattoo.” The minimum acoustic power output that can be heard as a whisper is ~0.1 microwatts (Section 4.9.1.5). From Eqn. 7.7, a 1-micron acoustic radiator driven by 1000 pW of input power at 3000 Hz produces ~4 x 10^-18 watts/radiator. The sound waves then lose 99.9% of their power passing through the skin-air interface (Section 4.9.1.6), so ~22 trillion devices are required to generate the required ~0.1 microwatt (by coherent emission) in order to be heard. But this many 1000-pW devices operated simultaneously produce 22,000 watts, well in excess of the suggested 100-watt whole-body thermogenic maximum. However, if the radiators position themselves near the skull to employ bone conduction into the tympanum, most of the skin-air interface loss is avoided reducing the required radiator count to 22 billion and total power to 22 watts, which might be acceptable. Such radiators may occupy 10% of local tissue volume make a patch ~350 microns deep covering the area of a postage stamp (~1 inch^2). (Skin-air interface losses may also be avoided by acoustic nanoradiators positioned exclusively supraepidermally.) The “talking wristwatch” approach (Section 7.4.2.7) is another possibility.

A more efficient nanorobot-aggregate user interface is the programmable dermal display. Pigment tattoos, port-wine stains, strawberry marks (common hemangiomas), and other birthmarks constitute an existence proof that small biocompatible particles can be permanently implanted in the dermis and do not migrate on timescales of decades or longer. This suggests that dermal displays can be positionably stable over very lengthy periods of time.

Consider a population of ~3 billion display nanorobots embedded 200-300 microns below the surface of the epidermis covering a 6 cm x 5 cm rectangle on the flat part of the back of the hand or on the smooth medial surface of the forearm. The nanorobots are ~1 micron in volume and occupy only 1% of the 300 mm^3 local tissue deployment volume. Each device consumes ~10 pW when generating visible photons of desired colors at a comfortable visible intensity of ~1 pW/micron^2 or ~1 watt/m^2 (Section 5.3.7), assuming an improved 10% ergo-optical conversion efficiency. Visible photons are completely scattered in 10-100 microns but almost none are absorbed, producing a diffuse glow as ~50% of the scattered photons eventually exit the surface of the skin. For installation and stationkeeping, display nanodevices require at least limited mobility, and a few additional nanodevices may be required to assist with computation.

* The tips of rod cells are renewed at the end of each night, the tips of cone cells at the end of each day, as regulated by the diurnal circadian clock (Section 10.1.1).
data storage and external communications, and other housekeeping chores.

If powered by continuously available chemical fuels, the display would only be operable for 0.1-10 sec before exhausting the entire oxyglucose supply present in the limited tissue deployment volume. Consequently, it is necessary to include a large energy storage buffer constituting 40% of device volume (assuming 10^{10} joule/m^3 stored energy density; Section 6.2.3) which allows ~1000 sec of operation if only 20% of all pixels are radiating at any given time. This power can be reabsorbed from local oxyglucose supplies in ~1 day, giving a long-term ~1% duty cycle for the display (~14 min/day) although the buffer can operate the display for up to ~21 minutes before being exhausted. This power restriction may be entirely avoided by adding supplemental power from, say, a wristwatch-sized extradermal acoustic source (Section 6.4.1), a dedicated transvascular energy organ (Section 6.4.4), photoelectric collectors (~30 pW/micron^2 in cloudless noontime sunlight, ~1 pW/micron^2 in a well-lighted room; Sections 4.9.4 and 6.3.6), or by employing a passive (reflective) display which might be satisfactory for some purposes.

The array of 3 billion nanorobots may be programmed to adopt any of many thousands of different displays. Each display configuration is capable of

1. presenting output data received from the larger in vivo nanorobot population scattered throughout the body (via a communication network), and

2. accepting input data from the patient to be conveyed (through the communication network) to appropriate internal nanorobot subpopulations.

Full-motion animation or video may also be projected, up to the 10^7 bit/sec maximum limit of the mobile communication network; fiber networks may allow up to ~10^9 bit/sec data transfer rates (Section 7.3.1). Figure 7.7A shows a small sample of alternative displays (at actual size) from among the many thousands that are conceivable. Acquisition of the information displayed requires a population of

Fig. 7.7A. A selection of possible dermal display screens.
sensory nanodevices distributed throughout the body and linked by a communications network. For instance, the “Vital Signs” panel requires at least one telemetering nanorobot stationed in each of the several hundred arteries of the human body (or possibly behind each vein valve cusp (Fig. 8.3) throughout the venous system), with each device reporting local blood pressure conditions on a periodic basis.

As for the display itself, in the first example (“Messager/Calculator”) there are 60 input keys measuring (0.5 cm)$^2$ each and a 2 cm x 6 cm output panel that can print 7 rows of 30 characters, each character measuring (2 mm)$^2$ with 1 mm of blank space between each line. Line segments used to write characters are 500 microns wide, more than ten times the limit of human visual acuity at a reading distance of 30 cm; ~500 million active nanorobots participate in drawing the images of the 60 input keys. Display nanorobots have a mean separation of ~5 microns in the tissue but remeasure their relative positions at least ten times per second; a finger-touch on an input key (Fig. 7.7B) depresses the skin by ~500 microns, producing massive and easily detectable displacements of the underlying nanodevices from which the chosen key may be inferred. Non-message related skin stretchings and proximate limb flexures are readily distinguished from input fingerings. The display is activated or deactivated by using finger tapping on the epidermis, temporally coded handclaps or other similar means, and becomes invisible under the epidermis when unilluminated and not in use.

The display may be partly obscured by excessive body hair or particularly dark dermal pigmentation. Post-installation bruises, scabs, incisions, scars, or even wrinkling on the back of the hand may slightly or temporarily impair readability. But nanorobot stationkeeping allows maintenance of near-perfect feature geometry during any trauma short of major avulsion or excision wounds involving significant loss of tissue, deep burns, or outright dermal excoriation or flaying.

In 1998 the closest analogous technology was an all-digital chip-mounted ~1 cm$^2$ projection display, created by Texas Instruments, which operated by independently twisting 307,200 individual tiny mirrors, each measuring 16 microns square, through ±10°, reflecting pulses of colored light onto a screen. Wristwatch-sized personal status monitors for regular biomonitoring were being developed by the Defense Advanced Research Projects Agency (DARPA) in the United States, and in the private sector for sports medicine.

### 7.4.7 Outmessaging to External Receivers

Much of the extensive self-diagnostic data that could theoretically be made available to the patient via dermal displays (Fig. 7.7) is probably beyond the comprehension of the typical user and thus will require professional interpretation. All such information may be rapidly downloaded to instruments under the physician’s control via a wide variety of outmessaging transducers including cable connections through transdermal output ports (Section 7.4.2.7), magnetic induction, infrared or acoustic links with internal communications organs (Section 7.3.4), or dermal displays configured for optimal high-capacity optical external data transfers. Physicians might also receive information from in vivo nanorobots which generate subtle physiological patterns (e.g., an artificial electrical or mechanical signal superimposed on the normal cardiac or respiratory “carrier wave”) requiring sensitive clinical diagnostic equipment to detect, though of course great care must be taken to ensure that such imposed patterns are non-nosogenic. Nanorobots could operate an implanted radio transmitter.

Artificial data secretion glands or other distributed data secretion sources may write huge quantities of data onto nonmetabolizable messenger molecules which are then excreted by the patient, recovered from the urine by laboratory nanodevices, then passed through a reading device, thus making the information available to the physician. A minimum urine flow of ~120 cm$^3$/day containing D$_{message} \approx 26$ bits/nm$^3$ messenger molecules (Section 7.2.1.1) at an innocuous concentration of only 100 parts per million by volume allows a data stream of ~4 x 10$^{15}$ bits/sec to flow from patient to physician—subject to unavoidable time delays for excretion, materials processing and message extraction. The total power requirement to originally encode this information in vivo is 1-1000 zJ/bit (Section 7.2.6) or 4-4000 microwatts for the entire stream, well within safe whole-body thermogenic limits. This technique may be especially valuable in long-term whole-body monitoring, neural recording and archiving (Chapter 25), and related data-intensive applications that require acquisition of information streams from millions or billions of scattered individual in vivo data-gathering sources.

Many of the inmessaging techniques summarized in Section 7.4.1 may also be operated in reverse to achieve outmessaging to the physician from in vivo nanodevices.
7.4.8 Transvenue Outmessaging

Can nanorobots stationed inside one patient communicate with nanorobots stationed inside another patient? Other than simple linkages of previously described standard inmessaging and outmessaging modalities, remote communication between two fully physically disjoint sets of intracorporeal nanorobots can probably be ruled out. However, in vivo and epidermal nanorobots can in theory be programmed to take advantage of serendipitous opportunities using mobile communicytes. Such opportunities for extracorporeal contacts may include patient-to-patient communicytes exchange by direct physical contact (e.g., handshaking or sexual activity), by indiscriminate broadcast transfers (e.g., sneezing, bleeding, desquamation, or sharing tools or utensils), by serendipitous anonymous contacts (e.g., doorknobs, public toilet seats, library books), or by deliberate airborne nanorobot migrations (Section 9.5.3). Security protocols (Chapter 12) are especially important in these applications.

Data is also readily transferred purposefully between patients, as for example via matching dermal communications patches embedded in the palms of each patient’s hands which are briefly pressed together while shaking hands. In 1998, IBM's Personal Area Network (PAN) used near-field oscillating galvanic potentials with a low carrier frequency (-330 KHz), inducing ~50 picoamp currents from a ~1.5 milliwatt power source, to exchange a few hundred bytes of personal data between users during a handshake.

Communication among large numbers of airborne exodermal nanorobots (Section 9.5.3) is a particularly interesting and useful application (e.g., Chapter 21). Such communication may be accomplished by chemical means (Section 8.6.2), but here we will review only optical and acoustic methods.

Consider a cloud of aerial nanorobots with number density \( n_{\text{bot}} \) (nanorobots/m\(^3\)) and mean separation \( X_{\text{path}} = n_{\text{bot}}^{-1/3} \) (meters) that is stationkeeping around a human user (Section 9.5.3). Each nanorobot has a communications energy budget of \( P_{\text{comm}} \) (watts) and broadcasts photons (or sound waves) of frequency \( \nu \) from a circular transmitter of radius \( r_{\text{antenna}} \) to a receiver of similar size located a mean distance \( X_{\text{path}} \) away.

In the case of optical communication, the information transfer rate \( I \) may be regarded as the number of photons received per second from a neighboring device, or

\[
I = \frac{e\%}{4h\nu X_{\text{path}}} \text{ (bits/sec)} \quad \{\text{Eqn. 7.21}\}
\]

where \( e\% \) is the ergophotonic efficiency of photon generation and \( h = 6.63 \times 10^{-34} \text{ joule-sec} \) (Planck’s constant). Conservatively taking \( e\% = 0.01(1\%) \), \( P_{\text{comm}} = 10 \text{ pW} \), \( r_{\text{antenna}} = 1 \mu m \), and assuming red photons with \( \nu = 6.3 \times 10^{14} \text{ Hz} \) (\( \lambda = 700 \text{ nm}, \sim 280 \text{ zl/photon} \)), then inter-nanorobot \( I = 10 \text{ bits/sec} \) for \( X_{\text{path}} = 100 \mu m \), or \( n_{\text{bot}} = 10^{12} \text{ nanorobots/m}^3 \). A nanorobot cloud of this density should appear visually dense, possibly nearly opaque, since \( 10^{12} \text{ nanorobots/m}^3 \) looks through a \( 1 \text{ m}^3 \) cubic cloud having a face area of \( 1 \text{ m}^2 \). The maximum broadcast optical intensity of a ~1 meter column uniformly distributed at this number density is \(-0.1 \text{ watts/m}^2\), which is safe for the human eye. Nanorobot cloud waste heat at this number density, due to communications activities, is \(-10 \text{ watts/m}^3\).

In the case of acoustic communication, the efficiency of ergoacoustic transmission efficiency \( e\% = P_{\text{comm}} / P_{\text{in}} \) as defined by Eqn. 7.7, and the input power budget for communication \( P_{\text{comm}} = P_{\text{in}} \) from Eqn. 7.9, except that in air the exponential attenuation must be replaced with \( e\% \alpha_{\text{air}} \nu^{-1} X_{\text{path}} \) as described in Section 4.9.1.3 for pure fluids. Using variables as previously defined, then \( I = \nu \cdot \text{duty} \) and

\[
P_{\text{comm}} = \frac{24}{\pi^2} \frac{\nu}{k} \frac{X_{\text{path}}^2}{\eta \nu_s \kappa T} \left( e\% \nu \cdot \text{duty} \right) \left( e\% \nu \cdot \text{duty} \right) \text{ (watts)} \quad \{\text{Eqn. 7.22}\}
\]

which is valid in pure fluids if \( r_{\text{antenna}}/\lambda < 1 \). We assume \( \eta = 0.018 \times 10^{-3} \text{ kg/m-sec} \) and \( \nu_s = 343 \text{ m/sec} \) for air at 20°C,763 kT ~ 4.3 Hz, \( \alpha_{\text{air}} = 2 \), \( \kappa = 2 \text{ for a piston radiator}, \rho = 1.3 \text{ kg/m}^3 \), \( r_{\text{antenna}} = 1 \mu m \), and \( \alpha_{\text{air}} = 1.4 \times 10^{-10} \text{ sec}^2/\text{m} \) (Table 4.2). For any given \( X_{\text{path}} \), there is an optimum transmission frequency \( \nu \) that minimizes \( P_{\text{comm}} \). Taking \( X_{\text{path}} = 100 \mu m \) \((n_{\text{bot}} = 10^{12} \text{ nanorobots/m}^3)\), internanorobot \( I = 40 \text{ bits/sec} \) at an optimum frequency of 4.2 MHz and \( \text{duty} = 0.001% \text{ duty cycle} \) (\( P_{\text{comm}} = 7 \text{ pW}\)). Note that Eqn. 7.22 is extremely sensitive to antenna size. At \( r_{\text{antenna}} = 2 \mu m \), \( I = 2000 \text{ bits/sec} \) at 4.2 MHz for \( X_{\text{path}} = 100 \mu m \).

What is the maximum acoustic intensity at the ear? Measuring from the outermost boundary of the nanorobot cloud and moving inward, ultrasound power extinguishes rapidly, falling to an intensity \( e\% (\sim 37%) \) in a distance \( X_{\text{ear}} = (2 \alpha_{\text{air}} \nu^3)^{1/3} \) ~ 200 \text{ microns at } \nu = 4.2 \text{ MHz} \). If the nearest boundary of the cloud is maintained a distance \( X_{\text{ear}} \) from the tympanic membrane of the human ear, then the acoustic intensity reaching the eardrum surface is approximately:

\[
I_{\text{cloud}} = P_{\text{comm}} \eta n_{\text{bot}} X_{\text{path}} \nu X_{\text{path}} \text{ (watts/m}^2\text{)} \quad \{\text{Eqn. 7.23}\}
\]

If the cloud fills the auditory canal, then \( X_{\text{ear}} = 0 \) and \( I_{\text{cloud}} = 10^{-3} \text{ watts/m}^2 \) (\(-93 \text{ dB}, \text{“shouting”}\), which is probably safe enough for the human ear and is inaudible at 4.2 MHz. At \( X_{\text{ear}} = 3 \text{ mm}, I_{\text{cloud}} = 10^{-9} \text{ watts/m}^2 \) \((-30 \text{ dB}, \text{“whispering”}\); if the cloud remains entirely outside the auditory canal, then \( X_{\text{ear}} = 2 \text{ cm} \) and \( I_{\text{cloud}} = 10^{-46} \text{ watts/m}^2 \).

Photic or acoustic modalities may serve as the basis for ex vivo mobile communication networks (Section 7.3.2). Since bit rates are significantly enhanced by proximity (e.g., smaller \( X_{\text{path}} \)), it may be useful to assign circumcorporeal animal nanorobots to condensed geometric or traffic patterns, and to allow opportunistic conferencing among them as required, which might appear to the user as irregular or stellate aggregations, diaphanous tendrils, ordered rows or clump arrays at Cartesian grid vertices, evanescent toroidal cloudlets, or other floating geometries conducive to rapid and effective data sharing. In the earlier examples given above, reducing \( X_{\text{path}} \) to 10 microns raises the optical transfer rate to 900 bits/sec and the acoustic transfer rate to 50,000 bits/sec at 13.4 MHz, for \( r_{\text{antenna}} = 1 \mu m \).

These modalities could also be used by a circumcorporeal nanorobot cloud to communicate visual and auditory messages directly to the user.

\*Airborne nanorobots traveling at a \( \Delta \nu - 1 \text{ m/sec} \) relative velocity will incur a Doppler shift in the received frequency of \( \nu \left(1 + (\Delta \nu / \nu_{\text{sound}}) \right) \sim (1.003) \nu \), defining a ± 3 KHz per MHz band bracketing the intended frequency.
8.1 Navigating the Human Body

It is difficult to imagine any significant application of medical nanodevices which does not involve navigation, however crude. Devices intended to monitor somatic states, assemble artificial internal structures, remove tumors or foreign matter, combat infections, or perform repairs, must normally be extremely tissue- or cell-specific. Navigation is also required to execute many control protocols (Chapter 12), to locate dedicated energy, communication, or navigational helper organs, or to stationkeep and coordinate with other nanodevices. Even bloodborne nanorobots intended to operate solely at the systemic level—such as nanobiots or immunocytes (Chapter 19) and respirocytes (artificial red cells; Chapter 22)—must know if they have been prematurely ejected from the vasculature so that they may cease functioning or at least modify their activities.

Perhaps the most important challenge of in vivo navigation is to determine how physicians may best direct nanorobots to specific target sites needing treatment within the human body. Two alternative strategies appear most likely to produce the best clinical results (both of which will be considered in this Chapter).

The first strategy is positional navigation, in which the nanorobot knows its position inside the human body to ~micron accuracy at all times in some clinic-centered or body-centered coordinate grid system. The nanodevice relies upon dead reckoning, cartotaxis, microtransponder network alignment, or triangulation on external beacon signals to establish its position continuously. This method requires some onboard computation, at least a basic set of sensors (e.g., acoustic), and probably also a good clock (Section 10.1). However, it is hardly foolproof—if the target coordinates are poorly specified or the beacon signals are misaligned or poorly calibrated, the nanorobots may go to the wrong place.

The second strategy is functional navigation, in which nanorobots seek to detect subtle variations in their environment, comparing diverse sensor readings with the profile of the target tissue or cell and congregating wherever this very precisely defined set of preconditions exists. These preconditions may be thermal, acoustic or barostatic, cytotoxic or immunochemical, mechanical or topological, or even genetic. The crudest forms of functional navigation may be called demarcation, wherein the doctor manually creates detectable artificial conditions at or near the target site such as dermal hot spots, injected chemical plumes, or focused ultrasonic beam spots of appropriate magnitudes and frequencies. Demarcation strategies can be implemented using extremely simple onboard sensors and control devices, possibly not even requiring a nanorobot computer, thus may prove useful early in nanomedical technology development.

More sophisticated forms of functional navigation can be extraordinarily flexible because targets may be specified without the physician having to know their exact physical location in the body—e.g., nascent cancer tumors, T cells reactive to specific antigens, infected deep-thoracic lymph nodes, bacteria of a particular species, broken capillary vessels, or virus particles having a specified protein coat chemistry. The physician need not know the exact number of targets, nor even if any targets are present at all. These forms of functional navigation require onboard computation, a resident database of relevant parameters and operational details, and a wider assortment of sensors and control protocols. But they also offer the greatest benefit at lowest risk for the patient and thus must be regarded as the preferred approach once the technology is available.

This Chapter opens with a survey of human somatography (Section 8.2), followed by general discussions of positional (Section 8.3) and functional (Section 8.4) navigation, cytonavigation (Section 8.5), and finally ex vivo navigation (Section 8.6).

8.2 Human Somatography

Somatography is, quite simply, the “geography” or map of the anatomical spaces, as seen from the viewpoint of a microscopic traveler in those realms. The human body is a complex and fascinating place to visit. The nanomedical theater of operations in adult patients may range in size from an extremely small ~0.005 m³ (Lucia Zarate, 5.9 kg, in January 1883 at age 20) to an extremely large ~0.5 m³ (Robert Earl Hughes, 485 kg, in February 1958 at age 32), but averages ~0.06 m³ of navigable volume for the standard 70 kg adult male having 15% body fat, medium build and good health. However, it is hardly foolproof—if the target coordinates are poorly specified or the beacon signals are misaligned or poorly calibrated, the nanorobots may go to the wrong place. The more typical tall but overweight late 20th-century American man (6-ft, 220 lbs, 25% body fat) measures ~0.1 m³ in size.) Thus the volume normally available for nanorobot navigation in a single patient is ~10³ microns³.

The first fully digitized comprehensive three-dimensional map of the human body was compiled in the early 1990s as part of the National Library of Medicine’s Visible Human Project. During this effort, a male and female cadaver were frozen in blocks of gel, sectioned into thousands of thin slices and digitally scanned in MRI (magnetic resonance imaging), CT (computerized x-ray tomography) and anatomical modes (optical photography), slice by slice. The male data set, released in 1994, consists of axial MRI images of the head and neck taken at 4 mm intervals and longitudinal sections of the rest of the body also at 4 mm intervals; each MRI image is 256 x 256 pixels with a 12-bit grey scale per pixel. The CT data consists of 512 x 512 pixel axial CT scans of the entire body taken at 1 mm intervals, also with a 12-bit grey scale per pixel. The axial anatomical images are 2048 x 1216 pixels with a 24-bit color scale per pixel, representing 60 megabits per image. The anatomical cross-sections are also at 1 mm intervals and coincide with the CT axial images. There are 1871 cross-sections for each mode (CT and anatomy) obtained from the male cadaver, a 0.112 terabit data set for the anatomical images (~111 million micron³ per voxel).
The female data set, released in 1995, has the same characteristics as the male set except that the axial anatomical images were obtained at 0.33 mm intervals instead of 1.0 mm intervals, resulting in more than 5,000 anatomical images and a data set of ~0.336 terabits. Even though this map has a resolution of only 37 million micron\(^3\) per voxel, this is still sufficient to resolve all terminal veins and terminal arterial branches in the body and all major gross anatomical features, and thus provides a good start toward a human somatographic atlas.

A complete human somatographic static map to cellular resolution in theory requires ~20 micron increments, a ~100 terabit data set using 8-bit 8000 micron\(^3\) voxels assuming a fixed scanning geometry and ignoring the indexing tables. Recording all major structural details in the capillary terminal bed demands ~4 micron resolution, a ~10,000 terabit data set assuming 8-bit voxels. A 10,000 terabit data set may be stored on a ~26 bit/nm\(^3\) hydrofluorocarbon memory tape (Section 7.2.1.1) in a cubic volume of ~72 micron\(^3\) within an in vivo ~100 micron\(^3\) library nodule. Divided into ~10\(^6\) independent spools, a datum located anywhere on the 3300-kilometer total tape length could be accessed in ~10 seconds assuming a read speed of ~30 cm/sec (Section 7.2.6). A smaller 100 terabit library nodule requires only ~16 micron\(^3\) of tape, with ~1000 spools and similar access times as in the previous example. Topological or functional mapping may permit data compression by a factor of 10-100 (e.g., Sections 8.2.1.2 and 8.3.2) without increasing access time or seriously reducing map utility.

A 1 micron\(^3\) storage block that could conveniently be carried aboard an individual medical nanorobot can hold ~0.01 terabits (Section 7.2.6), enough memory to contain a three-dimensional map of the entire human body to ~430 micron resolution or a map of a 1-kg organ to ~100 micron resolution, using 8-bit voxels. The required resolution for a given application is very mission-dependent. Additionally, two types of map are likely to be of value. The first is a precise static map of some generic reference cadaver, as described above, that may provide general navigational guidance. The second is a detailed map of the individual patient, assembled by exploratory or “surveyor” nanorobots (Chapter 19) prior to the deployment of therapeutic nanorobots which would be given this information to allow very specific navigational guidance. Map stability is an important issue (Section 8.2.1.2, Chapter 19).

The remainder of this Section offers a quick guided tour of the larger navigable volumes inside the human body, along with some useful quantitative details. The different regions of the body appear to have enough structural and chemical dissimilarity to allow a nanorobot traversing these volumes to determine its position to at least ~mm accuracy, based solely on simple landmark recognition, chemonavigational cues, and bifurcation and topological information, even without resorting to the precision positional systems described in Section 8.3.

### 8.2.1 Navigational Vasculography

Medical nanorobots may access the interior of the human body principally via the blood-carrying vasculature, representing ~5.4 liters or ~9% of total body volume. The blood vessels are of three types, each characterized by functional and structural differences. First, there is the high-pressure distributing system, made up of the arteries and smaller branches called arterioles, which convey oxygenated blood from the heart to all regions of the body (Section 8.2.1.1). Second, there is a network of minute vessels, called capillaries, through which biologically important materials are exchanged between blood and tissues (Section 8.2.1.2). Third, there is a low-pressure collecting system made up of veins and smaller branches called venules, which return the blood to the heart (Section 8.2.1.1). The lymphatic system, an additional 3.3 liters of very low-pressure vasculature (Section 8.2.1.3), returns cell-filtered plasma to the main circulatory system.

#### 8.2.1.1 Arteriovenous Macrocirculation

Topologically, the arteriovenous circulatory system resembles a “figure-eight” with a four-chambered heart positioned at the central junction. In the top half of the “figure-eight,” oxygen-depleted blood is pumped from the heart to the lungs via the pulmonary artery. Oxygen-rich blood leaves the lungs and returns to the heart via the left and right pulmonary veins. In the bottom half of the “figure-eight,” oxygen-rich blood received from the lungs is pumped into the aorta for distribution to the tissues, as shown in Figure 8.1. Oxygen-depleted blood is collected by the venous network (Fig. 8.2) and returns to the heart via the vena cava. (In both Figures, superficial vessels are drawn as solid lines and deep vessels as broken lines.) Most veins 1 mm in diameter or larger are fitted with a series of one-way valves to prevent backflow (Fig. 8.3). Such valves are most numerous in the lower extremities: the vena cava and the mesenteric, pulmonary and portal veins. Arteries have no valves. The smaller arteries are heavily Anastomosed except across the midline of the body.

Table 8.1 provides a summary of the ~19,000 kilometers of human arteriovenous vasculature (mostly capillaries); additional data for a few selected major vessels is in Table 8.2. Singhal et al have carefully surveyed the complete branching structure of the pulmonary arterial network, spanning the entire range from principal artery to final capillary (Table 8.3). These data may provide a useful model for estimating capillary branching systems in other tissues, bearing in mind that pulmonary capillaries are generally wider and shorter than capillaries found elsewhere in the body, and that capillary bed geometries are unique to each organ (Section 8.2.1.2). The measured thickness of the glycocalyx of the endothelium of the systemic

![Fig. 8.1. Human arterial system.](image)
arteries and vena cava may also be utilized for navigational purposes. For example, the thickness of rabbit glyocalyx ranged from 45 \pm 1 nm in the coronary artery to 81 \pm 2 nm in the carotid; the glyocalyx was 20 \pm 1.5 nm thicker on the downstream side of intercostal ostia than on the upstream side. 

How much time does a nanorobot flowing with the blood take to transit individual organs and the entire vascular circuit? It is generally accepted that the round-trip circulation time through the entire “figure-eight” averages ~60 seconds under resting conditions, a time that appears constant for nearly all mammals. During heavy exercise, the minimum circulation time may fall to 11-15 sec along most pathways primarily due to vasodilation.

Tissue transit times have been widely studied. For example, bloodborne particles in the lung normally take only 1-2 seconds to pass the alveolar sheet but 5-10 seconds to transit the entire pulmonary vasculature. Lung transit time is not significantly reduced if heart rate is simply increased; during heavy exercise, the minimum circulation time may fall to 11-15 sec along most pathways primarily due to vasodilation.

Tissue transit times have been widely studied. For example, bloodborne particles in the lung normally take only 1-2 seconds to pass the alveolar sheet but 5-10 seconds to transit the entire pulmonary vasculature. 

8.2.1.2 Arteriovenous Microcirculation

In most cases, the true capillaries do not directly join arterioles to venules. Rather, oxygenated blood passes out of the terminal branches of the arterioles (terminal arterioles are -10-50 microns in diameter) into metarterioles (10-20 microns), “preference channels” whose walls contain smooth muscle fibers in declining numbers from proximal to distal, which eventually merge with the non-contractile postcapillary (8-30 micron) and collecting (10-50 micron) venules (Fig. 8.4). Capillaries branch off directly from the metarterioles through precapillary sphincters constructed of single muscle fibers; other than these, there are no contractile elements in the capillaries. The state of contraction of the sphincters controls the rate of fluid flow through the capillaries. In skeletal muscle, with its widely varying oxygen requirements, there are 8-10 capillaries per metarteriole; in the mesenteric circulation, a stable metabolism requires only 2-3 capillaries per metarteriole; in the nail bed, the ratio is 1:1. The topology of the arterioles that supply the capillaries, and of the venules that drain them, is usually special for each tissue. Some beds are structured as trees, while others are organized into arcades, sinuses, or portal systems. Thus the vasculature of each organ is unique—useful navigational information for medical nanorobots.

The wall of a capillary vessel is a tube consisting of a single layer of rolled-up endothelial cells. Endothelial cells may be less than 1 micron thick with a flat surface area of 300-1200 micron², so the entire ~313 meter² vascular surface may be comprised of up to 0.25-1 trillion endothelial cells. Endothelial cells are laterally apposed to each other with a 10-20 nm gap between neighboring cell membranes.

1. First, there is the continuous type with cells joined tightly together. Striated muscles may have flat thin cells, while in postcapillary venules the cells are cuboidal, forming a thick layer.

2. Second, there is the fenestrated type, with cells so thin that internal vesicles form small pores 25 nm thick and 100 nm in diameter (typically ~100 pores/micron², compared to ~190 pores/micron² in arteriole endothelium and ~645 pores/micron² in postcapillary venule endothelium). This type is found in the renal medulla, endocrine glands, and in structures engaged in the production or absorption of fluids such as the renal glomerulus, choroid plexus of the brain, and the intestinal villus.

3. Third, there is the discontinuous type with distinct intercellular gaps and a broken basement membrane, commonly found in sinusoid vessels and in organs such as the liver, spleen, and bone marrow whose functions include injection or extraction of whole cells, large molecules and extraneous particles from the blood.

Capillaries average 8 microns in diameter, but what is the minimum human capillary diameter? This question is important because it sets an upper limit on the size of bloodborne nanorobots. For example, in one experiment 97% of all 15-micron radiolabelled
microspheres reaching the eye were trapped during the first pass.\textsuperscript{842} Standard anatomy and physiology textbooks variously report that the minimum capillary lumen measures 5-7 microns,\textsuperscript{834,839} 4-9 microns,\textsuperscript{836} 4-8 microns,\textsuperscript{517} or 4-6 microns\textsuperscript{841,843} in diameter. There are also references to capillaries as narrow as 3 microns for laboratory rodents and other mammals.\textsuperscript{834,840} Theoretical calculations\textsuperscript{841,844} suggest that a cylindrical vessel must be at least 2.7 microns in diameter to allow maximally deforming human red cells normally averaging 7.82 microns in diameter\textsuperscript{362} to pass. This theoretical minimum assumes a mean cell (surface) area (MCA) of ~135 micron\textsuperscript{2} and a mean cell volume (MCV) of ~94 micron\textsuperscript{3} for human erythrocytes, and allows for a maximum surface stretch of 10 micron\textsuperscript{2} to a total of MCA ~ 145 micron\textsuperscript{2}.\textsuperscript{*} From simple geometry for a red cell compressed into a hemisphere-capped cylinder during tube passage, the minimum tube diameter $D_{\text{tube}}$ is given by

$$MCV = \left( \frac{MCA}{4} \right) D_{\text{tube}} - \left( \frac{\pi}{12} \right) D_{\text{tube}}^3 \quad (\text{m}^3) \quad [\text{Eqn. 8.1}]$$

Experiments using polycarbonate sieves and micropipettes confirm that a tube diameter of at least 2.3 microns is necessary to avoid plugging.\textsuperscript{374} However, the above theoretical computation assumes mean erythrocyte geometry and ignores the considerable dispersion in volume and surface area of individual physiological red cells around the mean value. Specifically, the largest 1% of human red cells have diameter $d_{\text{RBC}}$ ~ 9.65 microns, MCA ~ 182 micron\textsuperscript{2} and MCV ~ 132 micron\textsuperscript{3},\textsuperscript{362} for example, dogs have 7.2-micron red cells with MCA ~ 123 micron\textsuperscript{2} and MCV ~ 69 micron\textsuperscript{3}, giving (assuming surface elasticity characteristics similar to the human erythrocyte) a theoretical minimum passable tube diameter of 2.4 microns. Cats have 5.6-micron red cells with MCA ~ 83 micron\textsuperscript{2} and MCV ~ 43 micron\textsuperscript{3}, giving $D_{\text{RBC}}$ ~ 2.2 microns, a full 0.5 micron less than for human red cells.

**Table 8.1. Approximate Quantification of the Human Arteriovenous System\textsuperscript{830-835,836,838,1616}**

<table>
<thead>
<tr>
<th>Human Blood Vessel</th>
<th>Diameter (mm)</th>
<th>Length (mm)</th>
<th>Wall Thickness</th>
<th>Internal Pressure</th>
<th>Number of Vessels</th>
<th>Total Length (mm)</th>
<th>Total Surface Area (m\textsuperscript{2})</th>
<th>Total Blood Volume (m\textsuperscript{3})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aorta</td>
<td>25.0</td>
<td>400</td>
<td>1500 (\mu)m</td>
<td>100 mmHg</td>
<td>1</td>
<td>400</td>
<td>31,400</td>
<td>200,000</td>
</tr>
<tr>
<td>Large arteries</td>
<td>6.5</td>
<td>200</td>
<td>1000 (\mu)m</td>
<td>100 mmHg</td>
<td>40</td>
<td>8000</td>
<td>163,000</td>
<td>260,000</td>
</tr>
<tr>
<td>Main artery branches</td>
<td>2.4</td>
<td>100</td>
<td>800 (\mu)m</td>
<td>95 mmHg</td>
<td>500</td>
<td>50,000</td>
<td>377,000</td>
<td>220,000</td>
</tr>
<tr>
<td>Terminal artery branches</td>
<td>1.2</td>
<td>10</td>
<td>125 (\mu)m</td>
<td>90 mmHg</td>
<td>11,000</td>
<td>110,000</td>
<td>415,000</td>
<td>120,000</td>
</tr>
<tr>
<td>Arterioles</td>
<td>0.1</td>
<td>2</td>
<td>20 (\mu)m</td>
<td>60 mmHg</td>
<td>4,500,000</td>
<td>9,000,000</td>
<td>2,800,000</td>
<td>70,000</td>
</tr>
<tr>
<td>Capillaries</td>
<td>0.008</td>
<td>1</td>
<td>1 (\mu)m</td>
<td>30 mmHg</td>
<td>19,000,000,000</td>
<td>19,000,000,000</td>
<td>298,000,000</td>
<td>375,000</td>
</tr>
<tr>
<td>Venules</td>
<td>0.15</td>
<td>2</td>
<td>2 (\mu)m</td>
<td>20 mmHg</td>
<td>10,000,000,000</td>
<td>20,000,000</td>
<td>9,400,000</td>
<td>355,000</td>
</tr>
<tr>
<td>Terminal veins</td>
<td>1.5</td>
<td>10</td>
<td>40 (\mu)m</td>
<td>15 mmHg</td>
<td>11,000</td>
<td>110,000</td>
<td>518,000</td>
<td>190,000</td>
</tr>
<tr>
<td>Main venous branches</td>
<td>5.0</td>
<td>100</td>
<td>500 (\mu)m</td>
<td>15 mmHg</td>
<td>500</td>
<td>50,000</td>
<td>785,000</td>
<td>1,590,000</td>
</tr>
<tr>
<td>Large veins</td>
<td>14.0</td>
<td>200</td>
<td>800 (\mu)m</td>
<td>10 mmHg</td>
<td>40</td>
<td>8000</td>
<td>352,000</td>
<td>1,290,000</td>
</tr>
<tr>
<td>Vena cava</td>
<td>30.0</td>
<td>400</td>
<td>1200 (\mu)m</td>
<td>5 mmHg</td>
<td>1</td>
<td>400</td>
<td>37,700</td>
<td>280,000</td>
</tr>
<tr>
<td>Heart chambers</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>120 mmHg</td>
<td>--</td>
<td>--</td>
<td>450,000</td>
</tr>
</tbody>
</table>

| TOTALS | ~19,000 km | 312,900,000 | 5,400,000 |

* Different species have red cells differing slightly in diameter, thickness, surface area and volume.\textsuperscript{842} For example, dogs have 7.2-micron red cells with MCA ~ 123 micron\textsuperscript{2} and MCV ~ 69 micron\textsuperscript{3}, giving (assuming surface elasticity characteristics similar to the human erythrocyte) a theoretical minimum passable tube diameter of 2.4 microns. Cats have 5.6-micron red cells with MCA ~ 83 micron\textsuperscript{2} and MCV ~ 43 micron\textsuperscript{3}, giving $D_{\text{RBC}}$ ~ 2.2 microns, a full 0.5 micron less than for human red cells.

**Table 8.2. Additional Data on Selected Blood Vessels\textsuperscript{836,834}**

<table>
<thead>
<tr>
<th>Blood Vessel or Element</th>
<th>Diameter (mm)</th>
<th>Mean Flow Rate (m\textsuperscript{3}/sec)</th>
<th>Mean Flow Velocity (mm/sec)</th>
<th>Velocity Range (mm/sec)</th>
<th>Mean Tube Reynolds Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascending aorta</td>
<td>23.0-43.5</td>
<td>364,000</td>
<td>630</td>
<td>245-876</td>
<td>3210-6075</td>
</tr>
<tr>
<td>Inferior vena cava</td>
<td>20.0</td>
<td>34,000-50,000</td>
<td>135</td>
<td>107-160</td>
<td>323-482</td>
</tr>
<tr>
<td>Descending aorta</td>
<td>16.0-20.0</td>
<td>54,000-85,000</td>
<td>270</td>
<td>--</td>
<td>1200-1500</td>
</tr>
<tr>
<td>Common carotid</td>
<td>5.9</td>
<td>5100</td>
<td>187</td>
<td>99-388</td>
<td>332</td>
</tr>
<tr>
<td>Carotid sinus</td>
<td>5.2</td>
<td>3300</td>
<td>156</td>
<td>85-325</td>
<td>244</td>
</tr>
<tr>
<td>Femoral artery</td>
<td>5.0</td>
<td>3700</td>
<td>188</td>
<td>-350 to 1175</td>
<td>283</td>
</tr>
<tr>
<td>External carotid</td>
<td>3.8</td>
<td>1800</td>
<td>157</td>
<td>83-327</td>
<td>180</td>
</tr>
<tr>
<td>Large veins</td>
<td>5-10</td>
<td>4000-12,000</td>
<td>150-200</td>
<td>--</td>
<td>210-570</td>
</tr>
<tr>
<td>Large arteries</td>
<td>2-6</td>
<td>1600-5700</td>
<td>200-500</td>
<td>--</td>
<td>110-850</td>
</tr>
<tr>
<td>Small arteries</td>
<td>0.3</td>
<td>3.5</td>
<td>50</td>
<td>--</td>
<td>2.3</td>
</tr>
<tr>
<td>Arterioles</td>
<td>0.025-0.10</td>
<td>0.025</td>
<td>5</td>
<td>--</td>
<td>0.038</td>
</tr>
<tr>
<td>Capillaries</td>
<td>0.004-0.015</td>
<td>0.000075</td>
<td>0.2-1.5</td>
<td>0.00036</td>
<td>0.0027</td>
</tr>
</tbody>
</table>

hypothetical male, age 30, weight 70 kg, blood volume 5.4 liters
Basic Capabilities • Navigation

Table 8.3. Branching Structure of the Pulmonary Arterial Network Entering the Lungs

<table>
<thead>
<tr>
<th>Pulmonary Branching Order</th>
<th>Number of Branches of Each Order</th>
<th>Number of End Branches of Each Order</th>
<th>Length of Pulmonary Vessel (mm)</th>
<th>Diameter of Pulmonary Vessel (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>300,000,000</td>
<td>90.5</td>
<td>30.0</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>100,000,000</td>
<td>32.0</td>
<td>14.83</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>30,210,000</td>
<td>10.9</td>
<td>8.06</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>13,760,000</td>
<td>20.7</td>
<td>5.82</td>
</tr>
<tr>
<td>5</td>
<td>66</td>
<td>3,983,000</td>
<td>17.9</td>
<td>3.65</td>
</tr>
<tr>
<td>6</td>
<td>203</td>
<td>1,159,000</td>
<td>10.5</td>
<td>2.09</td>
</tr>
<tr>
<td>7</td>
<td>675</td>
<td>347,000</td>
<td>6.6</td>
<td>1.33</td>
</tr>
<tr>
<td>8</td>
<td>2290</td>
<td>89,160</td>
<td>4.69</td>
<td>0.85</td>
</tr>
<tr>
<td>9</td>
<td>5861</td>
<td>48,050</td>
<td>3.16</td>
<td>0.525</td>
</tr>
<tr>
<td>10</td>
<td>17,560</td>
<td>16,040</td>
<td>2.10</td>
<td>0.351</td>
</tr>
<tr>
<td>11</td>
<td>52,550</td>
<td>5358</td>
<td>1.38</td>
<td>0.224</td>
</tr>
<tr>
<td>12</td>
<td>157,400</td>
<td>1787</td>
<td>0.91</td>
<td>0.138</td>
</tr>
<tr>
<td>13</td>
<td>471,300</td>
<td>598</td>
<td>0.65</td>
<td>0.086</td>
</tr>
<tr>
<td>14</td>
<td>1,413,000</td>
<td>206</td>
<td>0.44</td>
<td>0.054</td>
</tr>
<tr>
<td>15</td>
<td>4,226,000</td>
<td>67</td>
<td>0.29</td>
<td>0.034</td>
</tr>
<tr>
<td>16</td>
<td>12,660,000</td>
<td>24</td>
<td>0.20</td>
<td>0.021</td>
</tr>
<tr>
<td>17</td>
<td>300,000,000</td>
<td>1</td>
<td>0.13</td>
<td>0.013</td>
</tr>
</tbody>
</table>

(alveolar)

Fig. 8.4. Metarterioles, venules, and lymph capillaries (redrawn from Tortora853).

giving a theoretical minimum passable tube diameter $D_{tube} \approx 3.1$ microns for these cells. The largest cell in a population of $10^8$ cells (there are a total of ~300,000 of these largest cells in circulation at any given time, a not inconsiderable number) has $d_{RBC} \approx 16$ microns, $MCA \approx 500$ micron$^2$, $MCV \approx 450$ micron$^3$ giving $D_{tube} \approx 3.7$ microns. Thus if human capillaries were as small as 2.3-3.7 microns, a significant number of red cells would be unable to navigate passage and would quickly plug the tubes, resulting in angionecrosis. Experiments have also shown plugging by human leukocytes at tube apertures much below ~5 microns. Thus the minimum viable human capillary diameter appears to be ~4 microns.

The average density of capillaries in human tissue is ~600/mm$^3$, which implies a mean separation of ~40 microns between adjacent capillaries averaging ~1 mm in length. Most living tissue cells lie within ~1-3 cell widths of a capillary. To reach a particular cell, bloodborne nanorobots may normally travel most of the distance through the vascular network, then exit the capillary and cross at most one or two cells to reach the desired target cell. Capillary density varies considerably in the vascular beds of different organs. For instance, microvascular density is 2500-3000/mm$^3$ in the brain, kidneys, liver, and myocardium; 300-400/mm$^3$ in phasic units of the skeletal musculature; and <100/mm$^3$ in bone, fat, connective tissue, and in tonic units of the skeletal musculature. Blood flow rate per gram of tissue (Table 8.4) is very roughly proportional to the relative microvascular density in each organ or tissue. More than 70% of the water of the blood is exchanged with extravascular water every minute—the walls of smaller capillaries are veritable sieves with respect to water.

What is the size (in bits) of the minimum topological map needed to reliably navigate the entire ~19,000 kilometers of human blood vasculature? Map size is driven by the number of bifurcations (decision junctions; see Section 8.3.4) and capillaries. Taking the pulmonary arterial tree (Table 8.3) as representative, traveling from the heart to a final capillary requires passage through approximately 17 forks with an average of 3.2 branches per fork. In a simple bifurcation map with 17 forks and at most 4 branches per fork, each capillary can be assigned a unique topological address using 17 digits with each digit having $\log_2(4) = 2$ bits, a 34-bit address vector. This specification is very compact because the minimum number of bits required to name each of 19 billion capillaries in the body is also $\log_2(19$ billion vessels$) \approx 34$ bits. Using run-length coding, we can have $4^i$ addresses of length 34 bits, $4^{i-1}$ addresses of length 32 bits, and so forth, giving a whole-body minimum of: $34 + 32 + ... + 2 = 763,549,741,512$ bits for a single comprehensive bifurcation map.

Assuming some data redundancy (e.g., parity check bits) and an allowance for special cases (e.g., arterial anastomoses, shunts, or arteriovenous fistulas which are direct connections between a small artery and a small vein), a complete map of the human vasculature
and each capillary serves only an excellent framework for determining cytographic location in the cytographic map, the number of bits required to uniquely name within a ~50,000 micron³ volume or ~(40 micron)³, the approximate instructions needed to reach each vessel. This allows localization to capillaries, with each address representing the unique navigational

probably requires ~1 terabit to store the addresses of all 19 billion cells. To make a complete cytographic map, the number of bits required to uniquely name each of ~10^{13} tissue cells in the human body (Section 8.5.1) is \log_2(10^{13}) = 43 bits; given similar allowances as before, \approx 1000 terabits are required to store for individual bloodborne nanorobots to carry a whole-body vasculographic or cytographic map onboard (the latter of which might require \sim 40,000 micron³ of data storage volume). However, a modest 0.1-micron³ 10⁹-bit map will allow a bloodborne nanorobot to retain the addresses (with complete navigational instructions) of up to 30 million individual capillaries (~50 cm³ of typical tissue) or ~20 million individual cells (~0.2 cm³ of cell-dense tissue) for the duration of its mission. Note that sequential passage through ~10 million different named capillaries of average length ~1 mm at a swimming speed of ~1 cm/sec (Section 9.4.2.6) requires ~12 days, a not unreasonable mission duration.

### 8.2.1.3 Lymphatic System

The lymphatic system consists of a branching series of closed endothelium-lined vessels. These vessels slowly drain lymph (composition discussed below) from the extracellular spaces and convey it back to the arteriovenous circulation, a sluggish current easily navigated by lymphborne medical nanodevices. The lymphatic vessels have a total capacity of ~2 liters of lymph,\,852 vs. ~3.7 liters of fluid capacity for all the veins. No detailed quantitative summation of the lymphatic system is readily available in the literature, so Table 8.5 must be regarded as an estimate consistent with data gleaned from a variety of sources,\,863,838,848,850,853-860,874 but possibly still containing significant errors.

Lymphatic vessels resemble veins in structure but are greater in number. Compared to veins, lymphatics of similar size have thinner walls, more valves, much greater variation in their caliber but a less sinuous course through the tissue. They also contain lymph nodes at various intervals along their length. Lymph vessels are frequently connected with one another by short anastomotic branches.\,857 Lymphatics of the skin travel in loose subcutaneous tissue and generally follow veins and venules. Lymphatics of the viscera generally follow arterioles, formingplexuses (extensive two-dimensional meshworks) around them.

However, the lymph drainage system actually begins with the prelymphatics. The prelymphatics are a randomly interconnected network of non-endothelialized tissue channels 0.1-0.2 microns in diameter penetrating the cell masses of the body, although in certain regions where lymphatic capillaries are absent (e.g., the brain, eye, and bone), the prelymphatics take over the role of the capillaries and eventually discharge into them, outside the regions in question.\,861 In the viscera, these interstitial channels will normally be very short since there are frequent arteriovenous and lymphatic capillaries—for example, ~30 microns long in the capsules of knee joints (~0.02-0.05 channels/micron³). In the muscles, the prelymphatics are longer since the lymphatic capillaries extend only into the larger regions of connective tissue. In the brain the channels may be 10-30 cm long since they must reach (in the case of the brain) from the depths of the cortex to the outside of the skull.\,861 Following injury, tissue channel volume may enlarge as much as ~400 times in a week, in response to a simple incision wound.\,1339 Drainage of the prelymphatics is done by the lymphatic capillaries, thin-walled irregularly-contoured valveless endothelial tubes varying in diameter from 15-75 microns. The lymphatic capillaries are blind sacs (Fig. 8.4) with an inferred mean separation of ~86 microns or ~4 tissue cell widths. Their walls are more porous than those of blood capillaries, so that larger molecules and particles may pass through them (Fig. 8.5). The resting gap of the open junction of adjoining lymphatic endothelial cells usually ranges from ~0.1 micron to several microns.\,861 However, Allen\,851 injected intraperitoneally a variety of particles of various sizes up to 22.5 microns in diameter, and all sizes later appeared in the diaphragmatic lymph. This suggests that the peritoneal mesothelium and the lymphatic endothelium on either side of the fenestrations of the basement membrane can open

### Table 8.4. Typical Blood Perfusion Rates for Various Tissues and Organs\,818,837

<table>
<thead>
<tr>
<th>Tissue Type</th>
<th>Anatomical Location or Organ</th>
<th>Specific Blood Flow Rate (mm³/sec-gm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adipose tissue</td>
<td>Abdomen, 10-29 mm thick</td>
<td>0.507</td>
</tr>
<tr>
<td></td>
<td>Abdomen, 30-49 mm thick</td>
<td>0.358</td>
</tr>
<tr>
<td></td>
<td>Abdomen, &gt;40 mm thick</td>
<td>0.307</td>
</tr>
<tr>
<td></td>
<td>Thigh, 11 mm thick</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>Thigh, 20 mm thick</td>
<td>0.33</td>
</tr>
<tr>
<td></td>
<td>Thigh, 43 mm thick</td>
<td>0.15</td>
</tr>
<tr>
<td>Bone</td>
<td>Humerus, marrow flow only</td>
<td>0.055</td>
</tr>
<tr>
<td>Connective tissue</td>
<td>Typical, basal (max)</td>
<td>0.50 (2.5)</td>
</tr>
<tr>
<td>Joint</td>
<td>Knee, avg. @ 303 K skin temp.</td>
<td>0.222</td>
</tr>
<tr>
<td></td>
<td>Knee, avg. @ 308 K skin temp.</td>
<td>0.487</td>
</tr>
<tr>
<td>Muscle</td>
<td>Calf, anterior, resting flow (max)</td>
<td>0.46 (9.15)</td>
</tr>
<tr>
<td></td>
<td>Forearm, resting flow (max)</td>
<td>0.53 (8.38)</td>
</tr>
<tr>
<td></td>
<td>Thigh, anterior, resting flow (max)</td>
<td>0.43 (6.00)</td>
</tr>
<tr>
<td></td>
<td>Typical, basal (max)</td>
<td>0.50 (10.0)</td>
</tr>
<tr>
<td>Organ</td>
<td>Brain, basal (max)</td>
<td>9.0-9.2 (18.3)</td>
</tr>
<tr>
<td></td>
<td>Gastrointestinal tract, basal (max)</td>
<td>6.7 (26.7)</td>
</tr>
<tr>
<td></td>
<td>Heart, basal (max)</td>
<td>13.3-14.0 (64.0)</td>
</tr>
<tr>
<td></td>
<td>Kidney, basal (max)</td>
<td>67-70 (100)</td>
</tr>
<tr>
<td></td>
<td>Liver, basal (max)</td>
<td>9.6-14.2 (54.5)</td>
</tr>
<tr>
<td></td>
<td>Lung, basal (max)</td>
<td>90 (490)</td>
</tr>
<tr>
<td>Skin</td>
<td>Abdomen, resting flow</td>
<td>1.44</td>
</tr>
<tr>
<td></td>
<td>Arms, resting flow</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>Calf, resting flow</td>
<td>1.77</td>
</tr>
<tr>
<td></td>
<td>Face, resting flow</td>
<td>1.17</td>
</tr>
<tr>
<td></td>
<td>Foot, dorsal surface, resting flow</td>
<td>2.38</td>
</tr>
<tr>
<td></td>
<td>Forearm, sunburned (max)</td>
<td>9.22 (46.7)</td>
</tr>
<tr>
<td></td>
<td>Hands, resting flow</td>
<td>3.35</td>
</tr>
<tr>
<td></td>
<td>Head, resting flow</td>
<td>7.15</td>
</tr>
<tr>
<td></td>
<td>Thigh, resting flow</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>Thorax, resting flow</td>
<td>6.45</td>
</tr>
<tr>
<td></td>
<td>Typical, resting flow (max)</td>
<td>1.7 (25.0)</td>
</tr>
</tbody>
</table>
at least this wide, so micron-sized nanorobots should find easy passage into the lymphatic system through these pores.

Lymph capillaries originate throughout the body, though not in avascular tissue, the central nervous system, splenic pulp or bone marrow. They may occur singly or in plexuses, especially in regions where the connective tissue is lobulated (e.g., glands) or arranged in cylinders (e.g., muscles). Each lymphatic capillary plexus is typically 320-640 microns wide, encompassing 2-6 blood capillary loops. The terminal lymphatic network has different geometric patterns in different tissues. In the mesentery, it conforms to the modular configuration of the blood capillary network. In the skeletal muscle, lymphatics are found in the immediate neighborhood of the arterioles; in the dermis, near the venules. In the lung, the terminal lymphatics lie at the junctions of interalveolar septa. Simple observation of the local lymphocapillary topology (Section 8.3.4) might enable nanorobots to identify which tissue they are passing through.

The precollecting ducts are short vascular segments with valves (like all subsequent lymphatics) at intervals of 0.4-1.5 mm that connect the capillary plexuses to the larger collecting ducts. The lymphatic collecting ducts are denoted either prenodal or postnodal, though there are no morphological or structural differences between the two. Collecting ducts display many swellings and strictures along their course which alter their profile, due to the presence of valves. (Valve-counting may offer useful navigational information.) There is an inner and an outer muscular layer. The collecting ducts empty

---

**Table 8.5. Approximate Quantification of the Human Lymphatic System**

<table>
<thead>
<tr>
<th>Human Lymphatic Vessel or Lymphatic Component</th>
<th>Diameter (mm)</th>
<th>Length (mm)</th>
<th>Wall Thickness</th>
<th>Lymph Flow Rate (mm³/sec)</th>
<th>Number of Vessels, Tissues or Objects</th>
<th>Total Length (mm)</th>
<th>Total Lymph Volume (mm³)</th>
<th>Flow Velocity (µm/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cervical portals:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoracic duct</td>
<td>4.0</td>
<td>450</td>
<td>500 µm</td>
<td>21.4</td>
<td>1</td>
<td>450</td>
<td>5700</td>
<td>1700</td>
</tr>
<tr>
<td>Right lymphatic duct</td>
<td>3.0</td>
<td>14</td>
<td>500 µm</td>
<td>1.7</td>
<td>1</td>
<td>14</td>
<td>100</td>
<td>250</td>
</tr>
<tr>
<td>Main lymphatic trunks:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subclavian (left &amp; right)</td>
<td>2.5</td>
<td>50</td>
<td>~300 µm</td>
<td>0.7</td>
<td>2</td>
<td>100</td>
<td>500</td>
<td>100</td>
</tr>
<tr>
<td>Jugular (left &amp; right)</td>
<td>2.3</td>
<td>40</td>
<td>~300 µm</td>
<td>0.7</td>
<td>2</td>
<td>80</td>
<td>330</td>
<td>200</td>
</tr>
<tr>
<td>Intestinal</td>
<td>2.3</td>
<td>85</td>
<td>~300 µm</td>
<td>14.0</td>
<td>1</td>
<td>85</td>
<td>350</td>
<td>3400</td>
</tr>
<tr>
<td>Bronchomedialinal (left &amp; right)</td>
<td>2.2</td>
<td>125</td>
<td>~300 µm</td>
<td>2.0</td>
<td>2</td>
<td>250</td>
<td>950</td>
<td>500</td>
</tr>
<tr>
<td>Lumbar (right &amp; left)</td>
<td>1.9</td>
<td>100</td>
<td>~300 µm</td>
<td>0.3</td>
<td>2</td>
<td>100</td>
<td>570</td>
<td>100</td>
</tr>
<tr>
<td>Cisterna chyli (thoracic)</td>
<td>8-13</td>
<td>50-75</td>
<td>~500 µm</td>
<td>14.0</td>
<td>1</td>
<td>60</td>
<td>1700</td>
<td>500</td>
</tr>
<tr>
<td>Minor lymphatic trunks:</td>
<td>1.0</td>
<td>10</td>
<td>~100 µm</td>
<td>0.2</td>
<td>1</td>
<td>100</td>
<td>800</td>
<td>300</td>
</tr>
<tr>
<td>Post-Lymphonodal collecting ducts</td>
<td>0.5</td>
<td>~3</td>
<td>15-50 µm</td>
<td>0.01</td>
<td>1700</td>
<td>5100</td>
<td>1000</td>
<td>70</td>
</tr>
<tr>
<td>Pre-Lymphonodal collecting ducts</td>
<td>0.5</td>
<td>~3</td>
<td>15-50 µm</td>
<td>4 x 10⁴</td>
<td>617,000</td>
<td>1,850,000</td>
<td>363,000</td>
<td>2</td>
</tr>
<tr>
<td>Precollecting ducts</td>
<td>0.15</td>
<td>~1</td>
<td>~5 µm</td>
<td>3 x 10⁻²</td>
<td>68,000,000</td>
<td>68,000,000</td>
<td>1,200,000</td>
<td>0.02</td>
</tr>
<tr>
<td>Lymphatic capillaries</td>
<td>0.02</td>
<td>0.5</td>
<td>1 µm</td>
<td>3 x 10⁻²</td>
<td>6,800,000,000</td>
<td>3,400,000,000</td>
<td>425,000*</td>
<td>0.03</td>
</tr>
<tr>
<td>Lymphatics, subtotal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-lymphatic channels</td>
<td>0.1-0.2 µm</td>
<td>30 µm-30 cm</td>
<td>--</td>
<td>3 x 10⁻¹⁴</td>
<td>8 x 10⁻¹⁴</td>
<td>24 x 10⁻¹²</td>
<td>(370,000)</td>
<td>0.002</td>
</tr>
<tr>
<td>Lymphocytes</td>
<td>~10 µm</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>7 x 10⁻¹¹</td>
<td>--</td>
<td>(350,000)</td>
<td>--</td>
</tr>
<tr>
<td>Lymph nodes</td>
<td>4.0</td>
<td>--</td>
<td>--</td>
<td>0.05</td>
<td>450</td>
<td>--</td>
<td>(15,000)</td>
<td>10</td>
</tr>
<tr>
<td>Spleen</td>
<td>66</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>(75,000)</td>
<td>--</td>
</tr>
<tr>
<td>Thymus</td>
<td>22</td>
<td>--</td>
<td>--</td>
<td>1</td>
<td>--</td>
<td>--</td>
<td>(10,000)</td>
<td>--</td>
</tr>
<tr>
<td>Lymphatic vessel walls</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>(460,000)</td>
<td>--</td>
</tr>
<tr>
<td>Total Lymphatic System</td>
<td>~3,300,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Estimated for hypothetical adult male, weight 70 kg, consistent with references cited in text; * final volume after an assumed 90% fluid reabsorption.
into the lymphatic trunks, larger and thicker-walled vessels which receive the lymph drained from entire regions of the body.

By their confluence, the lymphatics form progressively larger vessels, which ultimately converge at two cervical portals (Fig. 8.6). The first is the right lymphatic duct, which is relatively short and typically opens into the right brachiocephalic vein at the junction of the right internal jugular and right subclavian veins. It carries lymph drainage from the right upper portion of the body (Fig. 8.7). The second is the thoracic duct, which arises in the abdomen and courses upward along the anterior aspect of the vertebral column (with valves every ~3 cm along its length), through the thorax into the base of the neck, where it typically joins the venous system at the junction of the left jugular and subclavian veins. At its termination, a bicuspid valve faces into the vein to prevent or reduce reflux of blood, although after death blood regurgitates freely into the thoracic duct, which then looks like a vein.

The thoracic duct is ~5 mm in diameter at its abdominal commencement, but becomes narrower at mid-thoracic levels, then in ~50% of patients grows slightly wider again before its termination. There is a tremendous variation in structure from one patient to another. The duct may divide in its mid-course into two unequal vessels which soon re-unite, or into several small branches which form a plexus before combining again to form one short wide trunk. Higher in the body, the duct occasionally bifurcates, the left branch ending as usual, the right branch diverging to join one of the right lymph trunks or even the right lymphatic duct, with the combined vessel opening into the right subclavian vein. After performing 529 human dissections, Kinnaert found that the thoracic duct terminated on the internal jugular vein in 36% of the cases, the subclavian vein...
in 17% of the cases, the jugulo-subclavian junction in 34% of the cases, both veins 8% of the time, and the transverse cervical vein in 2% of the cases where the thoracic duct terminated on the left. In 21% of the cases, the terminal openings were multiple; in the rest of the cases, the single thoracic duct terminus diameter ranged from 2-5 mm.

Various lymphatic organs are located along the length of the lymphatic system (Fig. 8.8). A major function of these organs is the production of lymphocytes which are added to the lymph passing through the vessels. Lymphatic tissue is a loose-structured material consisting of a spongelike stroma and free cells in the meshes of the stroma, and is part of the reticuloendothelial system (RES). Phagocytic cells in the sinuses serve as filters that scavenge particles from the lymph and destroy them. Such particles include red blood cells, pathogenic bacteria, and larger dust particles imported by the respiratory tract and collected by macrophage cells of the bronchial nodes. Lymphborne nanorobots can avoid these traps (Chapter 15); the Reynolds number of lymph flow is typically only ~0.0025.

The most numerous of the lymphatic organs is the lymph node, designed as a bacterial/particulate filter with channels, sinuses, valves, and fluid entering via 6-10 afferent lymph vessels (Fig. 8.9). Lymph nodes vary enormously in their size and structure, probably averaging ~4 mm in diameter but occasionally reaching 30-40 mm. A normal young adult body contains ~450 lymph nodes distributed as follows: ~30 nodes in the arm and superficial thoraco-abdominal wall down to the umbilicus; ~20 nodes in the legs and superficial buttocks, infraumbilical abdominal wall and perineum; 60-70 nodes in the head and neck; ~100 nodes in the thorax, including deep walls and contents; and ~230 nodes in the abdomen and pelvis, including deep walls and contents.

There is also the 10 cm³ of lymphoid tissue in the bilobate thymus gland, and additional lymphoid tissue in the 80-300 cm³ spleen. (The spleen slowly enlarges during alimentary digestion and varies in size with the content of blood and state of nutrition, being large in well-fed patients and small in starved patients. About 10% of the population has accessory spleens; these ~1 cm³ spleniculi may be numerous and widely scattered in the abdomen. Other lymphoid tissues include the lymph nodules or the lymph follicles in the intestines (Section 8.2.3), the subepithelial lymphoid aggregates including the palatine, lingual, and nasopharyngeal (e.g., adenoids, tonsils), and the lymphoid tissue in the submucosa of the appendix and in the bone marrow.

What, exactly, is lymph? Lymph is essentially an alkaline ultrafiltrate of the blood plasma formed by continual seepage of fluid constituents of the blood across the capillary walls and into the surrounding interstitial spaces. The lymph fluid has much the same composition as the interstitial fluid. Like blood, lymph consists of a plasmatic part and a corpuscular part (Table 8.6). Lymph contains almost no platelets but has about one-third the fibrinogen and five times the prothrombin as in blood serum, hence will thrombose spontaneously forming a yellowish clot. Lymph has 0.02-0.10% as many red cells as arterial blood, but most have passed near cells and become thoroughly deoxygenated, especially given the ~1 day transit time through the lymphatic system for cells. There is still some dissolved oxygen in the lymph water, however, roughly the same concentration as in blood plasma (see Appendix B). Glucose may be present in lymph at slightly higher concentration than in blood serum, so lymphoresident nanorobots have access to plenty of chemical fuel energy (Section 6.3.4). Lymph is slightly less viscous than blood plasma, with specific gravity 0.016-1.023.

In a fasting patient, the lymph coming from the intestine is a clear, transparent fluid. After a meal containing fat has been ingested, the intestinal lymph becomes white or milky. This is termed chyle—chyle is just lymph containing a surge (5-15% by volume) of emulsified fat which is transported in the form of chylomicrons measuring 0.5-0.75 microns in diameter. In ~65% of all patients, the intestinal trunk (which drains lymph from the stomach, intestines,
pancreas, spleen, and visceral surface of the liver) empties separately into the thoracic duct immediately above the junction of the two lumbar trunks, and so the thoracic duct widens to a diameter of 8–13 mm, forming a cisterna chyli. In the other ~35% of cases, the intestinal trunk joins the left lumbar trunk and there is no cisterna chyli (Fig. 8.10), a navigationally distinctive characteristic.

The chemical composition of lymph plasma varies significantly at different locations in the lymphatic system, reflecting changes in source of drainage and variations in capillary permeabilities. This provides another ready source of useful navigational information for lymphborne medical nanorobots (Section 8.4.3). For example, thoracic lymph is rich in histaminase originating in the kidneys and gut, while cervical duct lymph contains only very low concentrations of histaminase, as in blood plasma. In humans, fibrinogen is 4.1 mg/cm³ in blood plasma, 1.1 mg/cm³ in thoracic duct lymph; in dogs, prothrombin concentration expressed as a percentage of serum concentration was measured experimentally as 93.2% in dogs, prothrombin concentration expressed as a percentage of serum concentration was measured experimentally as 93.2% in hepatic lymph, 51.2% in thoracic lymph, and 7.6% in leg lymph. The enzyme tributyrinase has high concentration in the intestinal lymph and significantly lower concentrations elsewhere. Insulin is more plentiful in pancreatic lymph.

Cervical, leg, renal, and right duct lymph has 33%-50% of the protein concentration of blood plasma, while intestinal and thoracic duct lymph contains ~67% as much protein as the blood plasma. Protein is generally lowest in the capillary lymph, increasing in several steps with travel upstream toward the main trunks. Hepatic lymph has the highest protein content (about equal to blood plasma), as would be expected in view of the high permeability of the capillaries of the liver sinusoids. Central (trunk) lymph generally differs from blood plasma in having high concentrations of sodium, cholesterol, and tributyrinase, and significantly lower concentrations of total protein, albumin α2 fraction, potassium, nitrogen, aldolase, pyruvic transaminases, choline esterase, amylase, and diastase. The cellular composition of lymph also changes dramatically with lymphatic location. Most of the prenodal duct lymph contains ~1 x 10⁶ cells/cm³, except for hepatic lymph which has 4-6 x 10⁶ cells/cm³. The majority of these cells are red cells, but a small proportion, perhaps 10-20%, are nucleated cells. Such cells include ~85% lymphocytes, ~13% monocytes and macrophages, and ~2% neutrophilic granulocytes. The postnodal duct lymph has ~12 times more nucleated cells than the peripheral lymph, perhaps ~3 x 10⁶ nucleated cells/cm³. The central lymph has 8-11 x 10⁶ nucleated cells/cm³.

How fast does lymph move? Of the 20 liters/day of plasma water that leaves the blood circulation through ultrafiltration, 18 liters/day are reabsorbed by passing back out of the lymphatic capillaries and into the venous capillary loops; the remaining 2 liters/day returns to the circulation as lymph that passes through the entire lymphatic system. In other words, relatively little lymph reaches the collecting ducts compared with the total amount entering the lymphatic capillaries. Much more lymph is formed in the periphery than ever reaches the main lymph ducts. Basal lymph flow rates are ~2 liters/day, rising to a 6-12 liter/day rate during heavy exercise. Summing the vessel crossing times (~length/velocity from Table 8.5), it takes ~24 hours for a node-inert particle entering a lymphatic capillary to drift to the final terminus of the thoracic duct and rejoin the venous flow.

What moves the lymph? Valves are essential to the function of the lymph vascular system, which has no pump comparable to the heart. In the collecting ducts, one-way valves are spaced roughly 2-3 mm apart and can oppose retrograde pressures of up to 20 mmHg. Lymph flow in the extremities depends largely on the massaging effect of movement of the surrounding tissues that results from muscle contraction. Such movements can be produced indirectly by the motion of nearby arterioles, or by the contraction of the skeletal muscles, or by movement of the organs (e.g., motion of the arms or legs, peristalsis of the intestine, or breathing of the lung). Arterial pulsations contribute 1.5–2.2 mmHg pressure variation to the lymphatic ducts between systole and diastole. Larger collecting lymphatics are innervated, so movement can also be induced by contractions of smooth muscles in the lymphatic vessel.
The contractile activity seems to depend on the volume of lymph produced. If this is small, the vessel walls are quiescent, but slight distension initiates rhythmic contractions at ~0.1-0.2 Hz and can raise the lymphatic pressure by up to ~5 mmHg. By comparison, the mean interstitial hydrostatic pressure is ~1.4 mmHg and the mean intralymphatic pressure is ~0.9 mmHg.

The minimum topological map required to reliably navigate the entire ~3500 kilometers of human lymphatic vasculature should be roughly comparable to that required for a comprehensive blood capillary map, or ~1 terabit (Section 8.2.1.2).

### 8.2.2 Navigational Bronchography

Medical nanorobots may access the human body via the respiratory system. The airway begins at the mouth and nose, extends through the pharynx, larynx, the trachea and bronchial branchings, and ends at the alveoli. The conducting zone from the top of the trachea to the beginning of the respiratory bronchioles contains no alveoli, so gas exchange with the blood does not occur there. Gas exchange occurs only in the respiratory zone, which extends from the respiratory bronchioles down to the alveolar sacs.

Weibel provided the first quantitative measurements of the length, diameter, and area of successive segments of the human airway. Each branch gives rise to two narrower daughter branches which may vary considerably in length. However, the summary in Table 8.7 assumes mean values of length and diameter. Generations 0-16 are the conducting airways, while generations 17-23 constitute the respiratory airways which have alveoli on their walls. Generation 23 terminates in alveoli.

The respiratory airflow begins in the mouth and nose. In the nose are two nasal cavities (totalling ~160 cm² in area), one of which is illustrated in Figure 8.11. The nasal cavities are divided by a partition called the nasal septum, from which diverge three winglike projections called the conchae. The endings of the olfactory nerve lie in the mucosa in the cilia-free region near the ~1 cm² olfactory bulb above the superior concha. The nasal passages are flanked by four sinuses which may swell up during infection or inflammation, closing off the air passages and necessitating breathing through the mouth. Lacrimal ducts drain tears and other secretions from the eyes into the nose via the nasolacrimal duct, requiring noseblowing after crying. The human vomeronasal organ, located near the base of the nasal septum in adults, transduces pheromonal signals, although apparently this organ is not present, is inactive, or is very insensitive in some people.

Air passing through the nasal cavities is warmed to within 1 K of body temperature by an extensive capillary vasculature, and is humidified by nasal mucous glands (e.g., goblet cells which secrete mucoid fluid) to within ~1% of full saturation, before the air enters the pharynx. Coarse particles are removed from incoming air by nose hairs. Smaller particles are removed by turbulent precipitation, wherein obstacles in the nasal passages force the airflow to execute many sharp turns which the particles are too heavy to negotiate. The particles hit the nasal mucous membrane and become embedded in the mucus. The nasal turbulence mechanism is so effective that

### Table 8.7. Approximate Quantification of the Human Bronchial System

<table>
<thead>
<tr>
<th>Pulmonary Branch</th>
<th>Generation</th>
<th>Number</th>
<th>Branch Diameter (mm)</th>
<th>Branch Length (mm)</th>
<th>Cumul. Length (mm)</th>
<th>X-Section Area (cm²)</th>
<th>Volume (cm³)</th>
<th>Cumul. Volume (cm³)</th>
<th>Air Speed (cm/sec)</th>
<th>Reynolds Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trachea</td>
<td>0</td>
<td>1</td>
<td>18</td>
<td>120.0</td>
<td>120</td>
<td>2.6</td>
<td>31</td>
<td>31</td>
<td>393</td>
<td>4350</td>
</tr>
<tr>
<td>Main bronchus</td>
<td>1</td>
<td>2</td>
<td>12.2</td>
<td>47.6</td>
<td>167</td>
<td>2.3</td>
<td>11</td>
<td>42</td>
<td>427</td>
<td>3210</td>
</tr>
<tr>
<td>Lobar bronchus</td>
<td>2</td>
<td>4</td>
<td>8.3</td>
<td>19.0</td>
<td>186</td>
<td>2.2</td>
<td>4</td>
<td>46</td>
<td>462</td>
<td>2390</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>5.6</td>
<td>7.6</td>
<td>194</td>
<td>2.0</td>
<td>2</td>
<td>47</td>
<td>507</td>
<td>1720</td>
<td></td>
</tr>
<tr>
<td>Segmental bronchus</td>
<td>4</td>
<td>16</td>
<td>4.5</td>
<td>12.7</td>
<td>206</td>
<td>2.6</td>
<td>3</td>
<td>51</td>
<td>392</td>
<td>1110</td>
</tr>
<tr>
<td>Bronchi w/cartilage</td>
<td>5</td>
<td>32</td>
<td>3.5</td>
<td>10.7</td>
<td>217</td>
<td>3.1</td>
<td>3</td>
<td>54</td>
<td>325</td>
<td>690</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>6</td>
<td>64</td>
<td>2.8</td>
<td>9.0</td>
<td>226</td>
<td>4.0</td>
<td>4</td>
<td>57</td>
<td>254</td>
<td>434</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>7</td>
<td>128</td>
<td>2.3</td>
<td>7.6</td>
<td>234</td>
<td>5.1</td>
<td>4</td>
<td>61</td>
<td>188</td>
<td>277</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>8</td>
<td>256</td>
<td>1.86</td>
<td>6.4</td>
<td>240</td>
<td>7.0</td>
<td>4</td>
<td>66</td>
<td>144</td>
<td>164</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>9</td>
<td>512</td>
<td>1.54</td>
<td>5.4</td>
<td>246</td>
<td>9.6</td>
<td>5</td>
<td>71</td>
<td>105</td>
<td>99</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>10</td>
<td>1020</td>
<td>1.30</td>
<td>4.6</td>
<td>250</td>
<td>13.6</td>
<td>6</td>
<td>77</td>
<td>73.6</td>
<td>60</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>11</td>
<td>2050</td>
<td>1.09</td>
<td>3.9</td>
<td>254</td>
<td>19.7</td>
<td>7</td>
<td>85</td>
<td>52.3</td>
<td>34</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>12</td>
<td>4100</td>
<td>0.95</td>
<td>3.3</td>
<td>257</td>
<td>29.0</td>
<td>10</td>
<td>95</td>
<td>34.4</td>
<td>20</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>13</td>
<td>8190</td>
<td>0.82</td>
<td>2.7</td>
<td>260</td>
<td>44.2</td>
<td>12</td>
<td>106</td>
<td>23.1</td>
<td>11</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>14</td>
<td>16,400</td>
<td>0.74</td>
<td>2.3</td>
<td>262</td>
<td>70.0</td>
<td>16</td>
<td>123</td>
<td>14.1</td>
<td>6.5</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>15</td>
<td>32,800</td>
<td>0.66</td>
<td>2.0</td>
<td>264</td>
<td>113.3</td>
<td>22</td>
<td>145</td>
<td>8.92</td>
<td>3.6</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>16</td>
<td>65,500</td>
<td>0.60</td>
<td>1.65</td>
<td>266</td>
<td>180.0</td>
<td>30</td>
<td>175</td>
<td>5.40</td>
<td>2.0</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>17</td>
<td>131 x 10³</td>
<td>0.60</td>
<td>1.41</td>
<td>267</td>
<td>300.0</td>
<td>42</td>
<td>217</td>
<td>3.33</td>
<td>1.1</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>18</td>
<td>262 x 10³</td>
<td>0.54</td>
<td>1.17</td>
<td>269</td>
<td>534.3</td>
<td>61</td>
<td>278</td>
<td>1.94</td>
<td>0.57</td>
</tr>
<tr>
<td>Bronchi w/cartilage in wall</td>
<td>19</td>
<td>524 x 10³</td>
<td>0.47</td>
<td>0.97</td>
<td>270</td>
<td>944.0</td>
<td>93</td>
<td>370</td>
<td>1.10</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Average of five normal human lungs; includes both lungs; air speed and Reynolds number at 1 liter/sec flow.
almost no nasally-inhaled particles larger than 2-5 microns reach the lower airway.

Most of the surfaces of the nasal airway are covered by a layer of ciliated, pseudostratified columnar epithelium cells. Each epithelial surface cell has 25-100 cilia that beat forcefully and continually toward the pharynx at ~10 Hz. Respiratory cilia are ~0.2 microns in diameter and ~2-5 microns long (having a ~31 nm apical glycocalyx), with a mean separation between cilia of ~2-5 microns. The film of particle-carrying nasal mucus is 200 microns thick and is moved toward the pharynx by the cilia at a speed of ~1-3 cm/min, where it is then swallowed down the esophagus. Mucus velocity increases with luminal depth. In unciliated areas such as the front of the nose, where temperatures fall below levels the cilia can tolerate, the mucous layer creeps along the surface solely through traction from neighboring ciliated areas. Swallowed mucus volume totals ~0.1 cm³/min. The absolute viscosity of normal mucus is typically ~1 kg/m-sec, rising as high as ~500 kg/m-sec in the thick sputum of cystic fibrosis patients. Mucus rheology and mucociliary clearance mechanisms of the respiratory tract have been widely studied, along with the energy dissipation per cilium in the periciliary fluid. Airflow contributes little to mucus transport during normal breathing, except in patients with bronchial hypersecretions, although high-frequency ventilation and coughing may make significant contributions.

During inspiration, air passes through the nose or mouth into the pharynx (throat). The posterior wall of the pharynx rests against the cervical vertebrae; the lateral wall has openings communicating with the middle ear (auditory or eustachian tube). The pharynx branches into two tubes—the esophagus (through which food passes to the stomach; Section 8.2.3) and the larynx (part of the airways). The larynx houses the vocal cords, two strong bands of elastic tissue (covered by stratified scalelike epithelium) stretched horizontally across its lumen. The flow of air past the vocal cords causes them to vibrate, producing sounds. The ventricular folds or false vocal cords (in mid-glottis) point downward. When closed by the action of sphencter muscles, the folds form an exit valve that permits the building up of abdominal pressure as in straining at stool or in expulsion of the fetus. Coughing involves releasing the air explosively through the folds.

After passing the larynx, air enters the trachea, a cylindrical tube with 16-20 circumferential cartilaginous rings shaped like horseshoes and embedded in an external fibroelastic membrane (Fig. 8.12). The trachea branches into two main bronchi, one of which enters each lung. The right bronchus is shorter, wider, and more nearly vertical in direction than the left bronchus; both (and subsequent branchings) are supported by complete rings of cartilage to prevent collapse under high levels of suction. There are more than 20 generations of branchings in the lungs, each resulting in narrower, shorter, and more numerous tubes (Table 8.7). When the bronchi become smaller than ~1 mm in diameter, they lose their cartilage, and become bronchioles.

The lungs themselves are cone-shaped organs which lie in the pleural cavities of the thorax. The base of each lung contacts with the upper surface of the diaphragm, extending to the level of the 7th rib anteriorly and the 11th rib posteriorly. The right and left pleural cavities are formed by two serous sacs into which the lungs are invaginated. Two layers of pleura are separated by a thin layer of fluid from ~20-80 microns thick. Pleural fluid is formed on the parietal pleural surface at a rate of 7-11 cm³/hr, with up to 20-25 cm³ normally present in the pleural space; glucose is present at serum levels (Appendix B) and there are topographic differences in pleural pressure. The right lung, larger in size than the left, is divided into three lobes; the left lung has only two lobes (Fig. 8.12), presumably to make room for the heart. The interlobar surfaces are covered with visceral pleura where the lobes contact one another and are lubricated with the same thin mucoid pleural fluid that lubricates the outer surface of the lungs. The lobes slide against each other in the same way that the entire lungs slide within the thoracic cavity. The diaphragm is the principle respiratory muscle. Contraction of the diaphragm elongates the lungs, forcing them to inflate. Other muscles elevate the anterior thoracic wall or compress the abdomen, raising the ribs from an inferiorly slanting position to a horizontal position that increases the anteroposterior diameter of the chest.

Returning to the airflow, bronchiole walls are composed of smooth muscle and connective tissue. This smooth muscle normally
remains relaxed so that the bronchioles stay open, although particles of irritating substances entering these passageways can cause bronchiolar spasms (as in patients with asthma). The bronchioles branch several times. The last bronchiolar branch that still has a complete muscular coat is the terminal bronchiole. This too branches into several respiratory bronchioles with sparse smooth muscle. These bronchioles are the entryways into the respiratory lobules, the final air spaces of the lungs (Fig. 8.13). The walls of all parts of the respiratory lobules form the respiratory membrane which totals ~70 m² in total surface area in both lungs combined. The membrane is so thin (≤1 micron) that oxygen and carbon dioxide can diffuse freely between the air inside the lobule and the blood in the capillary surrounding the lobule.

As in the nasal passages, the bronchial tree from the lower pharynx down to the end of the respiratory bronchioles is lined with ciliated columnar epithelial cells. These cilia also wave constantly toward the pharynx, moving secreted mucus at ~1.4 cm/min which replaces the entire mucoid coating once every ~20 minutes. A second protective mechanism is provided by mobile phagocytes present in the airways and in the alveoli, that engulf inhaled particles and bacteria (≤100/μm² in a healthy person) and thus prevent this foreign matter from gaining access to other lung cells or from entering the blood by conveying this matter into the lymph system. Ciliary activity may be inhibited for several hours by smoking a single cigarette. Phagocytes are also injured by cigarette smoke, air pollution, and other noxious agents. Below the respiratory bronchioles, the ciliated columnar epithelium gives way to a nonciliated cuboidal epithelium.

Alveoli first begin to appear in the respiratory bronchioles, attached to the walls, and their frequency increases in the alveolar ducts until the airways end in grapelike clusters of alveoli (Fig. 8.13). The alveoli are tiny hollow sacs 100-300 microns in diameter that open onto the lumina of the airways (Fig. 8.14). Typically the air in two alveoli is separated by a single wall. The moist air-facing surfaces of the alveolar wall are lined by a continuous layer, one cell thick, of squamous type I epithelial cells (Fig. 8.15, enlargement of box in Figure 8.14). The alveolar surface contains smaller numbers of thicker specialized type II epithelial cells that secrete a detergent-like substance, or surfactant (a complex of protein with dipalmityl lecithin) in a fluid layer ~70 nm thick, which physically stabilizes alveoli of different sizes during inflation and deflation. (See also Section 9.2.3.)

The alveolar walls also contain capillaries, the endothelial linings of which are separated from the alveolar epithelial lining only by a basement membrane and a very thin interstitial space containing interstitial fluid and a loose meshwork of connective tissue (Fig. 8.15). In places, the interstitium is absent and the alveolar epithelium fuses with the capillary endothelium, with total thickness ≤ 1 micron. In some of the alveolar walls there are pores that permit the flow of air between alveoli, an important route when the airway leading to an alveolus is occluded by disease.
Table 8.8. Approximate Quantification of the Human Alimentary System

<table>
<thead>
<tr>
<th>Alimentary System Component</th>
<th>Length (cm)</th>
<th>External Diameter or Width (cm)</th>
<th>Internal Volume (cm³)</th>
<th>Luminal Cylinder (cm³)</th>
<th>Contents Passage Time</th>
<th>Velocity of Contents (cm/sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouth &amp; Pharynx</td>
<td>8</td>
<td>2-5</td>
<td>50</td>
<td>80</td>
<td>1-10 sec</td>
<td>1-8</td>
</tr>
<tr>
<td>Esophagus</td>
<td>25</td>
<td>1.3-2.5</td>
<td>100</td>
<td>200</td>
<td>5-20 sec</td>
<td>3-5</td>
</tr>
<tr>
<td>Stomach</td>
<td>12</td>
<td>8</td>
<td>230-1000</td>
<td>600</td>
<td>2-6 hrs</td>
<td>~0.001</td>
</tr>
<tr>
<td>Pyloric sphincter</td>
<td>0.3</td>
<td>1.3-2.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Intestine</td>
<td>400</td>
<td>3-6</td>
<td>1100</td>
<td>3500</td>
<td>3-5 hrs</td>
<td>0.03</td>
</tr>
<tr>
<td>Duodenum</td>
<td>25</td>
<td>3.8-6.0</td>
<td>600</td>
<td>2000</td>
<td>10-20 hrs</td>
<td>0.004-0.008</td>
</tr>
<tr>
<td>Jejunum</td>
<td>160</td>
<td>3.2-3.8</td>
<td>3500</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ileum</td>
<td>215</td>
<td>2.5-3.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large Intestine</td>
<td>~150</td>
<td>5.0-7.5</td>
<td>300</td>
<td>2000</td>
<td>10-20 hrs</td>
<td>0.004-0.008</td>
</tr>
<tr>
<td>Cecum</td>
<td>6</td>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Appendix</td>
<td>8-9</td>
<td>~0.8</td>
<td></td>
<td>0.07-0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ascending colon</td>
<td>13-20</td>
<td>~7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transverse colon</td>
<td>50</td>
<td>~6.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Descending colon</td>
<td>23-30</td>
<td>~6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sigmoid colon</td>
<td>41</td>
<td>~5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td>16-20</td>
<td>2.5-3.8</td>
<td>40</td>
<td>100</td>
<td>~1 hr</td>
<td>0.006</td>
</tr>
<tr>
<td>Rectal canal</td>
<td>13-16</td>
<td>3.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anal canal</td>
<td>3-4</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total, Average, or Range:</td>
<td>~600</td>
<td>~3.5</td>
<td>1800-2600</td>
<td>6500</td>
<td>16-32 hrs</td>
<td>~0.01</td>
</tr>
</tbody>
</table>

estimated for 70 kg adult male.

At rest, ~4 liters/min of air enter and leave the alveoli, while 5.4 liters/min of blood, the entire cardiac output, flows through pulmonary capillaries which total ~2400 kilometers in length and ~150 cm³ in volume. During heavy exercise, the air flow to the alveoli can increase to 120-160 liters/min and the blood flow to 25-30 liters/min.866 The volume of air taken in with each breath (resting tidal volume) is ~0.4 liter/breath at a respiration rate of 18/min (12-15/min during sleep). Volume flow may rise to ~3.2 liters/breath at a respiration rate of up to 62/min during the most strenous exercise.780 The two human lungs hold a combined ~6 liters of gas, of which ~3.7 liters is the maximum inspirational capacity leaving ~2.3 liters as residual capacity or dead space.

Blood pressure in the pulmonary artery is only 15-30 mmHg systolic and 4-12 mmHg diastolic, compared to 100-150 mmHg systolic and 60-100 mmHg diastolic in the aorta.661 This low-pressure system allows pulmonary capillaries to be quite flexible and thin-walled—they distend if higher blood pressures are applied to the pulmonary artery. Pulmonary blood vessels are richly supplied with nerve fibers but are largely free from neural and chemical control—although they do respond to hypoxia and to pharmacological doses of catecholamines, histamine and serotonin.661

A navigational map of the airways can be surprisingly compact. Only log₂(300 x 10⁶) ~ 28 bits are needed to uniquely name each alveolus, requiring 8.4 x 10⁸ bits to store all 300 million addresses. A complete bifurcation map of the airways down to the last generation of respiratory bronchioles (e.g., a lobule map for both lungs) can be stored in just ~10⁸ bits. If numerous positions within each alveolus must be specified, for example during a scan seeking pre-cancerous epithelial cells, the required map could be much larger. However, C. Phoenix notes that if nanorobots are given individual instructions to scan a certain set of cells for cancer, a contiguous (nonrandom) tissue volume could be searched, specified by a small number of partial addresses. For instance, if 10⁷ nanorobots were tasked to scan 30 (unnamed) alveoli each, within the contiguous volume, then each nanorobot must store one ~23 bit hardcoded kernel address plus one ~5 bit alveolus extension address (total 28 bits/alveolus) and one 30-bit cell address for each cancerous cell discovered, a modest memory requirement. On the other hand, general surveyor nanorobots lacking individualized instructions can keep track of the forks they pass and the branches they take as they enter the lungs, assembling the final address (of any cancer cell they ultimately discover) as they go. Upon reaching the terminus of the bronchial system and encountering another nanorobot already at work, the newcomer moves on to another location, keeping track of its current address (~28 bits). Once a cancer is detected, treatment nanorobots can follow the surveyor’s address to return to the specific cancer cell without requiring a comprehensive map of the rest of the lung. Thus a treatment protocol can be executed without any overall map of the lung, although the assignment of individual search territories may prove more efficient, and multiple connectivities within the bifurcation tree may demand storage of additional bits or require post-survey data compression.

8.2.3 Navigational Alimentography

Medical nanorobots may also access the human body via the alimentary canal—a ~6 meter long musculomembranous digestive tube which begins at the mouth and ends at the anus. With its accessory glands and organs, the alimentary system controls the intake, mastication, digestion, absorption and elimination of foods.

An approximate quantification of the gastrointestinal system is given in Table 8.8. (Passage times are for normal digestion and exclude pathological conditions such as constipation or diarrhea.) The general
plan as shown in Figure 8.16 is topologically very simple. The alimentary canal is formed by the mouth, pharynx, esophagus, stomach, small intestine, large intestine, and rectum. Associated with the canal are accessory structures and glands including the salivary glands, liver, gallbladder and pancreas. The tube itself has four layers in radial aspect. Starting at the luminal surface these include the mucosal (superficial epithelium and cells secreting mucus and digestive juices), submucosal (connective tissue with blood and lymphatic vessels), muscular (smooth muscle in circular and longitudinal layers) and serous (parietal and visceral peritoneum) layers. The entire tract, glands, and muscle are innervated from the involuntary or autonomic system and are also influenced by hormones, some from the tract itself and some secreted by glands elsewhere in the body.

It is instructive to imagine the "view" from a medical nanorobot as the nanodevice navigates the entire length of the alimentary system. The journey begins in the mouth (the sphincter oris) where food (typical viscosity 0.01-100 kg/m-sec) is chewed and mixed with saliva secreted by one of the three pairs of major salivary glands (Fig. 8.17). (There are also the minor salivaries consisting of the labial, buccal, molar, lingual, and palatine glands.) After consulting its map, the nanorobot first explores a duct (~1 mm wide, ~25 mm long) in the anteroinferior dental arch leading into the walnut-sized submaxillary salivary gland situated mostly under the lower jawbone. The gland is organized as several thousand separate small secretory units called acini. Each acinus with its microduct forms a single parenchymal unit called a salivon. The lumens of salivons flow together to form progressively larger ducts (all surrounded by contractile myoepithelial cells and a basal lamina), finally reaching the main salivary duct that empties into the mouth. Passing under the tongue, our mobile nanodevice traverses multiple ducts (~0.5 mm wide, ~5 mm long) leading from the sublingual, the smallest gland located in the floor of the mouth. After climbing up to the soft palate in the roof of the mouth, the nanorobot reaches a pair of ducts (~1 mm wide, 15-20 mm long) leading from the parotid glands—the largest of the three glands, lying anteroinferior to each ear.

The six major salivary glands together typically produce 500-1500 cm^3/day of saliva in response to mechanical, thermal, and chemical stimuli applied to the mucous membrane of the mouth, or as the result of psychological or olfactory stimuli. Unstimulated flow is 10-50 cm^3/hr during waking hours and 1-2 cm^3/hr during sleep, but may rise as high as 290 cm^3/hr during eating. Salivary fluid specific gravity is 1.01-1.02, mean viscosity is ~ 4 x 10^{-3} kg/m-sec but up to 7.4 x 10^{-3} kg/m-sec in bulimic patients, and pH averages 6.8 (range 5.6-7.6). Plenty of chemical energy is available to power a nanorobot exploring these ducts—including 1.96 (1.13-2.81) x 10^{-4} gm/cm^3 of glucose, 7.5 (2.5-9.0) x 10^{-5} gm/cm^3 of cholesterol, and 2-4 x 10^{-3} gm/cm^3 of protein for stimulated secretions, or 1 x 10^{-3} gm/cm^3 of glucose and 8 x 10^{-3} gm/cm^3 of cholesterol for unstimulated secretions. Saliva is 99.4% (99.1%-99.6%) water, but duct secretions vary enough in composition along the length of each duct and between glands to be useful in crude chemonavigation. The salivary microbial population is ~ 10^6/cm^3. For example, the submaxillary glands secrete 2.7 (0.8-6.0) x 10^{-5} gm/cm^3 of mucopolysaccharide (mucin) derived from mucous cells within the glands, plus a small amount of the important digestive enzyme ptyalin (salivary amylase) from serous cells. By comparison, the parotid glands are purely serous glands that secrete no mucin, while the sublingual glands secrete mostly mucus. The salivary microbial population is ~ 10^6/cm^3. After food has been chewed, moistened, and reduced to a semi-liquid state, the tongue rolls it into a 1-10 cm^3 bolus and pushes it backward into the pharynx. During the act of deglutition (swallowing), the uvula is pushed backward against the posterior pharyngeal wall, closing the nasal passages; the larynx elevates and the epiglottis folds back, forming a protective lead...
that covers the windpipe. The esophagus opens to receive the food, and contracting muscular waves convey the bolus to the stomach. These waves are strong enough to oppose at least several g's of negative vertical acceleration; esophageal manometry shows normal contractions of up to -190 mmHg in the distal esophagus.\textsuperscript{122} With each swallow, 2-5 cm\textsuperscript{3} of air are ingested.\textsuperscript{2180}

Like the mouth and pharynx, the esophagus is lined with mucus-coated stratified squamous epithelium which is part of the mucous membrane. Figure 8.18 illustrates a cross-section of the esophagus, showing a number of longitudinal pleats to allow radial stretching. Note that when no food is present, the tube is almost completely squeezed shut. The tube readily distends to a width of 1-2 cm to permit the passage of a bolus of food or drink. The movement of food from the bottom of the esophagus into the stomach is controlled by the opening and closing of a muscular ring called the lower esophageal or cardiac sphincter. At the approach of the bolus, the cardiac sphincter relaxes, the opening widens, and the food is shot into the stomach. Normal sphincter pressure is 10-50 mmHg.\textsuperscript{2122,2180}

The stomach is an irregularly pear-shaped bag located at the level of the lowermost ribs with its upper end just below the heart. It has a central portion (the body), a balloonlike portion to the left (the fundus), and a constricted portion to the lower right (the pylorus, or lower quarter of the organ). The upper part of the stomach is usually puffed out a bit with gas, even when the stomach is empty, so it protrudes slightly above the cardiac sphincter. However, when the stomach is filled with food in a man standing erect, the organ assumes an almost vertical position with a tubular shape.

When greatly stretched, the stomach of a normal male can store as much as 1000 cm\textsuperscript{3} of food and fluid, up to 1500 cm\textsuperscript{3} for very large people but as little as 60 cm\textsuperscript{3} in newborns.\textsuperscript{863,870} However, 500 cm\textsuperscript{3} usually gives a person a sense of fullness.\textsuperscript{863} The inner wall of the empty stomach has longitudinal expansion pleats called rugae. As the stomach fills with food, the rugae flatten out and disappear, leaving a "600 cm\textsuperscript{2} smooth mucous membrane surface when the organ is full.

The uppermost epithelial layer of the stomach lining is the mucosa, several millimeters thick (Fig. 8.19). Almost all of the epithelial cells that line the surface are simple columnar mucous cells that secrete mucus. The gastric mucus is especially viscous, \textasciitilde50-100 microns thick,\textsuperscript{3165} and is highly resistant to both the digestive juices and the acid secreted by the stomach itself. Absorption through the almost impermeable stomach wall is slight, although the stomach does absorb small amounts of water, electrolytes, certain drugs such as aspirin, and alcohol.

The mucosa contains the secreting cells of the stomach, arranged in small tubular units to form the gastric glands. Within these glands, the chief or zymogenic cells are simple columnar cells forming a continuous lining for the tubule, and secrete important protein-digesting enzymes such as pepsin and rennin (chymosin), and lipid-digesting enzymes such as lipase. The parietal cells are scattered along the tubule and secrete ~0.5% hydrochloric acid solution.\textsuperscript{49} Both the movements of the stomach and the flow of gastric juice may be stimulated by the nerve plexuses and by hormones such as gastrin and histamine.\textsuperscript{751}

Each day ~35 million gastric glands secrete 1000-3000 cm\textsuperscript{3} of gastric juice;\textsuperscript{1975} a residuum of 50 cm\textsuperscript{3} is always present in the stomach, even after lengthy fasting.\textsuperscript{49} Secretion rates in young adults average 77 cm\textsuperscript{3}/hr (male) and 70 cm\textsuperscript{3}/hr (female) while fasting, 54 cm\textsuperscript{3}/hr (male) and 38 cm\textsuperscript{3}/hr (female) while sleeping, and 114 cm\textsuperscript{3}/hr (male) and 99 cm\textsuperscript{3}/hr (female) after eating.\textsuperscript{385} (The gender distinction is at least partly due to differences in mean body size.) Gastric juice is 99% water with specific gravity ~1.006 (1.004-1.010) and pH ~2.0.\textsuperscript{585,575} Glucose content of gastric juice is 0.33-1.19 x 10\textsuperscript{-3} gm/cm\textsuperscript{3}.\textsuperscript{585}

If the stomach has been empty for some time, it contracts and produces uncomfortable sensations (hunger pangs). Soon after food is taken, the pangs cease and gastric peristalsis begins as a ring of constriction around the middle of the stomach. This ring progresses toward the pylorus, growing measurably deeper as it moves. These mixing waves recur at ~3/minute, each taking ~1 minute to travel from its origin to the pylorus, so there are usually 3-4 moving waves of contraction simultaneously present on the human stomach. This kneading of the food can produce a gurgling sound because of the gas that is usually trapped in the stomach.

The well-churned food mixture, now called chyme, is ejected through the pyloric valve into the duodenum of the small intestine. Nervous and hormonal signals (e.g., enterogastrone) arising mainly from the duodenum, but also partly from the stomach, control the degree of contraction of the pyloric sphincter and thereby control the rate at which the chyme is emptied from the stomach into the duodenum of the small intestine. About 8-10 cm downstream from the pyloric sphincter, the cruising nanorobot pauses to investigate a small elevation called the duodenal papilla. The orifice at the summit of the papilla is surrounded by muscle fibers which form the sphincter of Oddi.\textsuperscript{585,575}
beyond which lies the 6-mm-wide ampulla of Vater and the common bile duct, a tube ~3 mm wide and ~60 mm in length that carries various glandular secretions directly into the duodenum. About 10 mm up the ampulla of Vater the pancreatic duct (duct of Wirsung) veers to the left side of the body and runs ~15 mm to the pancreas, where it continues on for the length of the organ, another ~90 mm. In some patients, the narrower accessory pancreatic duct (duct of Santorini, draining the head of the pancreas) runs ~15 mm from the pancreas and empties into the duodenum ~25 mm upstream from the duodenal papilla. Through the ~2 mm wide pancreatic duct passes ~1000-1500 cm³/day of pancreatic juice.583 1975 This juice contains pancreatin, a mixture of the three digestive enzymes trypsin (which digests protein), lipase (which digests fat), and amylase (which digests starch). The specific gravity of the fluid is 1.008, mean viscosity is $1.61 \times 10^{-3}$ kg/m-sec (up to $5.8 \times 10^{-3}$ kg/m-sec in patients with chronic pancreatitis315), the pH is 7-8, and the glucose content is $0.85-1.8 \times 10^{-3}$ gm/cm³.585 Juice flows on signal from the hormone secretin, manufactured by the mucous membrane of the duodenum which sends its message as soon as partially digested food enters from the stomach.

Moving ~55 mm up the common bile duct from the ampulla of Vater, the nanorobot passes through the sphincter of Oddi and next encounters the cystic duct, a ~2 mm wide tube ~35 mm in length branching to the right that leads through the spiral valve directly to the gallbladder. The branch to the left (of the body) continues as the ~25 mm long common hepatic duct, which then bifurcates to form the left and right hepatic ducts (~2 mm wide) leading directly to the liver (Section 8.2.5), each duct up to ~50 mm in length. Some patients also have a few direct connections between the liver and the cystic duct, called the ducts of Luschka.935

The gallbladder is a 7-10 cm long pear-shaped bag composed of muscle and membrane lodged in a hollow on the underside of the right lobe of the liver behind the lower ribs, that serves as the reservoir for bile that is secreted continuously by the liver. The cystic duct carries bile both to and from the gallbladder, depending upon whether the sphincter of Oddi is open or closed. The liver secretes 800-1000 cm³ of bile each day, but the bile duct delivers only ~500 cm³/day of bile to the duodenum.853 870 The excess is shunted to the gallbladder, which holds up to ~50 cm³ of tenfold bile concentrate. About 90% of all digestion and absorption takes place in the small intestine, including up to 6 liters/day of the 8-10 liters/day of water that flows into it from swallowed saliva, ingested water, the acid fluid secreted by the stomach, bile and pancreatic juice, as well as fluid secreted by the upper small bowel itself. Food is passed along by muscular contractions in waves known collectively as peristalsis, with waves progressing arhythmically for distances varying from 10-100 cm in length, and occasionally over the entire length of the small intestine. Food is also broken up by rhythmic segmentation contractions within the irregular peristaltic motions, which are ringlike contractions of the circular muscle ranging in frequency from 10-30/minute, with higher rates at the upstream end of the bowel.

The small intestine is a continuous tube with three well-defined sections—the duodenum, jejunum, and ileum (Table 8.8). The total length is commonly reported as ~7 meters, but this measurement is for tissues taken from cadavers which have lost all muscle tone. In the living body, the small intestine is only 3-5 meters in length.863 In the duodenum and jejunum the submucosa elevates into a series of permanent transverse pleats called circular folds, plicae circulares, or Kerkring's folds (Fig. 8.20). These ridges do not disappear when the wall is distended by the passage of food, and may stand up to ~10 mm high in the mucosa. Some folds wrap all the way around the intestine, while others extend only partly around. Starting a short distance past the pylorus, the circular folds are numerous and high in the distal portion of the duodenum and in the proximal portion of the jejunum, after which they gradually become less numerous and smaller; by mid-ileum, they fade out completely. Thus the number and depth of the folds may be used as crude navigational surface markers of range from the pylorus. The folds increase the local absorptive area by about threefold and further enhance absorption by causing the chyme to spiral as it passes downstream,853 crudely analogous to a spinning bullet passing down a rifled bore. A nanorobot traversing the folded intestinal surface must travel at least ~3 times farther than another nanorobot pursuing a more linear axial course through the luminal (chymous) contents of the tube.

The duodenum is the first short part of the small intestine, arranged in a horseshoe shape enclosing the head of the pancreas. The acidic chyme is neutralized following exposure to the alkaline pancreatic juices and bile fluid. Hence the pH of the chyme

![Fig. 8.20. Layers of the small intestine (redrawn from Guyton863).](image-url)
increases along the length of the duodenum, a fact which may be useful for chemonavigation (Section 8.4.3). The duodenal glands (Brunner’s glands) are found only in the duodenum. They are tortuous and branching; their mucus-containing secretion has a pH of 5.8-7.6, a specific gravity of 1.01, and a highly variable cholesterol concentration of 3.61 (0-31.5) x 10^4 gm/cm^3. Total unstimulated secretory volume is ~30 cm^3/hr, rising to 181 cm^3/hr (male) and 126 cm^3/hr (female) with secretin stimulation.

In the jejunum, fats, starches, and proteins are broken down to their smallest components (or small peptides) and are absorbed by the lining cells of the bowel. Of particular interest, absorption of sugars takes place chiefly in the upstream portion of the small intestine, specifically in the duodenum and upper jejunum. Hence the concentration of glucose in the chyme peaks and then sharply declines in the jejunum because starchy molecules are enzymatically reduced to the simplest sugars there, prior to absorption, although the disaccharides are not as readily absorbed. Cholesterol is also absorbed mainly in the jejunum.

In the ileum, water is absorbed (~0.07-0.40 cm^3/sec) along with calcium, other minerals, and vitamins (especially vitamin B12). Bile is recaptured and returned to the liver via the hepatic portal vein and the lymphatic thoracic duct systems. Fat is also absorbed more rapidly in the ileum than in the duodenum or jejunum. These spatially differential absorptive properties of the small intestine produce numerous luminal chemical gradients that may be useful for both axial and radial chemonavigation by nanorobots traveling with the chyme flow.

Four additional somatographic features will be encountered by nanorobots crawling across the surface of the small intestine. First, lymphoid tissue is present in the form of nodules or follicles throughout the luminal wall. These may appear as solitary nodules that may extend into the submucosa all over the intestine, averaging 0.4-2 mm in diameter, up to 3-5 nodules/cm^2 of mucous membrane, ~2400-4500 per patient. The nodules become larger and more numerous as the nanorobot travels downstream. Aggregated nodules are found chiefly in the lower ileum, penetrating the mucosa, some as solitary nodules that may extend into the submucosa all over the intestine, averaging 0.4-2 mm in diameter, up to 3-5 nodules/cm^2 of mucous membrane, ~2400-4500 per patient. These oval-shaped aggregates, called Peyer’s patches, are slightly elevated areas 12-38 mm in length and 8-25 mm wide, oriented with the long axis pointing downstream.

Each patch is composed of 8-60 nodules in varying degrees of fusion. There are 20-30 patches per patient.

Second, the mucosal surface is pockmarked with ~100 million pit-like structures ~50 microns in diameter called intestinal glands or crypts of Lieberkühn (Fig. 8.21). Their depth ranges from 100 microns in the mucosa near the headwaters of the small intestine to 700 microns far downstream in the large intestine. The glands collectively secrete 2-3 liters/day of fluid. The secretion is an alkaline fluid (pH ~ 7.6) somewhat resembling pancreatic juice, but includes considerable cellular debris resulting from the sloughing of cells from the mucous lining of the intestine. The secretion contains mucus and digestive enzymes including enterokinase to activate the trypsinogen of the pancreatic juice, and trypsin, pepsin, and trypsin, maltase to break maltose into glucose, sucrose to break sucrose (table sugar) into halves forming glucose and fructose, lactase to break lactose (milk sugar) into glucose and galactose, and two nucleic acid digesting enzymes (ribonuclease and deoxyribonuclease). The secretory activity of the intestine is influenced by nervous, hormonal (e.g., enterocrinin), and mechanical stimuli, and again provides many navigationally-useful chemical gradients.

Third, there are ~5 million villi studding the entire mucosa of the small intestine (Fig. 8.21). Each villus is a fingerlike outgrowth of the mucous membrane averaging ~200 microns wide and ~1000 (range 500-1500) microns in length. (For comparison, a micron-scale nanorobot is about the size of one of the stipple dots in Figure 8.21.) Villi are taller and more numerous in the duodenum and jejunum (~20-40/mm^2) than in the ileum (~15-30/mm^2), and collectively increase the effective absorptive surface of the small intestine to ~10 m^2. They are leaf-shaped in the proximal duodenum, change to tongue-shaped by the upper jejunum, and finally become slender and finger-shaped in the ileum—probably all morphologically distinguishable for navigational purposes using a high-resolution navigational network (Section 8.3.3). Villi have a skin of simple columnar epithelial cells adapted for nutrient absorption and contain a capillary loop, a central lymphatic (lacteal) and connective tissue. Each villus is also provided with a small lengthwise strip of smooth muscle. During fasting, the villi lie flat on the mucosal surface and are inactive; when exposed to the intestinal contents during digestion, they become erect and perform a lashing movement driven by rhythmic shortening and lengthenings. These movements accelerate the flow of blood and lymph, and mechanically assist absorption by stirring the liquids of the gut in the immediate neighborhood.

Fourth, the surface area covered by the absorptive mucosal cells of the villi is also increased many times, estimated to be 14- to 39-fold, by the striated or brush border which consists of minute processes or microvilli projecting vertically into the lumen (not shown in Figure 8.21). These microvilli are extremely regular in size and evenly spaced: ~1 micron long, ~0.1 micron in diameter, and numbering ~1000 microvilli/cell.

Our touring nanorobot now leaves the ileum and enters the headwaters of the large intestine (the colon) through the ileocecal valve (valvula Bauhini), moving into the cecum. The ileocecal sphincter normally remains mildly contracted to slow the passage of chyme into the cecum. Immediately after a meal, a gastrointestinal reflex forces any chyme that remains in the ileum out into the cecum, and intensifies ileal peristalsis; the stomach hormone gastrin also reflex forces any chyme that remains in the ileum out into the cecum.

Fig. 8.21. Expanded view of the surface of the small intestine (redrawn from Mitchell and Arey).
is present as scattered solitary nodules but there are no aggregated glands are present, but they secrete mostly mucus. Lymphoid tissue stored in the transverse colon. All but ~100 cm³ of the 500-1500 dehydrated and compacted in the ascending colon, and temporarily stored in the colon is to absorb most of the remaining fluid and electrolytes and infected, causing acute appendicitis (see also Section 1.2.2).

There is almost no digestion in the large intestine. The task of the colon is to absorb most of the remaining fluid and electrolytes from the chyme. Each day -2000 cm³ of loose watery material is dehydrated and compacted in the ascending colon, and temporarily stored in the transverse colon. All but -100 cm³ of the 500-1500 cm³ of the water that enters the colon is absorbed, mostly in the colon and ascending colon. The mucous epithelium of the colon has a single layer of columnar cells, some of which have become specialized as unicellular slime glands (goblet cells); the colon coats feces with a thin layer of mucus to ensure smooth movement. (Amoeba move readily through glands (goblet cells); the colon coats feces with a thin layer of mucus.)

Structurally, the colonic tube is shunted into sacculations known as haustra, ~16 mm long segments running along the tube axis. Each haustrum is defined internally by opposing semilunar folds (corresponding to the creases on the outer surface) that act as mechanical baffles to hold the feces steady against lateral accelerations. There are mixing movements comprised of haustral contractions of the recesses of the colon, somewhat resembling the segmenting contractions of the small bowel. Three or four times a day mass peristalsis occurs, long-wave contractions that begin at the middle of the transverse colon and drive the more solid contents of the transverse colon into the ascending colon and sigmoidal colon. Food in the stomach triggers this reflex action. The large intestine has no mechanical baffles to hold the feces steady against lateral accelerations.

The mucous epithelium of the colon has a single layer of columnar cells, some of which have become specialized as unicellular slime glands (goblet cells); the colon coats feces with a thin layer of mucus. That's because gastric acids whose main purpose is to break down food more rapidly also act as a disinfectant to prevent the growth of microbes. The human colon, however, contains at least 400 species of bacteria, with -15 types of these accounting for the majority of the intestinal microflora. These bacteria ferment any remaining carbohydrates and produce surplus vitamins (e.g., vitamins K and B₁₂) and useful amino acids which are absorbed in the colon, so the relationship is quite symbiotic. About one-third of the material in the colon is microbial mass, and up to half of the stool consists of bacteria.

The bacterial population is also of navigational significance. From the esophagus to the ileocecal valve, the upper tract is relatively sterile with few live bacteria. That's because gastric acids whose main purpose is to break down food more rapidly also act as a disinfectant to prevent the growth of microbes. The human colon, however, contains at least 400 species of bacteria, with -15 types of these accounting for the majority of the intestinal microflora. These bacteria ferment any remaining carbohydrates and produce surplus vitamins (e.g., vitamins K and B₁₂) and useful amino acids which are absorbed in the colon, so the relationship is quite symbiotic. About one-third of the material in the colon is microbial mass, and up to half of the stool consists of bacteria.

The 100-200 grams of feces excreted daily is 65-75% water, 5-10% fatty material, and 20%-50% bacterial bodies, with a normal pH of 7.0-7.5. Even while fasting, 7-8 grams of feces are excreted daily. (Further details are in Chapter 26.)

The unpleasant fecal odor is due mainly to skatole and indole, produced by bacterial action in breaking down amino acids, but hydrogen sulfide and methyl mercaptan (more pronounced on a high-protein diet) also contribute. All of these chemical substances may be quantitatively measured in real time (Section 4.2) by medical nanorobots traversing the alimentary canal.

A simple map of the -0.65 m² cylindrical luminal surface of the gut (Table 8.8) to (200 micron)² villiary resolution requires just -16 million pixels or -130 million bits assuming 8-bit pixels. This data store can be compactly encoded on -0.005 micron³ of hydrofluorocarbon memory tape (Section 7.2.1.1) and may be carried aboard a micron-scale nanorobot in a storage device only -0.01 micron³ or -0.2 micron³ in volume. A complete map of the -10 m² absorptive surface to -20 micron² cellular resolution (excluding the brush border) requires -25 billion pixels or -0.2 terabits assuming 8-bit pixels.

8.2.4 Navigational Osteography

The skeleton is the single largest organ system, representing -14% of total mass and -11% of the navigable volume of the human body. The bony skeleton supports and protects the vital organs—the skull protects the brain; the spinal column shields the spinal cord and maintains erect posture; the ribs shelter the heart, lungs, and liver; pelvic bones protect the kidneys and internal sexual organs.

The total number of bones varies at different ages. At birth, the human body contains ~270 bones. This number declines slightly during infancy as a few separate segments join to form single bones. From young childhood through puberty, the bone count increases as wrist and ankle bones develop. Post-adolescence, the bone count steadily declines again with the gradual union of independent bones.

The adult skeleton (Fig. 8.22) consists of 206 bones: 28 skull bones (8 cranial, 14 facial, and 6 ear ossicles); the horsehoe-shaped hyoid bone of the neck (which breaks during hanging, choking the victim to death; see Figure 8.11); 26 vertebrae (7 cervical or neck, 12 thorax, 5 lumbar or loins, the sacrum which is five fused vertebrae, and the coccyx, our vestigial tail, which is four fused vertebrae); 24 ribs plus the sternum or breastbone; the shoulder girdle (2 clavicles, the most frequently fractured bone in the body, and 2 scapulae); the pelvic girdle (2 fused bones); and 30 bones in each of the four extremities (a total of 120). The paired bones include the 12 ribs on either side, 8 wrist bones, 5 hand bones, 14 finger bones, 7 ankle bones, 5 foot bones, 14 toe bones, 3 auditory ossicles, and the paretial, temporal, palatine, lacrimal, nasal, upper jaw, cheek, lower nasal concha, cartilagebone, shoulder blade, hip, arm, outer and inner
forearm, thigh, kneecap, calf, and shin bones. There are also a variable number of sesamoid bones, ranging from 8-18 in number, which are small rounded masses embedded in certain tendons and usually related to joints.

The total mass of the skeleton is ~10,000 gm; skeletal density averages ~1.45 gm/cm³, so total volume is ~6875 cm³ (Table 8.9). The total exterior surface area of the bones is ~1 m², which may be mapped to ~20 micron² cellular resolution using 2.5 billion pixels, requiring ~0.02 terabits of onboard nanorobot memory assuming 8-bit pixels.

More interesting are the interior spaces of bones in which medical nanodevices may travel. Bone is composed mainly of ~50 nm long crystals of very dense calcium salts having the hardness of marble (mostly hydroxyapatite, ~45% of bone mass) held together by bone matrix containing large numbers of extremely strong ~200 nm-wide collagenous fibers arranged in a dense mat together with a mucopolysaccharide cement sometimes called "ground substance." In living bone, ~25% of bone weight is water and another ~30% is organic material.

Embedded in this mineral structure are bone cells of three types. Osteoblasts, which secrete the substances that make up the bone matrix, line the outer surfaces of bone and also line many of the surfaces inside the internal cavities of the bone. Osteocytes, the most numerous bone cell type, originate from osteoblasts when those cells become trapped within small irregular matrix cavities called lacunae. After the trapped cells transform into osteocytes, they stop forming new bone but continue to support normal bone metabolism. Finally the osteoclasts (large multinucleated cells lining ~3% of internal bone cavity surfaces) remove old bone when it needs repair. Osteoclasts may be stimulated by parathyroid (endocrine) hormone to cause bone absorption when extra calcium ions are needed in the extracellular fluid. Bone is continually being resorbed and rebuilt to maintain its structural integrity.

Most of the compact bone is laid down in concentric layers, or lamellae, both on the outer surfaces of the bone and around the internal blood vessels that supply nutrition. Figure 8.23 illustrates this lamellar structure of bone, showing that the lacunae (irregular spaces) containing the osteocytes lie between the successive lamellae and are interconnected from one layer to another and between lacuna by small fluid-filled canaliculi averaging ~0.3 microns (range 0.1-1.0 micron) wide. The largest (Haversian) canal lies at the center of each system of concentric lamellae. Haversian canals are ~20 microns wide and carry blood vessels, lymph vessels, nerves, and also connect with the lacunae via the canaliculi. All osteocytes lie within ~200 microns of a blood capillary. The entire Haversian system allows nutrient and calcium flow through the bone interior volume and provides convenient navigable channels which medical nanorobots may utilize to gain direct access to bone cells and matrix materials.

Compact bone (providing the principal structural bearing strength of the body) is formed on the exterior of all bones. Deeper inside, the bone becomes a more open latticework with large spaces, called spongy or cancellous bone, whose solid parts nevertheless have an internal structure similar to compact bone. In the long bones of the extremities, the epiphyses (end knobs) are composed of spongy bone covered by a thin layer of compact bone. The diaphysis (central shaft) is made up almost entirely of compact bone surrounding a cavity containing marrow. Covering the entire bone is the periosteum, a nutrient-carrying fibrous membrane investing the surfaces of bones, except at the points of attachment of tendons and ligaments (where cartilage is substituted). The periosteum has an outer fibrous layer composed of dense fibrous tissue with blood vessels and an inner osteogenic layer containing many fibroblasts. Blood vessels

![Fig. 8.22. Anterior view of the human skeleton (redrawn from Guyton).](image-url)

![Fig. 8.23. Internal cellular structure of bone (redrawn from Guyton).](image-url)
traversing the periosteum enter the bone and pass through channels called Volkmann’s canals to enter and leave the Haversian canals. In humans, there are few true voids in the bones.

Bone marrow fills the spaces of spongy bone and the medullary cavity of long bones. It is composed of a supporting framework of reticular tissue in which there are blood vessels and blood cells in various stages of development. A thin membrane called the endosteum, more delicate than the periosteum but resembling it in structure, lines the medullary cavity. In the adult, there is red and yellow marrow. Red cells and some white cells are formed in the red marrow; in the newborn all marrow is red, but in the adult the red marrow is found in the spongy bone such as the proximal epiphyses of long bones and in the sternum, ribs, vertebrae and diploe of the cranial bones. Yellow marrow contains many fat cells and is found in the medullary cavity of long bones, producing macrophages and granulated white cells. There is ~3 kg of marrow in the adult body, occupying a volume of ~2400 cm³.

Even the relatively low capillary density of <100/mm² in bone implies a total investiture of ~3 x 10⁸ skeletal capillaries and a total luminal surface area of ~8 m² for the entire vasculature, which would require a ~0.16 terabit vascular map at cellular resolution. A full cellular map identifying each individual osteocyte in the entire skeletal system requires ~1 terabit.

An interesting navigable volume within the skeletal system is the human spine, which averages 71 cm in length in men and 60 cm in women with remarkably little variation. The body of each vertebra has a ring-shaped neural arch, through which passes the spinal cord. The spinal cord, measuring ~46 cm long and ~1 cm in diameter, descends from the brain and medulla oblongata through the foramen magnum and passes through the hollows of the spinal column, giving off 31 pairs of spinal nerves until it reaches the level of the disc between the first and second lumbar vertebrae, where the lower end tapers off to a point called the conus medullaris. The cord itself (Fig. 8.24) is composed of H-shaped gray matter in the center, extending forward as the anterior nerve roots (motor control) and backward as the posterior nerve roots (sensory), with solid white tracts of nerve fibers descending from the brain or ascending to it.

Cerebrospinal fluid (CSF) fills the ventricles of the brain and the subarachnoid cavity surrounding the spinal cord. The fluid is formed mainly by the choroid plexuses (large, reddish tufted organs) in the four ventricles of the brain. The plexuses are rich in blood vessels and separated from the cavity of the ventricles by only a single layer of secretory cells, thus affording ready access to the fluid by medical nanorobots (Chapter 16). The fluid, once formed in the brain, is later mostly reabsorbed by the arachnoid villi in the brain, but a small amount drifts slowly downstream along the spinal cord and is absorbed by the arachnoid villi through the spinal regions. Among other things, CSF is regarded as “the drainage system of the brain,” crudely analogous to urine, and the entire fluid volume is replaced once every ~30,000 sec.

The human body contains 90-150 cm³ of cerebrospinal fluid, normally maintained at an average pressure of 11 mmHg (range 5-13 mmHg). The fluid is clear and watery, normally containing 1-10 white cells/mm³ (all mononucleocytes) and similar chemical constituents as blood plasma, but with only 0.4% as much protein as the plasma; to first approximation, it may be regarded as an almost protein-free filtrate of the blood. The specific gravity of cerebrospinal fluid is 1.007 (1.0062-1.0082); viscosity at 310 K typically ranges from 0.7-1 x 10⁻³ kg/m·sec; and pH averages 7.4 (7.35-7.70), with ~1% (0.85-1.7%) total solids, 6 x 10⁻⁶ gm/cm³ glucose, 2.4-5.0 x 10⁻⁶ gm/cm³ cholesterol, and no acetylcholine or fibrinogen. Protein content increases from 1.0 x 10⁻⁴ gm/cm³ in the ventricles of the brain, to 1.5 x 10⁻⁴ gm/cm³ in the cisterna of the upper spinal column, to a high of 2.5 (2.0-4.0) x 10⁻⁴ gm/cm³ in the lumbar vertebrae—a potentially useful chemonavigational gradient.

Most of the ~150 bone joints in the human body are freely movable diarthroses (Fig. 8.25). There are ball and socket joints, saddle joints, hinge joints, and pivot joints. In the diarthrodial joint, two or more bones are united by an encircling band of fibrous tissue called the articular or fibrous capsule. The articular capsule is lined with...
Synovial membrane, and the apposed ends of bone are covered by a layer of hyaline cartilage, called articular cartilage. The fibrous capsule is reinforced and strengthened by ligament cords woven into it. Joint cavities sometimes extend into pouches or recesses or communicating bursae, and the shape of the sac changes with motion, as in the knee joint.

Synovial fluid, which lubricates and nourishes the joint, is exuded by the synovial membrane and fills the capsule cavity. There may be a total of ~30 cm³ of synovial fluid in all of the human body joints. Synovial fluid is a colorless, viscid mucin-containing material resembling egg white. The human knee contains -1.1 cm³ (range 0.13-3.5 cm³) of synovial fluid at an average osmotic pressure of 11 mmHg. The specific gravity of the fluid is 1.008-1.015; absolute viscosity is highly variable but averages ~160 gm/cm² mucin, and no fibrinogen.

There are also ~140 bursae (synovial sacs) in the body, located in fibrous tissue wherever friction occurs, probably containing at least ~20 cm³ of additional synovial fluid. The best-known include the subdeltoid bursa, which lies beneath the shoulder muscle; the prepatellar bursa, located in front of the kneecap; and the Achilles bursa, which lies between the heel bone and the Achilles tendon at the back of the heel.

8.2.5 Organography and Histonavigation

Measured by volumes, the typical -0.06 m³ adult male body contains ~38% hydrated protein, ~15% fat, ~14% expandable volumes and ~13% fluids. More importantly, ~60% of the volume of the human body is comprised of well-defined and highly specialized organs (Table 8.9) if skin and blood are included, as is commonplace.

Earlier Sections have described numerous physical routes through which all the organs can be accessed by a traveling medical nanorobot. Simple maps describing each of these routes can be made available to the nanodevice within a modest onboard data storage budget.

In addition to map-reading and precision positional navigation to ~3 micron accuracy (Section 8.3.3), with the help of a navigational network a nanorobot might verify that it had arrived at its intended organ by detecting various asymmetries that exist throughout the human body. As trivial examples, the heart lies mostly on the left side of the chest while the liver lies mostly on the right. The larger right lung has 3 lobes with 10 secondary bronchi segments, while the left lung has only 2 lobes and 9 secondary bronchi segments. The left ventricle of the heart has a muscular wall twice as thick as the right ventricle. The right lobe of the liver is much larger than the left lobe, and has a completely different shape. The right ear statistically is slightly more sensitive than the left ear, undoubtedly involving small-scale anatomical asymmetries. The left and right halves of the face have distinct physiognomies.

Aside from such gross anatomical measurements, there are many other (and more convenient) ways for a nanorobot to determine its histological location in the body. For example, in the kidney, the tissue pressure (easily measurable by nanosensors) is lower than the pressure under the tight renal capsule, and tissue pressure is lower in the brain than in the surrounding cerebrospinal fluid. (See also Section 8.4.2.) The speed of sound varies markedly from tissue to tissue (Table 6.7). For instance, a grid of nanorobots equipped with acoustic radiators and -nanosecond clocks (Section 10.1) could measure and distinguish the speed of sound in the brain (1550 m/sec), the kidney (1575 m/sec), and muscle (1600 m/sec) over a grid distance of ~100 microns (Section 8.3.3). In the human eye, at 20 MHz, the highest acoustic sound velocities are found in the sclera (1597 ± 20.3 m/sec); the lens exhibits the highest attenuation in the eye.

Each tissue has a unique spectral transmission signature at optical wavelengths. For example, liver looks distinct from bowel because of differences both in absorbance and in the way the tissue scatters light (see Section 4.9.4), making feasible automated discrimination among tissue types. Both scattering and absorption measurements will require at least ~100 micron path lengths (~5 cell widths) for reliable tissue discrimination, requiring a cooperative activity among several nanorobots using timed test pulses. A ~1 pW radiator power budget producing up to ~10⁷ optical photons/sec omnidirectionally using a ~1 micron² emitter transmits up to ~10⁶ photons/sec to a ~1 micron² receiver located ~100 microns away (Section 7.2.3), which should be adequate for tissue-type discrimination. Tissues might also be distinguished based on cell membrane electrical conductivity (Section 4.8.7).

In some cases, the local distribution of vitamin concentrations alone might be probative. Table 8.10 gives vitamin concentrations for various tissues, which may reflect chemical variations present in the intracellular fluids, the constituent cells, or both. Thiamine (B₁) occurs both free and phosphorylated, with elevated concentrations in the heart, liver, and kidneys, and low amounts in the stomach and skin; larger stores of riboflavin (B₂) exist in the kidney and liver than elsewhere; niacin (B₃) is more prominent in the liver and skeletal muscle; inositol is concentrated most heavily in the brain, kidney, and spleen. Cobalamin (B₁₂) is stored primarily in liver, kidney, lungs, and spleen. Vitamin A (retinol) is fat-soluble and is stored principally in ester form in the liver but also in the lungs and kidneys. Water-soluble vitamin C (ascorbic acid) is present in highest concentration in tissues of high metabolic activity, most notably the adrenal and pituitary glands and in the intestinal wall. Vitamin D (cholecalciferol) is most concentrated in the skin, while vitamin E is stored mainly in the heart, lungs, muscles, and fatty tissues, and vitamin K is absorbed through the colon and stored in the liver. All three are fat-soluble. With enough chemical markers to choose from, a nanorobot equipped with suitable chemosensors can make high-probability locational inferences.

Tissues also have unique extracellular matrices (ECM) that can be sampled and identified by traveling nanorobots (Section 9.4.4.2). For example, 19 different forms of collagen have been identified in various tissues, including 5 fibrillar types, 1 network-forming type, 4 fibril-binding types, 2 short-chain types and 1 long-chain anchoring type. Types III and VIII are found only in cardiovascular tissue, Type VII is found only in the skin, and so forth.

* Normally the major lobe of the liver is on the right side of the abdomen and the spleen is on the left side, but about one in 10,000 patients is born with situs inversus—a reversal of the left-right positions of these organs.127
Glycosaminoglycans (a component of the glycocalyx and ECM, usually complexed with proteins to form mucoproteins) are also tissue specific. For instance, chondroitin sulfates are found in cartilage, bone, and cornea. Keratin sulfate I is found in the cornea, keratin sulfate II in loose connective tissue, heparin in mast cells, heparan sulfate in skin fibroblasts and the aortic wall, and hyaluronic acid in synovial fluid, vitreous humor, and loose connective tissue. It appears that ECMs contain embedded structures and glycoproteins (e.g., osteonectin, tenascin) unique to each tissue type that should be recognizable to nanorobots even in the complete absence of cells. Intermediate filaments comprising the cytoskeleton inside the cell are also tissue-specific (Section 8.5.3.11).

Finally, organs may be identified by examining the surface antigens of their constituent cells. This is the most powerful technique. For

### Table 8.9. Mass, Volume, and Scale Size of the Organs of the Human Body

<table>
<thead>
<tr>
<th>Fluid, Tissue, Organ, or Organ System</th>
<th>Total Mass (gm)</th>
<th>Total Volume (L (^3) ~ cm(^3))</th>
<th>Size Scale per Organ (L ~ cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avg. Adult Male Body</td>
<td>70,000</td>
<td>60,000</td>
<td>~ 40</td>
</tr>
<tr>
<td>Muscle</td>
<td>30,000</td>
<td>23,000</td>
<td>--</td>
</tr>
<tr>
<td>Fat</td>
<td>10,500</td>
<td>12,000</td>
<td>--</td>
</tr>
<tr>
<td>Integument (total:)</td>
<td>6,100</td>
<td>5,500</td>
<td>--</td>
</tr>
<tr>
<td>Skin</td>
<td>2,000</td>
<td>1,800</td>
<td>--</td>
</tr>
<tr>
<td>Subcutaneous tissue</td>
<td>4,100</td>
<td>3,700</td>
<td>--</td>
</tr>
<tr>
<td>Skeleton (total:)</td>
<td>10,000</td>
<td>6,875</td>
<td>--</td>
</tr>
<tr>
<td>Cortical bone</td>
<td>4,000</td>
<td>2,235</td>
<td>--</td>
</tr>
<tr>
<td>Red marrow</td>
<td>1,500</td>
<td>1,200</td>
<td>--</td>
</tr>
<tr>
<td>Yellow marrow</td>
<td>1,500</td>
<td>1,200</td>
<td>--</td>
</tr>
<tr>
<td>Cartilage</td>
<td>1,100</td>
<td>845</td>
<td>--</td>
</tr>
<tr>
<td>Trabecular bone</td>
<td>1,000</td>
<td>670</td>
<td>--</td>
</tr>
<tr>
<td>Periarticular tissue</td>
<td>900</td>
<td>725</td>
<td>--</td>
</tr>
<tr>
<td>Lymph</td>
<td>~ 2,000</td>
<td>~ 2,000</td>
<td>--</td>
</tr>
<tr>
<td>Lymphoid tissue</td>
<td>1,400</td>
<td>1,350</td>
<td>--</td>
</tr>
<tr>
<td>Gastrointestinal tract</td>
<td>2,000</td>
<td>1,800</td>
<td>--</td>
</tr>
<tr>
<td>Contents (chyme/feces)</td>
<td>~ 2,000</td>
<td>~ 2,000</td>
<td>--</td>
</tr>
<tr>
<td>Blood vessels</td>
<td>1,800</td>
<td>1,700</td>
<td>--</td>
</tr>
<tr>
<td>Contents (blood)</td>
<td>5,600</td>
<td>5,400</td>
<td>--</td>
</tr>
<tr>
<td>Liver</td>
<td>1,650</td>
<td>1,470</td>
<td>11.7</td>
</tr>
<tr>
<td>Brain</td>
<td>1,400</td>
<td>1,350</td>
<td>11.1</td>
</tr>
<tr>
<td>Nerve trunks</td>
<td>--</td>
<td>330</td>
<td>6.7</td>
</tr>
<tr>
<td>Cerebrospinal fluid</td>
<td>90-150</td>
<td>90-150</td>
<td>--</td>
</tr>
<tr>
<td>Spinal cord</td>
<td>38</td>
<td>36</td>
<td>--</td>
</tr>
<tr>
<td>Lungs (2)</td>
<td>825</td>
<td>775</td>
<td>9.8</td>
</tr>
<tr>
<td>Contents (air)</td>
<td>~ 7.7</td>
<td>~ 6,000</td>
<td>--</td>
</tr>
<tr>
<td>Heart</td>
<td>330</td>
<td>300</td>
<td>6.7</td>
</tr>
<tr>
<td>Chamber volume</td>
<td>--</td>
<td>450</td>
<td>--</td>
</tr>
<tr>
<td>Kidneys (2)</td>
<td>300</td>
<td>270</td>
<td>5.2</td>
</tr>
<tr>
<td>Spleen</td>
<td>155</td>
<td>145</td>
<td>6.6</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>150</td>
<td>140</td>
<td>~12</td>
</tr>
<tr>
<td>Contents (urine)</td>
<td>~ 500</td>
<td>~ 500</td>
<td>--</td>
</tr>
<tr>
<td>Digestive fluids</td>
<td>~ 150</td>
<td>~ 150</td>
<td>--</td>
</tr>
<tr>
<td>Pancreas</td>
<td>110</td>
<td>100</td>
<td>4 x 15</td>
</tr>
<tr>
<td>Salivary glands (6)</td>
<td>50</td>
<td>48</td>
<td>2.0</td>
</tr>
<tr>
<td>Synovial fluid</td>
<td>~ 50</td>
<td>~ 50</td>
<td>--</td>
</tr>
<tr>
<td>Teeth (32)</td>
<td>42</td>
<td>14</td>
<td>0.8</td>
</tr>
<tr>
<td>Thyroid gland</td>
<td>40</td>
<td>36</td>
<td>3.3</td>
</tr>
<tr>
<td>Testes (2) male</td>
<td>40</td>
<td>37</td>
<td>2.6</td>
</tr>
<tr>
<td>Uterus (virgin/post-preg.)</td>
<td>35/110</td>
<td>30/100</td>
<td>~5</td>
</tr>
<tr>
<td>Eyeballs (2)</td>
<td>~12</td>
<td>10.8</td>
<td>2.4</td>
</tr>
<tr>
<td>Hair (avg. haircut)</td>
<td>21</td>
<td>16</td>
<td>--</td>
</tr>
<tr>
<td>Prostate gland</td>
<td>20</td>
<td>18</td>
<td>2.6</td>
</tr>
<tr>
<td>Adrenal gland (2)</td>
<td>20</td>
<td>18</td>
<td>2.1</td>
</tr>
<tr>
<td>Thymus gland</td>
<td>16</td>
<td>14</td>
<td>2.4</td>
</tr>
<tr>
<td>Ovaries (2) (female)</td>
<td>8</td>
<td>7.2</td>
<td>1.5</td>
</tr>
<tr>
<td>Gallbladder</td>
<td>7</td>
<td>7</td>
<td>3 x 8</td>
</tr>
<tr>
<td>Contents (bile)</td>
<td>~ 50</td>
<td>~ 50</td>
<td>--</td>
</tr>
<tr>
<td>Fingernails &amp; toenails (20)</td>
<td>1.1</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Pituitary gland</td>
<td>0.61</td>
<td>0.57</td>
<td>0.8</td>
</tr>
<tr>
<td>Pinal gland</td>
<td>0.2</td>
<td>0.18</td>
<td>0.5</td>
</tr>
<tr>
<td>Parathyroid glands (4)</td>
<td>0.15</td>
<td>0.14</td>
<td>0.3</td>
</tr>
</tbody>
</table>

*estimated for 70 kg adult male.*
instance, T cell receptors may be organ-specific: those comprised of $\gamma_2$ subchains are found in the spleen, $\gamma_3$ in the skin, $\gamma_4$ in the female reproductive organs and the tongue, and $\gamma_5$ in the lining of the intestinal tract. Protein components of the matrix, such as laminin and fibronectin, bind to specific integrin molecules on cell surfaces; through these 20+ integrins, the ECM transduces signals that regulate intracellular tissue-specific gene activity. The plasma membrane class B scavenger receptor for high-density lipoproteins, known as SR-BI, is expressed primarily by liver, ovarian, and adrenal cells. Cell type identification is a major topic that is addressed at length in Section 8.5.2.2.

Must a bloodborne nanorobot exit the bloodstream and enter the tissues in order to determine its proximity to the target organ? Most endothelial cells that coat the blood vessels servicing an organ or tissue system are likely to display organ-specific antigens on their luminal surfaces (Section 8.5.2.2). In these cases, organ homing can be accomplished without exiting the blood vessels, simply by sampling the antigenic signature of the vessel luminal walls. However, in some cases it may be necessary for a nanorobot to penetrate the endothelial layer and directly examine the surfaces of the underlying tissue cells or the underlying ECM to determine proximity to the target organ.

Once a nanorobot verifies that it has arrived at the intended organ, the next major challenge is intraorgan navigation. This may be accomplished by following standardized maps which may be customized to the individual patient by prior somatographic surveys (e.g., Section 8.4.1.4) and supplemented with direct positional and chemonavigational techniques. A complete somatographic description of all the organs in the body is beyond the scope of this book, especially since relevant material is readily available in the well-known *Handbook of Physiology* series published by the American Physiological Society, and in anatomy,853,854,863,866 histology,567,755,956 and cytology,531,938,939 textbooks generally. However, it is instructive here to convey a sense of place and scale by briefly reviewing the internal structure of the largest gland in the human body, the liver (see position in Figure 8.16).

The liver is a soft, plastic organ whose surface is pushed in by the surrounding organs. Thus the smaller left lobe bears concave impressions of the esophagus and stomach, while the larger right lobe bears concave impressions of the duodenum, the transverse colon, and the kidney.

Histologically, the liver is comprised of ~250 billion hepatic cells averaging ~18 microns in size. The human liver contains ~6 different cell types, but hepatocytes constitute ~80% of the cell population of the liver.938 The hepatocyte has 1-3 spherical nuclei (depending

<table>
<thead>
<tr>
<th>Human Tissue</th>
<th>Thiamine (gm/cm³)</th>
<th>Riboflavin (gm/cm³)</th>
<th>Niacin (gm/cm³)</th>
<th>Biotin (gm/cm³)</th>
<th>Inositol (gm/cm³)</th>
<th>Pantothentic acid (gm/cm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serum</td>
<td>0.01-0.09 x 10^-4</td>
<td>0.03-0.04 x 10^-4</td>
<td>0.02-0.15 x 10^-5</td>
<td>0.09-0.16 x 10^-7</td>
<td>0.03-0.07 x 10^-5</td>
<td>0.06-0.35 x 10^-6</td>
</tr>
<tr>
<td>Whole blood</td>
<td>0.03-0.10 x 10^-4</td>
<td>0.15-0.60 x 10^-4</td>
<td>0.5-0.8 x 10^-5</td>
<td>0.07-1.7 x 10^-7</td>
<td>--</td>
<td>0.15-0.45 x 10^-6</td>
</tr>
<tr>
<td>Mammary gland</td>
<td>0.43 x 10^-4</td>
<td>2.4 x 10^-4</td>
<td>1.0 x 10^-5</td>
<td>0.4 x 10^-5</td>
<td>2.7 x 10^-4</td>
<td>3.9 x 10^-6</td>
</tr>
<tr>
<td>Skin</td>
<td>0.52 x 10^-4</td>
<td>1.2 x 10^-4</td>
<td>0.86 x 10^-5</td>
<td>0.22 x 10^-7</td>
<td>2.0 x 10^-4</td>
<td>3.1 x 10^-6</td>
</tr>
<tr>
<td>Ileum</td>
<td>0.55 x 10^-4</td>
<td>4.2 x 10^-4</td>
<td>1.9 x 10^-5</td>
<td>0.6 x 10^-7</td>
<td>7.9 x 10^-4</td>
<td>5.3 x 10^-6</td>
</tr>
<tr>
<td>Stomach</td>
<td>0.56 x 10^-5</td>
<td>5.2 x 10^-5</td>
<td>1.9 x 10^-5</td>
<td>1.9 x 10^-7</td>
<td>7.6 x 10^-5</td>
<td>6.1 x 10^-6</td>
</tr>
<tr>
<td>Spleen</td>
<td>0.61 x 10^-4</td>
<td>4.3 x 10^-4</td>
<td>1.8 x 10^-5</td>
<td>0.25 x 10^-7</td>
<td>5.8 x 10^-5</td>
<td>3.9 x 10^-6</td>
</tr>
<tr>
<td>Seminal duct</td>
<td>0.69 x 10^-4</td>
<td>1.0 x 10^-4</td>
<td>0.92 x 10^-5</td>
<td>0.15 x 10^-7</td>
<td>&lt;1.0 x 10^-4</td>
<td>2.0 x 10^-6</td>
</tr>
<tr>
<td>Testicle</td>
<td>0.80 x 10^-4</td>
<td>2.0 x 10^-4</td>
<td>1.6 x 10^-5</td>
<td>0.9 x 10^-7</td>
<td>16.0 x 10^-4</td>
<td>5.0 x 10^-6</td>
</tr>
<tr>
<td>Colon</td>
<td>1.0 x 10^-4</td>
<td>2.1 x 10^-4</td>
<td>1.3 x 10^-5</td>
<td>0.9 x 10^-7</td>
<td>7.8 x 10^-4</td>
<td>5.0 x 10^-6</td>
</tr>
<tr>
<td>Spleen</td>
<td>1.1 x 10^-4</td>
<td>3.6 x 10^-4</td>
<td>2.3 x 10^-5</td>
<td>0.6 x 10^-7</td>
<td>10.3 x 10^-4</td>
<td>5.4 x 10^-6</td>
</tr>
<tr>
<td>Skeletal muscle</td>
<td>1.2 x 10^-4</td>
<td>2.0 x 10^-4</td>
<td>4.7 x 10^-5</td>
<td>0.35 x 10^-7</td>
<td>4.5 x 10^-7</td>
<td>12.0 x 10^-4</td>
</tr>
<tr>
<td>Smooth muscle</td>
<td>1.2 x 10^-4</td>
<td>2.3 x 10^-4</td>
<td>3.1 x 10^-5</td>
<td>0.6 x 10^-7</td>
<td>5.6 x 10^-4</td>
<td>6.2 x 10^-6</td>
</tr>
<tr>
<td>Lung</td>
<td>1.5 x 10^-4</td>
<td>1.9 x 10^-4</td>
<td>1.8 x 10^-5</td>
<td>1.9 x 10^-7</td>
<td>4.0 x 10^-4</td>
<td>5.0 x 10^-6</td>
</tr>
<tr>
<td>Adrenal gland</td>
<td>1.6 x 10^-4</td>
<td>8.2 x 10^-4</td>
<td>2.6 x 10^-5</td>
<td>3.5 x 10^-7</td>
<td>6.9 x 10^-4</td>
<td>8.0 x 10^-6</td>
</tr>
<tr>
<td>Brain</td>
<td>1.6 x 10^-4</td>
<td>2.5 x 10^-4</td>
<td>2.0 x 10^-5</td>
<td>5.8 x 10^-7</td>
<td>15.1 x 10^-4</td>
<td>15.0 x 10^-4</td>
</tr>
<tr>
<td>Liver</td>
<td>2.2 x 10^-4</td>
<td>16.0 x 10^-4</td>
<td>5.8 x 10^-5</td>
<td>7.4 x 10^-7</td>
<td>6.6 x 10^-4</td>
<td>43.0 x 10^-4</td>
</tr>
<tr>
<td>Kidney</td>
<td>2.8 x 10^-4</td>
<td>20.0 x 10^-4</td>
<td>3.7 x 10^-5</td>
<td>6.7 x 10^-7</td>
<td>12.4 x 10^-4</td>
<td>19.0 x 10^-4</td>
</tr>
<tr>
<td>Heart</td>
<td>3.6 x 10^-4</td>
<td>8.3 x 10^-4</td>
<td>4.1 x 10^-5</td>
<td>1.7 x 10^-7</td>
<td>5.0 x 10^-5</td>
<td>16.0 x 10^-4</td>
</tr>
</tbody>
</table>

Fig. 8.26A. Geometric arrangement of hexagonal lobules in the liver (redrawn from Cormack936).
upon cell volume and position in the structure), each containing a
nucleolus. These cells are bathed in flowing blood and arranged
into ~1 million roughly hexagonal functional units -1 mm wide
called hepatic lobules (Fig. 8.26A). There are no interlobular septa
(membranes separating the lobules) in the human liver.936

Around each hepatic lobule, liver cells form single-cell thick
trabecular or hepatic plates (Fig. 8.26B). Portal blood from the gas-
trointestinal tract enters the portal vein through the hepatic plate,
carrying absorbed nutrients into the portal venules. Each portal
venule opens directly into large numbers of venous sinusoids (open
cavities) formed by liver cells. The hepatic sinusoids form rich
vascular networks with extensive but irregular anastomoses, making a
three-dimensional spongework. Blood filters through the sinusoids
and exits via the central vein (Fig. 8.26C), from which it flows into
the hepatic venous system and from there into the inferior vena
cava. Hepatic arterioles diverging from the hepatic artery supply an
additional ~25% of the blood flow through the lobule, enriching
the venous blood with freshly oxygenated red cells.

Bile canaliculi 1-2 microns wide936 lie between adjacent liver
cells, forming networks of polygonal meshes each surrounding an
individual hepatic cell. The canaliculi empty into a bile ductule at
the periphery of the lobule. The ductules, in turn, join other ductules,
producing vessels of increasing diameter, leading eventually to the
wide hepatic duct and the common bile duct beyond.

The lymphatic system of the liver begins in the spaces of Disse,
which lie between the surfaces of the hepatic plates and the endothelial
walls of the venous sinusoids (Fig. 8.27). This perisinusoidal space
contains blood plasma but no red cells or platelets, and is lined with
vast numbers of microvilli projecting from the sinusoidal surface of
the hepatocytes bordering the space.936 The spaces of Disse empty
into small lymphatic vessels that run alongside the portal venules.
The sinusoid endothelial cells are flat and thin, with only their
nuclei protruding slightly into the sinusoid lumen. Their walls are
perforated with large fenestrae (~1 micron openings) numbering
~0.1/micron². The fenestrations are dynamic structures that undergo
changes in size and number in response to hormones and
cytoskeletal fiber inhibitors.884,885 There are also numerous smaller
pores measuring ~0.1 microns in diameter, mostly organized into groups
of 10-15 pores called sieve plates (Fig. 8.27). Sieve plates are ~0.7
microns wide and number ~0.3 plates/micron².936 These openings
admit particulate material into the Disse space, for disposal by
Kupffer cells.

Kupffer cells are ~15-20 micron stellate macrophagic cells,
mechanically attached to the sinusoid endothelial cells. They partially
occlude the sinusoid lumen but have no functional attachments to
the endothelial cells or the hepatocytes. Kupffer cells are motile and
have the ability to phagocytize (engulf and digest) particles of dirt,
inert nanodevices (engulf and transport; Chapter 16), worn-out
blood cells, and bacteria. Their customary positioning, predominantly
at the periportal end of the sinusoids, confirms that they monitor
arriving blood, looking for particles to remove from the flow. There
are ~25 billion Kupffer cells in the liver,884,886 with a lifespan of many
months.

With ~10⁶ hepatic lobules each containing ~250,000 liver cells,
a connectivity map listing every hepatocyte requires a 20-bit lobular
address plus an 18-bit cellular address (within the lobule). Thus a

Fig. 8.26B. Plate structure of a liver lobule (redrawn from Guyton863).

Fig. 8.26C. Schematic of flow geometry across a liver lobule (modified
from Cormack936).

Fig. 8.27. Expanded view of hepatocyte neighborhood (Kupffer cells
omitted; modified from Elias883).
simple hepatic lobular map (-1 mm resolution) takes just 20 megabits but a complete hepatic cellular map (-18 micron resolution) requires 250 billion 38-bit addresses, for a total map size of ~9.5 terabits. Of course, such complete maps may rarely be needed, since most medical nanorobotic tasks involve operations to be performed on a relatively small number of cells, or on a serendipitous basis, or can be guided by chemonavigation (Section 8.4.3).

8.3 Positional Navigation

Positional navigation should allow a medical nanorobot to know its three-dimensional coordinate position to -micron accuracy at all times. A nanodevice may rely upon dead reckoning, cartotaxis, microtransponder network alignment, or triangulation on external beacon signals, as described below.

8.3.1 Dead Reckoning

Perhaps the simplest but least accurate method of positional navigation is dead reckoning—the determination of position by keeping an account of the distance and direction traveled, without reference to any exogenous sources of information other than the beginning point. For example, starting from a well-defined initial location, a legged nanorobot counts the exact number of footfalls taken in all directions; the measured length of each footfall gives the distance traveled. However, anchor points on biological membranes are positionally unstable (Section 9.4.3.1), so a micron-sized bipedal walker with a 100-nm leg stride that achieves even a very optimistic 1% footpad positional stability (~1 nm) during each leg placement cycle will find that the accumulated error in its computed position has reached one body length (~1 micron) after just 1000 strides, or ~100 microns (~100 body-lengths) of locomotion. This level of accuracy is equivalent to ~one cell-width of error per ~2 mm (~20,000 strides) of travel.

Accuracy may be at most 2 orders of magnitude better when negotiating hard or very firm surfaces such as tooth enamel or bone. The classical positional variance of a telescoping nanomanipulator capable of 100 nm of horizontal travel is at best 0.01 nm (Section 9.3.1.4); a device stiffness of ~10 nN/nm gives a limb deflection of ~0.1 nm if ~nN forces are applied, resulting in a measurement error of 0.01%-0.1% at each footfall. Additional measurement errors due to varying strains caused by fluctuating normal loads are of similar magnitude. For example, strain $\varepsilon = m / E A = 0.01\%$ for the $2.5$-cm diameter human femur, taking a human mass of $m = 10^2$ kg supported by two femurs of total cross-sectional area $A = 10^2$ cm$^2$, with Young's modulus $E = 10^{10}$ N/m$^2$ for wet compact bone (Table 10.3) and acceleration of gravity $g = 9.81$ m/sec$^2$. Pressures exerted during chewing may reach 10-100 atm (Chapter 28), giving a maximum natural strain on tooth enamel of $(10^7$ N/m$^2$) / $E = 0.01\%$, taking $E = 7.5 \times 10^{10}$ N/m$^2$ for enamel (Table 10.3).

Navigational accuracy via dead reckoning may be up to 1-2 orders of magnitude poorer in other cases, especially during nanorobot swimming (Section 9.4.2) which requires additional corrections for fluid motions relative to surfaces. Sequential monitoring of accelerations and rotations alone produces even less accurate results. Assuming typical environmental accelerations of ~0.4 g's experienced by a 1-micron nanorobot, then data sampling at ~10 KHz to an accuracy of ~4 x 10$^{-5}$ g per measurement (Section 4.3.3.2) yields a cumulative error of ~0.4 g after ~10$^4$ measurements, a mere 1 second of travel. Pendular orientation sensors accurate to ~2 milliradian (Section 4.3.4.2) produce ~1 radian of accumulated error after a non-recalibrated chain of just 500 measurements.

Dead reckoning is most useful in two special circumstances:

1. hard-frozen tissues presenting immobile and highly anchorable surfaces; and
2. localized navigation requiring only very short distances to be traversed, such as locomotion between organelles in cyto or where abundant functional or positional cues allow frequent recalibration of the estimated current location.

8.3.2 Cartotaxis

Another positional navigation technique is simple landmark-centered map-following, or cartotaxis. Consider a nanorobot that wishes to crawl to a specific 1 micron$^2$ patch on the subvilliary mucosal surface of the small intestine. Ignoring the villi and microvilli but including the circular folds, the mucosal surface of the small intestine is ~1 m$^2$ (Section 8.2.3), at worst requiring a map of ~8 terabits to achieve 1 micron$^2$ resolution using 8-bit pixels.

A more compact map might include only the positions, sizes and shapes each of the ~100 million 50-micron-wide intestinal glands dotting the surface of the small intestine at least ~100 microns apart. These dots form a pattern unique to each person that changes slowly over time. After defining a two-dimensional intestinocoordinate system, two coordinates of log$_2$(10$^7$) - 20 bits each can specify a location to 1-micron accuracy on a 1 m$^2$ surface. Using 20 bits each for the two surface coordinates, 20 bits each for the major and minor elliptical axes framing the gland aperture, and another 920 bits to record unique topographic features of each hole, a complete small intestinal gland map requires at most ~1000 bits/gland or ~0.1 terabits.

This data storage requirement may be significantly further reduced by recognizing that on its way to its destination, an efficient ~5 micron$^3$ surface-walking nanorobot need pass over at most 0.002% of the 1 m$^2$ of small intestinal surface recorded on the map described in the previous paragraph, even assuming a maximum-length 4-meter whole-intestine traverse, passing at most a total of 4 m / 100 microns = 40,000 identified glands each requiring ~1000 bits to describe. Hence the minimum intestinoglandular map required for this longest traverse could in theory contain as few as ~40 megabits of data for a perfectly navigating nanorobot using a perfectly compiled and absolutely stable map. As a practical matter, to ensure reliability the onboard map conservatively should contain complete intermediate annular ring segment recalibration maplets at least ~500 microns in length along the tube axis and spaced ~2 cm apart, to be sequentially encountered axially while traveling down the tube. Nanorobots use dead reckoning to navigate between these guide rings until the ring closest to the destination is reached, analogous to the mid-course correction areas employed by cruise missiles. Having arrived at this nearest ring, the nanorobot follows a final map swath ~500 microns wide leading directly to the three glands lying closest to the destination patch; the final 1 micron$^2$ target patch is reached by interpolation and dead reckoning between these last three glands. Reliability and efficiency are enhanced by using large numbers of intercommunicating nanorobots simultaneously (Sections 7.3.2 and 8.3.3). This ~1 gigabit “practical map” contains complete descriptions of ~1 million individual intestinal glands and may be stored in a ~0.1 micron$^2$ onboard data spool using hydrofluorocarbon memory tape (Section 7.2.1.1).

The above scenario describes the ideal case. In actual living systems, biological surfaces are constantly in motion, are coated with slime, and are frequently being remodeled. Some mucosal surfaces may...
replace their entire luminal cell population every \(-10^7\) sec \((-1\text{ day})\). Thus the precise shapes and features of individual intestinal glands are constantly changing, although their larger structures and the pore pattern may remain intact for considerably longer periods of time. Intestinal glands are easily located by moving in the direction of rising concentration of enzyme-loaded intestinal juice, a basic chemonavigational technique (Section 8.4.3), and villiary trunks are available to provide additional cartographic guidance. But cartotaxic nanorobots will achieve the best results by using only the most recently prepared maps.

### 8.3.3 Microtransponder Networks

The archetypal high-resolution (-3-micron) internal navigational network may be described as a set of \(-10^{11}\) mobile acoustic transponder nanodevices, or navicytes, uniformly deployed throughout a \(-0.1\text{ m}^3\) human body volume with an average 100-micron spacing, (Power consumption rises sharply for much larger spacings.) Total navicyte fleet volume is a relatively unobtrusive \(-1\text{ cm}^3\). Navicytes emitting omnidirectional \(-100\text{ MHz}\) acoustic signal packets using \(t = 1\) micron radiators are \(-50\%\) energy efficient (Section 7.2.2). Navigational signal packets are \(-1\text{ microsec}\) in duration (conveying \(-100\text{ bits/packet}\) and repeat at \(-1\text{ millisecond}\) intervals \((-1\text{ kHz})\), giving a duty cycle of \(f_{duty} = 0.1\%\) which leaves plenty of clear air time for positional information inquiries from non-navigational nanorobots, communications and sensor traffic, local acoustic microscopy and the like, and also to allow staggered time slots to avoid noise and crosstalk.

How is a 100 MHz signal packet detected in a 1 nanosec signal processing time? At \(SNR = 2\) (Section 4.5.1), a navicyte acoustic sensor must receive at least \(kT \varepsilon^{SNR} \approx 30\text{ zJ}\) within a 1 nanosec integration, an energy influx rate of \(-30\text{ pW}\) at the receiver. To produce 30 pW at a 1-micron receiver located 100 microns away, a transmitter of equal area must radiate \(-300\text{ ,000}\text{ pW}\) of acoustic output (see Section 4.9.1.5) or \(-10^{-5}\text{ watts/m}^2\) for \(-1\text{ nanosec}\). Each navigational signal packet is prefaced by a triangular pulse \(-1\text{ nanosec}\) in duration. This triangle pulse is used for ranging. Because the transmitter is \(-50\%\) efficient, each triangle pulse requires \(-600,000\text{ zJ}\) of input energy to the transmitter in order to produce an acoustic output energy of \(-300,000\text{ zJ/pulse}\) at the transmitter surface. Broadcasting \(-1000\text{ pulses/sec}\) brings the required transmitter input power to \(-600\text{ million}\text{ zJ/sec or -0.6 pW}\) continuous. Transmitting the non-pulse portions of all packets costs \(-60\text{ pW}\) (Section 7.2.2.2). Note that the 1 nanosec triangular ranging pulse has the highest frequency that can be used without significant absorption. Given a pulse which deposits a small fixed multiple of \(kT\) at the receiver, then using the highest frequency produces the least uncertainty in distance. Very roughly, a pulse which is just barely detectable can have its timing localized to \(-1\text{ pulse width}\).

The triangle pulses make \(-9\text{ atm}\) acoustic spikes in water, probably low enough to avoid transient cavitation at this frequency and too brief for shock wave formation or stable cavitation (Section 6.6.1.1)—though it might be a good precaution to vary the time intervals between packets to preclude any possibility of unexpected resonances. Acoustic torque and related effects (Section 6.4.1) cannot yet be ruled out in this application and should be investigated further. Each triangular pulse represents an energy discharge event requiring a momentary power density of \(-10^{12}\text{ watts/m}^3\) inside a \(-0.8\text{ micron}\) powerplant, comparable to power densities available in mechanical (Section 6.3.2), chemical (Section 6.3.4.5), and electrical (Section 6.3.5) systems. At \(-100\text{ pW/navicyte}\), total navicyte network system power consumption is \(-10\text{ watts}\), well within proposed in vivo thermogenic limits (Section 6.5.2).

How is the navigational network established? For convenience, a navigational Prime Centrum is established on the ventral surface of the 10th thoracic vertebra (T-10) at the midsagittal plane of the vertebral body (see Figures 8.22 and 8.24), defining a permanent origin for a body-centered coordinate system. This origin is centrally located, lying directly posterior to the xiphoid process (Fig. 8.22) and the liver, directly inferior to the heart and between the two lungs. The site is easily accessible by bloodborne nanorobots, well-supplied with oxygen and glucose for power, securely anchorable due to the dense bone, and reasonably well-protected from injury; also, movements in the thoracic vertebrae are the most restricted because the ribs and costal cartilages resist distortion.

The Prime Centrum is comprised of four “monument” type bloodborne navicytes which are injected, or migrate stochastically using cytoidentification (Section 8.5.2), or are directed to the site by demarcation or other means. These nanorobots congregate around the chosen navigational origin approximately at the vertices of a square. Each monument navicyte attaches one end of a retractable fullerenic cable rule to each of its three brethren and crawls backward on an approximately radial course, paying out the slippery cable and adjusting mutual positions until the two diagonals measure exactly 141.42 microns center-to-center and the four normals measure exactly 100.00 microns center-to-center, guaranteeing a perfect 100-micron square with exact 90.00° corners. The four navicytes then secure themselves to the bone beneath the peristeum using biocompatible permanent anchors, becoming fully sessile, and retract the cable rules. If vertebral curvature radius \(R \approx 2\text{ cm}\) and diagonal navicyte separation \(S = 141.42\text{ microns}\), then maximum anteroposterior geometric deviation due to off-sagittal positioning is \(R(1-\sin((\pi/2)-(S/R))) = 0.5\text{ micron}\). The square may not be precisely aligned with the conventional anatomical coordinate axes. At least 10 independent navicyte monument quartets should be established as alternate or backup sites on T-10 to satisfy customary redundancy requirements (Chapter 13). Additionally, \(-10^3\) regional monuments are established at \(-cm\) intervals at fixed positions on all major skeletal surfaces of the body. Relative positions of regional monuments vary only within well-defined envelopes depending upon macroscopic joint rotations and limb flexures, or almost not at all on inflexible surfaces such as the diaphysis of the long bones.

Mobile navicytes deployed throughout the remaining body volume receive message packets emitted by neighboring devices, which in turn have received packets from their neighbors, ultimately stretching back in an unbroken chain to a regional monument or to the Prime Centrum. Assuming a simple cubical array, each stationary navicyte has two neighbors per directional axis—a total of six neighbors within acoustic communication range (100 microns).

Navicyte positional stability is enhanced by stationkeeping activities to avoid drift \((-1\text{ micron/sec}\) for a \(-1\text{ micron}\) nanorobot due to Brownian motion; Section 3.2.1), and anchorage to nonsignalling elements of the omnipresent extracellular matrix.

Each navicyte possesses an onboard clock capable of continuous \(\Delta t \approx 1\text{ nanosec}\) temporal accuracy between recalibrations or over mission times of \(\approx 10^3\) sec (Section 10.1). Message packets received

---

* In the fundus, the cell turnover cycle is \(-5\text{ days}\), while in the body of the stomach the whole cycle takes \(-1\text{ day}\). The time required for complete renewal of intestinal epithelium is \(-2.3\text{ days}\) for the duodenum and \(-2.8\text{ days}\) for the ileum.\(^{351}\) At the other extreme is the lens of the eye—one of the few structures containing cells that is preserved without turnover\(^{531}\).
periodically from each neighbor include data describing the exact universal time of packet transmission. Since each recipient has a synchronized clock (Section 10.1.3), the triangle pulse travel time between navicytes (τ ~ 65 nanosec) is known to an accuracy of -1 nanosec. The speed of sound vsound ~ 1540 in soft tissue (Table 6.7), so the -1 nanosec temporal uncertainty adds (vsound Δτ) ~ 1.5 microns of uncertainty to the range estimate.

The speed of sound varies between 1400-1600 m/sec for most nonosseous tissues. This speed is reasonably uniform within specific tissues (Table 6.7) over time, and is essentially frequency-independent over the nanomedically-relevant range. Knowledge of its own approximate histological location (based on chemical sampling, etc.) allows a navicyte to consult an onboard data table or a previously-compiled low-resolution map (see Chapter 19) to estimate local sound velocity to within -25 m/sec. Also, two navicytes (at least one of which is mobile) can directly measure the local speed of sound by briefly conjugating (Section 9.4.4.4), extending a cable rule of known length between them, then transmitting and timing a test pulse traveling through the medium. A 100-micron rule length allows local sound velocity to be measured to an average Δvsound ~ 25 m/sec accuracy using a 1-nanosec clock. A measurement uncertainty of -25 m/sec in the local speed of sound adds another (Δvsoundτ) ~ 1.6 micron of positional uncertainty, giving a total of ΔXmin ~ 3 microns uncertainty in each range estimate. Additional small uncertainties in sound velocity and in acoustic reflection and refraction power losses may occur when capillaries or other microvessels cross the line of sight between two navicytes.

Only four of the six neighbors are absolutely required for positional triangulation. The data packet received by a navicyte from its first neighbor (containing that neighbor’s correct three-dimensional coordinates) narrows the navicyte’s possible position to a spherical surface of radius equal to the computed range, centered on the first neighbor. The data packet received from the second neighbor defines a second geometric sphere, further narrowing the navicyte’s possible position to a circle formed by the intersection of the first and second spheres. The data packet from the third neighbor adds a third sphere, reducing the possibilities to two points on the intersection circle, and the signal packet from the fourth neighbor selects one of these two points as the navicyte’s true position. Additional minor corrections may be made for the bending of sound waves crossing known thermal, pressure, or salinity gradients (Section 6.4.1).

What positional accuracy can this system achieve? Consider the simplest case with many parallel coplanar rows of N navicytes lying -Xrowi apart, each row of total length Lrow = Σ Xrowi emanating from a common sagittal surface and extending deep into the tissues. The terminal navicyte of each row estimates its cumulative length as Lrow, which also contains an unknown range error εrow. The series of randomly distributed errors in the positions of each navicyte in the row, individually of magnitude ± ΔXmin do not cancel to zero but instead constitute a random walk with maximum error excursion from the mean of εrow ~ 2N1/2 ΔXmin. The longest rows will occur in the viscerum, farthest from any bony surface. Taking Lrow = 15 cm, ΔXmin = 3 microns and Xrow = 100 microns, then N = Lrow/Xrow = 1500 navicytes per row and εrow ~283 microns. Hence in this simple case, the minimum accuracy at the terminus of each row is 15 cm ± 230 microns, or -0.2%. Shorter rows accumulate less error at the terminus—the cumulative error at -2 mm from the common surface is at worst -27 microns (-1 cell width). The average error per navicyte in a 15-cm row is only εrow / N ~ 0.2 microns. Even for Lrow = 2 meters (-longest possible transverse path length in the human body), N = 20,000 and εrow = 800 microns. Note that an additional small systematic error in position may accrue if there is a systematic change in local sound velocity between recalibrations and if the subject viscera have a free surface unbounded by bone. For example, bruised tissue in which foreign fluids are accumulating will produce a slightly warped coordinate system; sound velocity differentials between, say, blood and interstitial fluid may be as large as 70 m/sec (Table 6.7).

Accuracy within each row may be significantly improved by introducing active error checking and continuous recalibration. Consider the terminal navicytes of two parallel rows A and B. As before, the rows lie -Xrow apart in their common plane. Each terminal navicyte computes that its total length is precisely Lrow, but in fact navicyte B is in error by a distance εB along the row. Given the minimum range error ΔXmin, terminal navicyte B can only detect that an error exists when its range measurement to neighboring terminal navicyte A increases from Xrow to Xrow + ΔXmin, whereupon from simple geometry:

\[ ε_B = \left(\frac{X_{row} + ΔX_{min}}{X_{row}}\right)^2 - X_{row} \right]^{1/2} \]  

[Eqn. 8.2]

For Xrow = 100 microns and ΔXmin = 3 microns, then εB ~25 microns or -1 cell width. This range error is relatively insensitive to Xrow; for instance, if Xrow increases to 110 microns then εB only rises to -26 microns.

However, row terminus errors (measured in integral units of εB) are likely to be normally distributed between ± εrow ~ 230 microns in the earlier simplest-case example. The error correction procedure involves nearest neighbors collectively polling the nearest nrow = 6000 rows (a bundle with termini covering -0.6 cm² when Xrow = 100 microns) for their measured εB’s and then computing the mean of the distribution, which has an uncertainty of εrow / nrow1/2 ~ 3 microns = ΔXmin, matching the uncertainty of a single range measurement, the smallest possible error. The estimated mean is used to produce a correction factor which is applied to the erroneous estimated value of Lrow, yielding a corrected Lrow accurate to -ΔXmin or -3 microns. This corrective process is repeated:

a. for each row,

b. for planar cross-sections at -2 mm increments along the entire row length (to eliminate error compression or rarefaction waves), not just along the terminal plane, and

c. at regular time intervals (-3 sec) to ensure continuous recalibration to -3 microns at all points in each row.

Note that parts of the human body frequently may deform up to -30% during normal activities, so Xrow is not constant but may vary between 85-115 microns in as little as 0.1-1 sec in working situations.

Detailed specification of a complete recalibration protocol is beyond the scope of this book. One procedure to further improve row-length measurement accuracy is to use the navicyte grid as a phonon gain medium by configuring each navicyte as a repeater station during the calibration cycle, effectively allowing the propagation of row-long pulses thus permitting independent row-length measurements. Multiple measurements can reduce independent row-length measurement error to arbitrarily low levels. Alternatively, J. Soreff suggests treating the entire calibration bundle as a coherent amplifier using lower acoustic frequencies to increase range. This enables phase detection to within -1 nanosec if phase-locked detection is used and if a few kT of energy are present in the time resolution window, allowing navicytes to hear a propagating plane wave averaged across an entire bundle. An analysis of soliton-like systemic effects analogous to intrinsic local modes in lattices is beyond the scope of this text.
The ability to make the corrections described above assumes that navicytes can distinguish $-\varepsilon_B$ from $+\varepsilon_B$, an angular spread of $\pm 14^\circ$ for $\varepsilon_B = 25$ microns and $X = 100$ microns. Since sound crosses the width of a micron-sized nanorobot in -1 nanosec, the relative angular location of a distant acoustic source can be directly measured by placing two sensors on either side of the nanodevice at precisely calibrated separations. For angles of acoustic wave incidence $\leq \theta$, measured from the axis joining the two sensors which are separated by a width $x_{sensor}$ on either side of the nanorobot, the wave arrival times will differ by more than $\Delta t - 1$ nanosec and thus will be received as two distinct pulses, rather than one. For incidence angles $>\theta$, the wave arrival times at either side cannot be distinguished. $\theta$ thus defines a permissive angular detection cone of size

$$\theta = \frac{\pi}{2} - \sin^{-1} \left( \frac{\Delta v_{sound}}{x_{sensor}} \right) \quad \text{(Eqn. 8.3)}$$

For $\Delta t = 1$ nanosec and $v_{sound} = 1540$ m/sec, a sensor separation of $x_{sensor} = 1.59$ microns allows up to $\theta = 0.25$ radian = $14^\circ$ to be distinguished. The uncertainty in $v_{sound}$ of $-25$ m/sec imposes a minimum measurement uncertainty of $\Delta \theta = 3^\circ$. The permissive detection cone shrinks to $\theta = 0^\circ$ at $x_{sensor} = 1.54$ microns.

Relative navicyte angle can also be computed from the coordinates of neighbors by simple geometry. Using the coordinates of a nearest neighbor with range $X_{range} = X_{row} = 100$ microns and lateral positional uncertainty $X_{error} = \Delta X_{min} = 3$ microns, then relative navicyte angle can only be computed to an accuracy of $\Delta \theta = \sin^{-1} \left( X_{error} / X_{range} \right) < 1.7^\circ$ (30 milliradians). However, relative angle uncertainty is reduced if the coordinates of more distant neighbors are available to the navicyte. For example, using a neighbor located 15 rows away ($X_{range} = 15X_{row}$) and again taking the minimum $x_{sensor}$, then $\Delta \theta < 0.1^\circ$ (-2 milliradians). Coordinates from a distant-neighbor navicyte with $X_{range} = x_{sensor} = 15 - 15$ cm with an uncorrected lateral positional uncertainty $X_{error} = v_{sound} = 230$ microns would give $\Delta \theta < 0.09^\circ$ (-2 milliradians); taking $X_{error} = \Delta X_{min} = 3$ microns in the ideal case, $\Delta \theta < 0.001^\circ$ of arc (-0.02 milliradian). By comparison, from Eqn. 3.2 Brownian tumbling of a micron-scale nanorobot amounts to $-10^{-6}$ (-0.02 microradian) in -1 nanosec, or $-1^\circ$ (-20 milliradians) between -1 millisecond signal packet repeat intervals.

Each navicyte obtains its own orientation relative to gravity using onboard gravity sensors, which are accurate to within -2 milliradians of verticity, with new measurements available every -0.1 millisecond (Section 4.9.2.2) or up to the limits of the communication network capacity (e.g., recalibration protocols). Regional monitors monitor and disseminate the local grid’s angular orientation relative to the gravity field at -1 millisecond intervals, thus fixing each navicyte’s orientation to the local grid on a virtually continuous basis. Absolute spatial orientation relative to a fixed onboard standard such as a nanogyroscoper (Section 4.3.4.1) may be determined to -100 microarcsecond accuracy during nonrealcalibrated deployment lifetimes of $10^5-10^7$ sec. Direct gyrostabilization may be possible in some applications (Section 9.4.2.2).

### 8.3.4 Vascular Bifurcation Detection

The ~5 billion navicytes circulating in the 5400 cm$^3$ blood volume can be used to form a dynamic transulinal geometric “virtual lattice” (with constantly changing membership) that remains positionally stationary near the bifurcation of each blood vessel whose diameter exceeds $X_{row} = 100$ microns. A spatial progression of such rows allows the bifurcation break point to be registered at most -100 microns past the junction and in at most -20 millisece above and below the larger-diameter vessel, for nanorobots drifting with the blood flow in arterioles. In the high velocity aorta (Table 8.2), minimum bifurcation detection range increases to ~630 microns in a detection time of -1 millisecond. Transmission of local navicyte coordinates allows traveling nanorobots to determine which branch of the bifurcation they are about to enter, or have already entered. (Major organs, lymphatic topology, and other important landmarks can be accurately located by similar methods.)

Bifurcations involving capillaries and other vessels with diameters less than $X_{row}$ may be readily detected by direct acoustic reflection echolocation from vessel walls (Section 4.8.2). Echoes from red blood cells are relatively weak (though still measurable) in comparison with those from the solid tissues within the body because the characteristic impedance of red blood cells is very similar to that of the plasma in which the cells are suspended. The speed of sound in blood is somewhat reduced by completing a whole-body scan of vsound as a function of position, then loading this map into each nanorobot using the navigational system.* In some cases, satellite transmissions (e.g., the 24-satellite Global Positioning System or GPS) to determine their position on Earth’s surface. Such a system might involve externally generated signals from beacons placed at fixed positions outside the skin, or could employ dedicated internally-installed emitter organs. As described in the microtransponder example (Section 8.3.3), at least four beacon signals must be detectable simultaneously to allow a fix to be determined in three-dimensional space. Macrotransponder networks may be useful in some applications, but also have several important drawbacks as described below.

Since in a macrotransponder system the nanorobots determine their positions by direct detection of beacon range signals rather than by polling their neighbors, accuracy may be solely a function of the total range measurement uncertainty $\Delta X_{min}$. As before (Section 8.3.3), given a typical signal path length $X_{path} = 15$ cm and a mean $v_{sound} = 1540$ m/sec, signal travel time ($\tau = 97,400$ nanosec) is known to an accuracy of $\Delta \tau = 1$ nanosec if beacon clocks and nanorobot clocks are stable and synchronized, adding $v_{sound} \Delta \tau = 1.5$ microns of range uncertainty.

However, by far the greatest source of uncertainty in a macrotransponder system is the variation in the speed of sound along the randomly chosen linear paths through the human body which the beacon signals must follow. The speed of sound is reasonably uniform within specific tissues or local regions, but an arbitrary 15-cm path through the body may plausibly involve sound velocities ranging from 630 m/sec for a path entirely through the lungs up to -4090 m/sec for a path entirely through the densest skull bone (Table 6.7). This implies a velocity uncertainty range of $\Delta v_{sound} = 3460$ m/sec and thus a maximum range uncertainty of $\Delta X_{range} = 33.7$ cm—approximately the anteroposterior thickness of the human body. There is also some variation in $v_{sound}$ as a function of temperature (Section 10.5.5) and other factors.

While local (~100 micron range) sound speed may be measured on site by one or a small number of nanorobots (Section 8.3.3), the net sound speed along an arbitrary -15 cm path cannot be similarly assessed without a comprehensive survey effort. Pathwise uncertainty in sound velocity is somewhat reduced by completing a whole-body survey of $v_{sound}$ as a function of position, then loading this map into each nanorobot using the navigational system.* In some cases, standardized maps with modest customization will suffice. A whole-body $v_{sound}$ map to ~100 micron resolution, giving $\Delta v_{sound} = 25$ m/sec,
would require ~1 terabit of onboard storage, a minimum 38 micron$^3$
memory tape volume inside a (~4 micron)$^3$ memory mechanism. 
Assuming a more reasonable ~1 micron$^3$ tape volume allows a ~0.026
terabit velocity map to be stored onboard, providing (~340 micron)$^3$
resolution and giving $\Delta v_{\text{sound}} = 84$ m/sec.
Unfortunately, a measurement uncertainty of ~84 m/sec in the
local speed of sound still adds ($\Delta v_{\text{sound}}$) ~ 8 mm of positional
uncertainty to each ~15 cm range measurement, a ~5% error.
Shortening the average beacon-nanorobot distance to $X_{\text{path}} = 1$ cm
(e.g., by installing ~60,000 internal emitter organs, which may be
unduly intrusive) reduces positional uncertainty to ~500 microns,
possibly sufficient in some applications.

Adaptively stable surfaces on the skin or even inside the body
may be difficult to find for large macroscopic transmitters, causing
significant source movement and degrading the positional accuracy
of the beacons as system monuments. This problem is partially
overcome by using acceleration compensators (e.g., gimbals and
gyros to "float" transmitters), thus neutralizing the effects of minor
body motions. However, because of the minimum ~8 mm range
uncertainty noted above, it will be difficult for beacons to mutually
recalibrate their own positions to high accuracy, so there will be
unavoidable unmonitorable monument movements especially during
rapid joint rotations and limb flexures. And macroscopic trans-
mitters are physically more responsive to the small random centrifugal
and gravitational forces normally experienced by the body than are
microtransmitters, adding to their moment-by-moment positional
uncertainty.

By now, the alert reader will be wondering if measurements of
just the angular positions of at least three externally located macro-
scopic acoustic beacons, taken without regard to range measurements,
can suffice for accurate navigation. The answer is that is an angles-only
system is quite possible, but unfortunately offers little improvement.
(GPS uses 4 satellites so that inexpensive handheld receivers don’t
need accurate clocks and thus can avoid ranging errors. Orbiting
satellites do have atomic clocks and transit time codes, but the
ground-based receiver need only compare time differences between
received signals.)

Consider an acoustic sensor pair with $x_{\text{sensor}} = 1.54$ micron, giving
the minimum possible angular measurement error $\Delta \theta = 3^\circ$ (Section
8.3.3). The positional uncertainty $\Delta x$ of the nanorobot is then

$$\Delta x = \frac{X_{\text{path}} \sin(\Delta \theta)}{\sin\left(\frac{\pi}{2} - \Delta \theta\right) \left(\frac{1}{3} N_{\text{beacon}}\right)^{1/2}} \quad \text{(Eqn. 8.4)}$$

where $X_{\text{path}}$ is the average distance between beacon and nanorobot
and $N_{\text{beacon}}$ is the number of noncollinear beacons sampled during
each measurement cycle. Assuming an epidermal beacon network
with $X_{\text{path}} = 15$ cm, $\Delta x = 8$ mm for $N = 3$ beacons; $\Delta x = 400$
microns for $N = 1000$ beacons.

Beacon signals will also travel many adjacent pathways with sound
speeds that differ only very slightly, producing a multiplicity of
received pulses and a smearing effect, further degrading signal quality.
Signals suffer refraction over such lengthy courses, causing the beam
to bend significantly from its original path in unknown amounts.
By contrast, the acoustic energy received from a pulse traversing
a ~100 micron path lies close to the sensor detection limit; these signals
are rapidly attenuated by the medium beyond this limit, and thus
produce no confusing detectable signal at the next nanovite station
located ~200 microns from the transmitter. Another problem with
macrobeacons is that they must use lower frequencies (~100 KHz;
Section 6.4.1) due to the long path lengths their signals must traverse
unamplified, hence maximum data flow rate may be up to ~1000
times slower than for the microtransponder system. Finally, macrobeacons permit only a one-way flow of navigational information,
a major limitation.

8.3.6 Dedicated Navigational Organs

By direct analogy to navigation organs (Section 7.3.4),
dedicated macroscopic organs may be implanted in the human body
to facilitate navigation. One important function of such organs
would be to act as central clearinghouses for systemic information
that is normally available only regionally or to individual navicyles.
This information might include updated positions of all regional
monuments, orientations of various grid sectors relative to gravity,
or information on macroscopic rotation rates (e.g., compasses,
gyros) to allow centrifugal and gravitational forces to be distinguished.
Highly accurate chronometers could also be maintained in these
organs (Section 10.1.4) to allow periodic systemwide resynchronization.

Dedicated navigation organs could serve as the "map rooms" of
the body, collating and organizing new information as it comes in,
maintaining accurate navicyle grid maps, organ maps, vascular maps,
fuctional maps (Section 8.4), and so forth, possibly in coordination
with dedicated computational organs (Section 10.2.5). A 1 mm$^3$
navigational library node can contain up to ~10 million terabits of
map data, which may be downloaded to nanorobots berthed at
docking ports at rates of up to 0.01 terabit/sec drawing ~50 pW
per docking port during the transfer (Section 7.2.6).

Dedicated navigation organs may be employed as direct interfaces
between internal navigational systems and related macroscopic
external modalities including operating theaters, clinical equipment,
various environmental entities, satellite uplinks, radio antennae,
transportation vehicles, and the like. They could also serve as emitter
organs in GPS-like acoustic navigational systems (Section 8.3.5).

8.4 Functional Navigation

Functional navigation allows a medical nanorobot to detect and
respond to subtle variations in tissue characteristics, often in the
absence of precise positional knowledge. Such tissue characteristics
may include thermal, acoustic or barostatic, cytochemical or
immunochemical, electrical (Section 4.9.3.3) or magnetic,1256
mechanical or topological variations. A few of the many possibilities
are described below.

8.4.1 Thermographic Navigation

8.4.1.1 Thermography of the Human Body

The human body presents a complex and temporally varying
spatial temperature field. The external parts of the body have a lower
mean temperature than the internal parts, with temperature decreasing
along the longitudinal axis of the extremities, producing both axial
and radial temperature gradients. The differing heat production
of individual organs, geometric irregularities, changes in insulation
and evaporation, convective heat transport via the blood, and the
diurnal and other periodic variations$^{327}$ add further complexities to the
thermal map.890

The homeothermic core of the body is distinguishable from the
shell, which most readily responds to environmental fluctuations.
The core generally consists of the interior of the thorax and abdomen,
the brain, and part of the skeletal muscles. With moderate changes
in ambient temperature, the shell normally comprises the outermost
20%-35% of the human body.$^{894,895}$ However, during extreme
or hot environment (modified from Wenger and Hardy,865 and from Fig. 8.28. Distribution of isotherms in a human body placed in cold or hot environment (modified from Wenger and Hardy,865 and from Aschoff and Wever891).

Skin temperatures display the greatest thermographic variability in response to external factors. For example, nude humans standing for 3 hours in a cold room (5°C, 50% relative humidity, 0.1-0.2 m/sec wind speed) experience skin temperature differentials up to 15°C (= 13°C to 28°C), with the lowest temperatures in fingers and toes, the highest in trunk and forehead, and average core/surface gradient -15°C.896 Heat loss is ~10% lower for females due to their thicker layer of subcutaneous fat, making their cold-room skin temperature slightly lower than for males.898 After 3 hours in a hot room (50°C), skin temperature differentials amounted to only 2.5°C (= 35°C to 37.5°C), with an average core/surface gradient of -1°C.896 With normal clothing in a room at 15-20°C, mean skin temperature is 32-35°C.

Skin thermography of the human head was first reported by Edwards and Burton;897 skin vs. rectal temperatures at various ambient temperatures are well-studied.895,898 Skin temperature patterns in neonates reflect near uniform heat conduction through the tissues.917 In childhood, specific patterns develop into a stable, permanent adult pattern.916 The dermothermal patterns differ in lean and obese patients, and exhibit a continual state of small rhythmic change. These changes —probably a result of active vasodilation due to sympathetic innervation over most of the human skin area919—are observed over the arms, hands, trunk and head, but are not all in phase with each other, nor even of the same amplitude.918,919,3334-3338 Skin thermology, including infrared thermography, is now an important branch of medical diagnostic imaging.899,912 Subcutaneous (shell) temperatures generally increase with depth.

Human core (rectal) temperature averages 37.0°C,894 but this simple number hides considerable natural variation (Table 8.11). The temperatures of inner organs vary by 0.2-1.2°C under normal room conditions, and by up to 0.9°C within individual organs.890 Temperature gradients within the brain amount to 1.4°C; the cortex is cooler than the basal regions, with incoming blood cooler than the central brain tissue.900 Brain temperature also decreases during sleep and rises during periods of emotional arousal.901 Cooling or warming the skin of the head causes temperature changes in the tympanic membrane (Fig. 7.3) of up to 0.4°C due to the returning venous blood.900 Oral temperature can have wide variation, depending for example on the thermal character of recently ingested food and drink, and thus has little direct correlation with swings in rectal temperature.910 Liver temperatures lie 0.2-0.6°C subrectal,902 possibly due to heat loss via the respiratory tract through the diaphragm. The temperature of the testes in the scrotum is ~3°C lower than that of the abdominal cavity.890 The temperature of the hypothalamus, the seat of human thermoregulation, might be the most logical choice as a reference value.890

Table 8.11. Core Temperatures of Organs and Blood Vessels in the Human Body890,894,904,907

<table>
<thead>
<tr>
<th>Blood Vessel, Organ or Tissue</th>
<th>Normal Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>32-35°C</td>
</tr>
<tr>
<td>Scrotum</td>
<td>34.0°C</td>
</tr>
<tr>
<td>Sternal</td>
<td>34.5°C</td>
</tr>
<tr>
<td>Subclavian vein</td>
<td>36.4°C</td>
</tr>
<tr>
<td>Liver</td>
<td>36.4-36.8°C</td>
</tr>
<tr>
<td>Oral cavity</td>
<td>36.5-36.6°C</td>
</tr>
<tr>
<td>Transverse facial artery</td>
<td>36.5-36.8°C</td>
</tr>
<tr>
<td>Superior vena cava</td>
<td>36.65°C</td>
</tr>
<tr>
<td>Urine</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Esophagus</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Aorta</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Inferior vena cava</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Pulmonary artery</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Lungs</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Pulmonary vein</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Heart, right ventricle</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Femoral artery</td>
<td>36.75°C</td>
</tr>
<tr>
<td>Common carotid</td>
<td>36.8°C</td>
</tr>
<tr>
<td>Kidney</td>
<td>36.85°C</td>
</tr>
<tr>
<td>Spinal cord</td>
<td>36.95°C</td>
</tr>
<tr>
<td>External jugular vein</td>
<td>37.0°C</td>
</tr>
<tr>
<td>Hepatic vein</td>
<td>37.0°C</td>
</tr>
<tr>
<td>Stomach</td>
<td>37.0°C</td>
</tr>
<tr>
<td>Rectum, mean</td>
<td>37.0°C</td>
</tr>
<tr>
<td>Rectum, range</td>
<td>36.2-37.8°C</td>
</tr>
<tr>
<td>Hypothalamus</td>
<td>37.3°C</td>
</tr>
<tr>
<td>Brain</td>
<td>37.3°C</td>
</tr>
<tr>
<td>Uterus</td>
<td>37.3°C</td>
</tr>
</tbody>
</table>

Human blood temperature also displays considerable variation. Outflowing blood is warmer than inflowing blood in organs with high metabolic rates, so the blood has a cooling effect in these organs, the opposite of the skin. Venous blood temperature in the extremities may be ~3°C lower than the arterial supply, but even the arterial blood can sink to 22°C in human extremities at cold room temperatures.903 Blood in the common carotid artery remains within 0.2°C of oral temperature, but falls 0.2-0.5°C as it cools the face and possibly also due to cooling from the internal jugular vein.905 However, inhalation of cold air has little effect on pulmonary blood temperature; even for patients in rooms cooled to -18°C, blood in the pulmonary vein and artery differ at most by 0.03°C.906

There are also periodic variations in body temperature. The diurnal fluctuation is the best known. This cycle is absent in neonates, but develops during the first weeks of life.913 Children have higher daytime temperatures than adults and a daily range of 1.7°C,890 with the daily range peaking near the age of 1 year.914  

Fig. 8.28. Distribution of isotherms in a human body placed in cold or hot environment (modified from Wenger and Hardy,865 and from Aschoff and Wever891).
standardized tests on young adults found that rectal core temperature fluctuated between 36.83-38.32˚C in males (range 1.20˚C) and 37.16-38.36˚C (range 1.20˚C) in females, and temperatures normally reaching their lows near 6 AM local time and their highs near 6 PM. The timing is keyed to environmental influences such as light/dark cycles, not personal activity.

The menstrual cycle also involves a complex long period fluctuation. During this monthly cycle, morning core temperature falls from 37.0˚C to 36.7˚C just before and during menstruation, after which there is an abrupt dip to -36˚C at midperiod near ovulation, followed by a return to normal levels for the remaining ~2 weeks. Menstrual-related skin temperature pattern variations are also observed.

Finally, there are various irregular influences on body temperature. Rectal temperature may rise up to 3.5˚C during the most extreme exercise, or may fall 2˚C after ~1 hour immersion in water at 23˚C. Ingesting food elevates skin temperature, and to a lesser extent the rectal temperature, for 1-2 hours. Alcohol ingestion raises skin temperature but lowers core temperature, while nicotine ingestion (e.g., smoking) lowers skin temperature in the extremities by 2-7˚C. In persons with fever, the temperature may rise from 37.0˚C up to 41.5˚C with daily fluctuations of 2.8-5.0˚C, usually peaking in late afternoon. Transient elevations to 44-45˚C have been recorded but are very rare.

In view of this measurable spatial and temporal variability, maintaining a high-resolution whole-body thermal map in real-time is impractical using in vivo nanodevices. Given the availability of -microkelvin sensors with -microsecond measurement cycles (Section 4.6.1), sensor technology is not the limiting factor. Map stability is the main concern. A volumetric map with only ~1 mm³ resolution, comparable to the best thermographic images, requires 10⁸ voxels and 17 bits/voxel for 0.001˚C accuracy. A hand placed on a stove produces thermal gradients >10°C/sec in the shell, but periodic temporal variations are 0.1-40 microkelvins/sec in the core and 0.005-1°C/sec in the shell during changes in physical exercise levels (~3˚C in 10-100 sec), movement between hot and cold environments (~10˚C in 10-100 sec), and fevers (~5˚C in ~1000 sec). Thus to maintain a map that reports all changes of ~1 millikelvin requires the entire field to be resampled every 0.001˚C / (1°C/sec) = 1 millisecond, generating ~2 terabits/sec. This dataflow possibly could be accommodated by a dedicated high-capacity fiber network (Section 7.3), a nanorobot traveling within a microvolume of blood that could crudely infer its distance from the epidermis based on the local conductivity, and the data storage capacity of individual bloodborne nanorobots.

Maintaining a hemathermographic map employing capillary-volume resolution to millikelvin accuracy generates an even more intractible >10¹⁵ bits/sec.

8.4.1.2 Thermal Demarcation

Thermographic navigation to a target site is possible without the use of any maps. Simple thermal demarcation is a good example of this. In a controlled room-temperature clinical setting, the relatively narrow homeostatically-maintained temperature band of 36-38˚C allows easy demarcation of selected target regions by applying thermal stimuli lying just outside the normal band. Examples might include hot (T > 38˚C) or cold (T < 36˚C) packs applied to the skin; diathermic tissue heating using different frequencies to achieve selective heating at various tissue depths (Section 6.4.2); adaptive phased array (APA) heating of small target volumes deep inside the body; thermally active catheter probe tips inserted transdermally or otherwise, directed to specific internal targets; or multiple intersecting focused infrared, visible photon, or ultrasound beams directed to specific convergence volumes for maximum heating effect. Depending upon the technique employed and the target tissue, volumes as small as ~1 mm³ may be demarcated in this manner. Nanorobots of relatively simple design may be injected near the target or allowed to circulate in the bloodstream, and will detect the outlier temperature data and gather in the vicinity of the target at a statistically determined accumulation rate and number density.

Outside of controlled environments or during prolonged treatment periods, simple demarcation becomes less reliable for several reasons:

1. Error in the physical placement of the instrument of demarcation will produce erroneous nanorobot localization, which may increase the likelihood of treatment error. Such iatrogenic mistakes are especially likely during time-critical or emergency crises, or in situations where the diagnosis is unclear, or in cases of self-administered medical care.

2. Even when initially correctly placed, the artificial hot or cold region will diffuse outward from the original spot, enlarging the radius of action in imprecisely known directions (due to nonisotropic tissue conductance), again increasing the risk of iatrogenic effects.

3. Load error in the human thermoregulatory control system is ~0.2-0.5˚C. Any thermal deviation of this magnitude or larger from local temperatures will provoke a natural counteractive response to restore the local temperature, including capillary sphincter action (both vasodilator and vasoconstrictor activity), increased sweat gland secretions, accelerated counter-current exchanges, more rapid heart rate, or local thermogenesis.

4. Increased susceptibility to spoofing, as when a patient enters a hot shower or bath, lies down on snow or on a cold concrete surface, washes his hands and face with hot water, drinks a hot or cold beverage, basks in the sun, and so forth. Temperatures hot or cold enough to definitively prevent spoofing might damage tissues if maintained for very long. This problem can be avoided by employing an oscillating artificial thermal gradient instead of a time-invariant source (especially across layers of highly insulating subcutaneous fat), but at the cost of increased required overall nanorobot sophistication to detect the oscillating demarcation. For energy delivered to a target L ~ 1 mm from the source in a medium of heat capacity C V and thermal conductivity K T, the maximum frequency that may be resolved is ω < \sqrt{\frac{K_T}{C_V L^2}} \approx 0.15 Hz for aequous tissue at 37˚C.

General positional information is also available to nanorobots equipped with thermal sensors in the absence of any maps. For instance, assuming the temperature of the external environment has been measured and disseminated via the communication network (Section 7.3), a nanorobot traveling within a microvolume of blood can crudely infer its distance from the epidermis based on the local blood temperature, especially when passing through the thermal shell of the body surrounding the core. Other cues including velocity measurements or histonavigational data may allow further refinement of the estimate.

8.4.1.3 Low-Resolution Thermographics

Even simple thermographic maps may have great utility. If the deviation from mean temperature to 3 significant figures (~0.01˚C accuracy requiring ~10 bits/measurement) is resampled at ~10 Hz, then 0.1˚C/sec gradients may be detected in each of ~10⁵ voxels (volume ~1 cm³/voxel) using ~10⁵ resident thermographicites, producing a 10⁶ bit/sec dataflow. A dermal thermographic network divided into 1 cm² squares requires ~20,000 resident nanorobots and produces a ~2 x 10⁶ bit/sec whole-network dataflow at
0.01°C accuracy (vs. 0.5-1.0°C accuracy for traditional thermograms).\textsuperscript{899} In either case, each thermographicyte in the network (Section 7.3) maintains in its \(-10^{10}\) byte memory (Section 7.2.6) the entire continuously updated network map and can store \(-1\) hour of historical data (~7 x 10\(^9\) bits) for the whole-body thermographic map or \(-1\) year of historical data (\(-3\) x 10\(^9\) bits) taken at its own position.

Measured skin temperature can be used for macrosensing (Section 4.9). As a simple example, knowledge of average skin temperature allows estimation of the environmental wind chill (air temperature and velocity) and relative wind direction using the theoretical equation given by Lampietro\textsuperscript{890} or the experimental data compiled by Mitchell et al\textsuperscript{893} over the range of 10-50°C and wind velocities of 1-5 m/sec, although the presence of clothing may be a complicating factor.

Thermal patterns appearing on centimeter-resolution volume or surface maps may be useful either for macrosensing or for the determination of various medically relevant states. Some examples of measurable patterns include:

1. bimodal thermal variations on the soles of the feet, while the patient is standing on a cold concrete floor or a hot asphalt pavement;
2. cooling of the back and seat after settling into a chair whose surface is initially at room temperature (chair material may be inferred from the subsequent temperature warm-up profile);\textsuperscript{5330}
3. warm pattern of a hand placed on the chest while patient hears the national anthem at a sports event;
4. cool vertical lines caused by tears running down cheeks;
5. asymmetric temperature profile caused by lying on one side or the other, in bed or at the beach;
6. warmth or coolness on palm and thumb of one hand while carrying a hot or cold plate of food, or grasping a cold can of beer;
7. cool knee spots while kneeling on a cold floor;
8. comparison of facial and scalp temperatures permitting inference of hair length and coiffure;
9. patterns of muscle warmth indicating physical activity involving specific tendons, limbs or appendages;
10. approximate arrangement and density of clothing\textsuperscript{5332}
11. increased warmth of the sexual organs, during excitation or tumescence, or of the skin, stomach, and liver, following a meal;
12. dangerous thermal gradients (e.g., hand placed on stove) detected in 0.1 sec, faster than the usual \(-0.2\) sec human response time; and
13. specific thermographic dermal patterns that signal the presence of a wide variety of medical conditions such as ischemic limbs, Raynaud’s Syndrome, cerebral apoplexy, reflex sympathetic dystrophy, muscular injuries, abnormal joints, ankylosing spondylitis (Pott’s disease), spinal root syndrome, osteoarthritis, tennis elbow, osteoid osteomas, headaches, breast cancers, or melanomas.\textsuperscript{899}

Assuming exogenous positional information is available to fixed thermographicytes, a population of these devices can constitute a permanent monitoring network that can issue alerts and detect:

1. temperature profiles of internal or external lesions located anywhere in the body;
2. thermal anomalies reflecting the presence of interior hematomas, lipomas, myomas, edemas and hydromas, new fibrous deposits or air pockets, or dental cavities;
3. slowly growing tumors (e.g., by searching for localized, shallow, but monotonic thermal gradients in the \(-\)microkelvin/sec range in the historical data);
4. unusual skin patterns due to defects in nervous control of the peripheral circulation;
5. intestinal surface patterns reflecting identifiable thermal characteristics of materials passing through the bowels, or alterations in cell metabolism resulting from food toxins or inflammatory reactions; or
6. alterations in the thermal behavior of individual organs at specific times of day, under varying workloads, or over long timespans, possibly indicating a change in functionality or general health (measurable by taking advantage of the different thermal characteristics of body tissues; Table 8.12).

Mobile nanorobots passing near any thermographicyte can interrogate the device’s thermographic data library and thus may examine all or part of the current whole-body coarse thermal map. Similarly, historical and current data can be transmitted to medical personnel who interrogate the network.

8.4.1.4 High-Resolution Thermographics

High-resolution maps can be assembled in designated tissue monitoring volumes, using a network of fixed thermographicytes whose physical positions are accurately known (Section 8.3.3). If \(L_{\text{therm}} = 100\) microns is the mean spacing between devices, then for a maximum transorgan temperature differential \(\Delta T \approx 1\) °C for an organ with scale size \(d_{\text{organ}} = 10\) cm (Table 8.9) the largest endogenously generated temperature differential between neighboring thermographicytes is \(\Delta T / L_{\text{therm}} / d_{\text{organ}} \approx 10,000\) microkelvins; the smallest detectable thermal variation of \(-1\) microkelvin (Section 4.6.1) corresponds to a fluctuation of \(-60\) watts/m\(^2\) in power density within the 0.001 mm\(^3\) monitoring volume (the equivalent power output of \(-100\) mitochondria). Recording a single temperature measurement to microkelvin accuracy requires at most \(-24\) bits.

Many useful configurations of such a network are readily conceived. As a simple example, consider a monitoring unit consisting of 100 thermographicytes reporting at \(-1\) KHz to one centrally located data processing nodule. The thermographicytes transmit a total of 2.4 megabits/sec to the nodule from within their 0.1 mm\(^3\) patrol volume. A nodule with \(-8\) micron\(^3\) of storage tape can retain up to \(-1\) day’s worth of historical data generated within its monitoring unit. Nodule nanocomputers can scan the data for a wide range of specified anomalies and thresholds, emitting an appropriate alert message (containing all relevant details) when one is found. In principle, an expanding (warm) tumor could be detected\textsuperscript{*} after

\* Tumors which are metabolically less active are harder to detect by this means. Also, measuring heat from tumors located in organs having high metabolic activity (e.g., liver, kidney) is even more difficult, especially given the large caloric variations after consuming food or drink, or during normal hormonal responses, and given the relatively high normal rate of cell division in some organs (e.g., liver, gut) that must be distinguished from abnormal tumor growth.
growing to a volume of at most \( V_{\text{thermal}} \), a maximum pathological cell count of \( \sim 100 \). Each monitoring unit consumes at most \( \sim 20,000 \) pW, including 100 pW/thermographicyte and 10,000 pW in the nodule for a \( \sim 1 \) micron\(^3\) computational facility drawing \( \sim 1000 \) watts/m\(^3\)-Hz (Section 6.5.6 (E)) and operating internally at a \( \sim 10 \) MHz clock speed. For continuous surveillance of an entire organ to microkelvin accuracy, up to \( \sim 16 \) million monitoring units are required, dissipating at most \( \sim 0.3 \) watts, which must also be taken into account. (The liver, for example, generates \( \sim 10 \) watts metabulously.)

Patients may also receive thermographicytes installed in a mobile configuration, with the nanodevices constantly traversing the tissues looking for anomalies, outmessaging only when such anomalies are detected, and accurately fixing their location to within \( \sim 3 \) microns at all times by interrogating a separate microtransponder positional navigation network (Section 8.3.3). At the extreme limits of accuracy, nanorobot waste heat may contribute significant error to thermal measurements. A nanorobot generating waste heat \( P_{\text{heat}} = 100 \) pW produces \( \Delta T = 10 \) microkelvin temperature differential over an \( X = P_{\text{heat}} / K_t \) \( \Delta T = 16\)-micron aqueous path; this nanorobot heat is

---

### Table 8.12. Thermophysical Characteristics of Various Body Tissues, Organs, and Other Materials\(^{460,567,817-819,1865,2153}\)

<table>
<thead>
<tr>
<th>Material, Organ or Tissue</th>
<th>Thermal Conductivity ( K_t ) (watts/m-K)</th>
<th>Heat Capacity ( C_v ) (MJ/m(^3)-K)</th>
<th>Approx. Density (kg/m(^3))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td>Very warm</td>
<td>2.80</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td>Normal hand</td>
<td>0.960</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td>Cool</td>
<td>0.545</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td>Upper 2 mm</td>
<td>0.376</td>
<td>3.77</td>
</tr>
<tr>
<td></td>
<td>Cold hand</td>
<td>0.335</td>
<td>3.77</td>
</tr>
<tr>
<td>Subcutaneous fat</td>
<td>High values</td>
<td>0.450</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Pure fat</td>
<td>0.190</td>
<td>1.96</td>
</tr>
<tr>
<td>Muscle</td>
<td>Living muscle</td>
<td>0.642</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td>Excised, fresh</td>
<td>0.545</td>
<td>3.64</td>
</tr>
<tr>
<td></td>
<td>Skeletal muscle, living</td>
<td>0.372</td>
<td>--</td>
</tr>
<tr>
<td>Bone</td>
<td>Mineral</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Cortical</td>
<td>2.28</td>
<td>2.70</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>1.16</td>
<td>2.39</td>
</tr>
<tr>
<td></td>
<td>Cancellous</td>
<td>0.582</td>
<td>2.07</td>
</tr>
<tr>
<td>Blood</td>
<td>Water at 310 K</td>
<td>0.623</td>
<td>4.19</td>
</tr>
<tr>
<td></td>
<td>Plasma (Hct = 0%) at 310K</td>
<td>0.599</td>
<td>4.05</td>
</tr>
<tr>
<td></td>
<td>Whole blood (Hct = 40%)</td>
<td>0.549</td>
<td>3.82</td>
</tr>
<tr>
<td>Organs</td>
<td>Heart (excised, near fresh)</td>
<td>0.586</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td>Liver (excised, near fresh)</td>
<td>0.565</td>
<td>3.78</td>
</tr>
<tr>
<td></td>
<td>Kidney (excised, near fresh)</td>
<td>0.544</td>
<td>4.08</td>
</tr>
<tr>
<td></td>
<td>Abdomen core</td>
<td>0.544</td>
<td>3.89</td>
</tr>
<tr>
<td></td>
<td>Brain (excised, near fresh)</td>
<td>0.528</td>
<td>3.86</td>
</tr>
<tr>
<td></td>
<td>Brain (living)</td>
<td>0.805</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Lung (excised, bovine)</td>
<td>0.282</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td>Whole body (average)</td>
<td>--</td>
<td>4.12</td>
</tr>
<tr>
<td>Air</td>
<td></td>
<td>0.009246</td>
<td>0.00119</td>
</tr>
<tr>
<td>Cotton fabric at 310 K</td>
<td></td>
<td>0.0796</td>
<td>0.0267</td>
</tr>
<tr>
<td>C(<em>{60})/C(</em>{70}) bulk compacts/crystals</td>
<td>0.1-0.4</td>
<td>2.3-2.5</td>
<td>1540-1676</td>
</tr>
<tr>
<td>Rubber</td>
<td></td>
<td>0.156</td>
<td>2.41</td>
</tr>
<tr>
<td>Ethanol at 310 K</td>
<td></td>
<td>0.163</td>
<td>1.96</td>
</tr>
<tr>
<td>Teflon</td>
<td></td>
<td>0.399</td>
<td>2.20</td>
</tr>
<tr>
<td>Concrete</td>
<td></td>
<td>0.934</td>
<td>1.93</td>
</tr>
<tr>
<td>Glass, plate</td>
<td></td>
<td>1.09</td>
<td>1.94</td>
</tr>
<tr>
<td>Ice at 2x9 K (-42˚C)</td>
<td></td>
<td>2.21</td>
<td>1.76</td>
</tr>
<tr>
<td>Sapphire (normal to c axis) at 310 K</td>
<td>2.20</td>
<td>2.89</td>
<td>3970</td>
</tr>
<tr>
<td>Stainless steel</td>
<td></td>
<td>13.8</td>
<td>3.68</td>
</tr>
<tr>
<td>Aluminum</td>
<td></td>
<td>204</td>
<td>2.45</td>
</tr>
<tr>
<td>Silver</td>
<td></td>
<td>405</td>
<td>2.59</td>
</tr>
<tr>
<td>Diamond, natural</td>
<td></td>
<td>2000</td>
<td>1.82</td>
</tr>
</tbody>
</table>
dissipated by thermal conduction to below measurable levels in $-C_n T^2 / K_t$ - 2 milliseconds, approximating the ~1 KHz sampling time.

8.4.2 Barographic Navigation

The internal pressure field of the human body (already briefly summarized in Sections 4.9.1, 6.3.3, and 8.2) is as complex as the temperature field. (Section 9.4.1 discusses the rheology of the circulation.) In general, fluid pressures range from 60-150 mmHg in the heart, arteries and arterioles; ~30 mmHg in the capillaries and 5-20 mmHg in the veins; 4-30 mmHg in the pulmonary artery; 9-13 mmHg in the cerebrospinal and synovial fluids; ~1.4 mmHg in the interstitial fluid between cells; and ~0.9 mmHg mean pressure in the lymphatics, with arterial pulsations contributing 1.5-2.2 mmHg oscillations with a maximum of ~5 mmHg in the lymphatics. (Systolic pressure may reach a maximum of ~230 mmHg during the heaviest exercise.)

As in thermographic, simple barographic demarcation provides a ready means for directing medical nanorobots to a desired treatment location within the body. Simply pressing down on the surface of the skin creates an anomalous region of elevated tissue pressure; tapping, slapping, or palpitating the body's surface is also detectable by in vivo nanorobots (e.g., Section 7.4.2.1). Acoustically active catheters inserted transdermally can place a sonic beacon signal near almost any target location. Highly focused but relatively low power (to reduce reflection echo detection) ultrasonic columnar or planar beams of, say, three different frequencies can be directed into the body on various trajectories and from various sources. The tiny spot where all three frequencies are audible at precisely equal intensity marks the target, providing $\leq 1 \text{ mm}^3$ localization accuracy at frequencies $\geq 1.5$ MHz. Focused ultrasound surgery (FUS) and high-intensity focused ultrasound (HIFU) may use ~2-second bursts of 1.7-MHz, 150-mm-deep focused beams at $5 \times 10^7$ watts/m$^2$ to occlude blood flow or destroy tumors by creating a spot of intense heat so small that there is a boundary of only ~6 cells between destroyed tissues and completely unharmed tissues, a ~100 micron spatial resolution.

As in the case of temperature fields, low-resolution and high-resolution barographic maps can be compiled and stored by barographicites, then used to diagnose a wide variety of medically relevant states in organs and throughout the body including tumors, edemas, excess cranial pressure, general hypertension or hypotension, swollen lymph glands, muscle spasms and the like. Such maps are also useful in macrosensing, as for instance to help determine if the patient is sitting, standing, prone, inverted, falling, or floating (Section 4.9.2). Blood pressure increases in patients who are exercising, eating a meal, or emotive (e.g., angry), and decreases in patients who are relaxing or sleeping. The typical daily readings in a hypertensive 35-year-old male might be 150-220 mmHg systolic and 90-140 mmHg diastolic.

However, the amplitude and timing of pressure waves in the blood vessels permit somewhat clearer positional inferences than is possible with thermographic, in part because the heart is such a strong, reliable, and stably-positioned acoustic pulse generator. Functional hemobaric mapping thus may provide useful positional information regarding vascular range estimates.

For example, Figure 8.29 shows the change in shape and amplitude of the pressure wave with increasing distance downstream from the aortic valve (where the left ventricle of the heart initially discharges into the descending aorta). There are three features of interest. First, each successive profile shifts to the right, suggesting wave propagation. Second, the sharp dicrotic notch in the pressure record (marking the closure of the aortic valve) is gradually lost in successive profiles, a clear navigational marker of aortic range.

![Fig. 8.29. Blood pressure profiles recorded at various distances along the aorta (data from dog aorta; distances measured from beginning of descending aorta; modified from Olson'923).](image-url)

* In 1998, heart sounds originating from multiple sites within the four chambers of the heart, its valves and the great vessels remained difficult to analyze. Despite the close coincidence of valve closure and high frequency components in the first heart sound (echocardiographic corroboration), the exact mechanism of sound production was poorly understood. Phonocardograms (PCGs) are distorted during transmission through the chest wall; cardiac auscultation, traditionally used to detect a wide variety of normal and pathological heart conditions, was still more art than science.
Third, there is a steepening and increasing of amplitude in successive profiles, indicating a rise in peak systolic pressure with distance from the heart. This counterintuitive increase is a dynamic phenomenon of an elastic branching system with tapered tubes—with steady flow in a rigid tube of similar diameter and taper, pressure always decreases in the direction of flow unless there is deceleration.\textsuperscript{361} Note that the mean pressure averaged over a full heartbeat still decreases with increasing distance from the aortic valve—by ~4 mmHg over the full length of the aorta—while the amplitude of the systolic/diastolic pressure oscillation nearly doubles.\textsuperscript{361} Measurements in dogs show that this peak pressure pulse amplification process continues until about the third generation of arterial branchings, after which both the oscillation and the mean pressure decrease gradually downstream along the arterial tree.\textsuperscript{361} These pulse profiles can probably be resolved with sufficient accuracy to allow a medical nanorobot equipped with appropriate pressure sensors to fix its position along all the main trunks of the arterial tree to within at least ±1 cm, and perhaps better.

Pressure and velocity waveforms change in different ways along the branches of the tree,\textsuperscript{924} in part due to reflection at bifurcations, which may allow bloodborne nanorobots to distinguish the organ in which they currently reside. For instance, the waveforms found in the renal and iliac arteries are quite distinct.\textsuperscript{924} Approximate position in the venous tree is readily inferred as well. For example, mean pressure in the venules is ~18 mmHg with fluctuations of amplitude ~5 mmHg; by the time bloodborne nanorobots reach the upper vena cava, near the right atrium of the heart, the mean pressure has fallen to ~2.2 mmHg with pressure fluctuations of amplitude ~5 mmHg.\textsuperscript{361}

Microvascular position can also be estimated from various local pressure measurements.\textsuperscript{929} For instance, Figure 8.30 shows a typical pressure-velocity distribution in arteriovenous microvessels 8-60 microns in diameter in the cat mesentery. Knowledge of mean local pressure and flow velocity identifies vessel diameter to within ±5 microns. Additionally, these pressure distribution profiles differ markedly from one tissue to another (Fig. 8.31). Comparison of the measured profile against a library of such profiles will at least narrow the possibilities, if not produce a unique tissue identification.

Local hematocrit (the volume fraction of red cells in whole blood, or Hct) varies as a function of microvessel diameter (Fig. 8.32). Simple omnidirectional echolocation (Section 4.8.2) measures the average distance to the nearest red blood cell over many samples—e.g., the mean center-to-center distance between red cells is ~9.8 microns at Hct 10%, ~6.8 microns at Hct 30%. An acoustic test pulse travels ~1.5 microns in 1 nanosec, hence Hct 10% and Hct 12.7% are distinguishable using one round-trip measurement, allowing a medical nanorobot to estimate the local microvessel diameter to within approximately ±10 microns. Given the scatter in the data, multiple measurements should improve the accuracy of the estimate.

Maps of mechanical properties such as strain fields in tissues can also provide useful navigational information but will be more difficult for individual nanorobots to employ directly. Magnetic resonance elastography (MRE) has already been used experimentally to map the shear modulus field of a porcine kidney using shear wave excitation at 200-400 Hz.\textsuperscript{928} Shear waves with displacements under 100 nm were readily observed, and it is believed the technique may be useful in mapping other viscoelastic parameters such as attenuation and dispersion, allowing more detailed tissue characterization. Elastographic parameters are highly variable in time as well as space. For example, the viscoelasticity of cervical mucus secretions swings widely over the course of the menstrual cycle, with a major feature occurring at mid-period near ovulation time,\textsuperscript{922} mirroring the temperature changes (Section 8.4.1.1).

### 8.4.3 Chemographic Navigation

The human body is a cauldron of chemical complexity. Of particular interest to us here are those extracellularly occurring chemical species reliably associated with specific locations, functions, or processes that might be useful in navigation. As trivial examples, high concentrations of hormones or neuropeptides are found in the endocrine glands or in the nerve tissues; lactic acid is produced in muscles driven anaerobically; uric acid is heavily concentrated in urine in the bladder (although high blood levels may indicate gout); capillary lymph is more watery than thoracic lymph, which in turn is far richer in histaminase than cervical duct lymph (Section 8.2.1.3).

As with thermal and pressure markers, crude chemical demarcation markers may be employed to direct medical nanorobots to particular locations. Such markers may include injected chemical plumes, time release implants, remote-triggered releasers (e.g., pressure-release multi-chemical-bearing tattoos), and tissue-coded markers (analogous to radiolabeled iodine concentrating in the thyroid gland, alkali metal ions seeking out bone, or cancer-cell-targeted light-sensitive drug molecules in photodynamic therapy)—all of which may establish useful localized chemical gradients. The principle difficulty for natural signal molecules is signal persistence, which may be relatively brief unless continuous broadcast and artificial (nonabsorbable) signal molecules are employed. Continuous broadcast creates a detection sphere of maximum radius $R_\text{max}$ around the broadcast emission point (Section 7.2.1.5). Use of artificial non-biodegradable molecules maximizes the signal/noise ratio. Diffusion causes small molecules to drift ~1 cell width in ~1 second, or ~1 mm in ~1000 seconds (Section 3.2 and Table 3.4); bulk flow microconvection and increased effective diffusivity due to natural stirring (Section 3.2.2) further limit positional precision.

Use of chemical beacons for chemonavigational triangulation is not efficient over large distances, since the required volume of signal...
molecules to the limit of nanodevice rotation during a Brownian motion-limited measurement time (Section 3.2.1). A pair of 10-probe sensors taking measurements every ~1 millisec (concentration differential $\Delta c / c \approx 20\%$ over the full nanomedically-relevant range of $K_d = 10^{-4}$ to $10^{-13}$ molecules/nm$^3$; Section 4.2.1) can resolve $\Delta \alpha \approx 3^\circ$ of arc (Eqn. 3.2), or two chemical point-emitters separated by $X_{path} \sin(\Delta \alpha) \approx 1$ micron at a range of $X_{path} \approx 20$ microns (~1 cell width) from the nanorobot.

Chemonavigation can also provide useful positional information. For example, glucose absorption in the jejunum establishes chemical gradients both in the radial and axial directions in the lumen and also across the surface of the small intestine (Section 8.2.3). If chyme leaving the duodenum starts with 1% glucose concentration and exits the upper jejunum at 0.01% concentration ~50 cm downstream, then the same 10-probe chemical sensor as above can measure 25 distinguishable glucose levels as a function of distance from the pylorus. This allows jejunal position to be estimated to within ~2 cm using glucose concentration alone.

Chemical navigation in the gut may incur significant positional uncertainties due to the large chemical variation of ingested foodstuffs, the generation and transport of gases, the volume and character of liquids consumed with the food, the state of mind (and bowel peristalsis) of the patient, the presence of biochemistry-altering disorders, and other factors. On the other hand, the results of measurements of nutrient levels in the stomach and duodenum may be broadcast downstream to jejunal nanorobots, allowing real-time assessment of the efficiency of digestive processes and thus reducing positional uncertainties. Similar gradients allow positional chemonavigation in gland ducts such as the salivary and pancreatic ducts.

Sulfate groups are more numerous on the surfaces of the rabbit aortic arch and carotid than on the surfaces of the rest of the systemic arteries, and post-staining fluorescence intensity (measuring abundance of sialyl groups in the endothelial glycocalyx) is 1.65 times higher in the carotid than in the aorta. In cyto chemonavigation is possible as well. For example, intracellular diffusion gradients of $O_2$ and ATP that
appear under some conditions could be useful for detecting the distribution and clustering of mitochondria.

Detailed chemomapping of individual patients may allow customization of standardized human chemographic profiles, e.g., to identify problem areas. Chemographicists may assist in the mapping process at up to -KHz sampling frequencies (Section 8.5.2.1; see also Chapter 19). For instance, regions of hypoxia develop in all solid tumors where rapidly dividing cells are supplied by inadequate or poorly developed vasculature. A continuously updated tissue oxygenation map allows recognition of growing hypoxic regions, permitting prompt detection of nascent tumors, myocardial and cerebral ischemia, cardiorespiratory failure, and rheumatoid arthritis (e.g., loss of fluids that provide joint oxygenation). Similarly, a whole-body map of the concentration of intracellularly sampled telomerase (or associated proteins such as TRF1/TRF2, or even telomerase mRNA levels) may provide a good first-cut toward a whole-body late-stage cancer map, since 85%-90% of all primary tumors display this biochemical marker. Cancer cells also display above-normal concentrations of β1 integrins, survivin, sialidase-sensitive cancer mucins and leptin receptors such as galectin-3, and below-normal concentrations of β2 integrins. Other examples include GM2 ganglioside, a glycolipid present on the surface of ~95% of melanoma cells, with the carbohydrate portion of the molecule conveniently jutting out on the extracellular side of the melanoma cell membrane. GM2 and another ganglioside, GD2, are expressed in several types of cancer cells including small-cell lung, colon, and gastric cancer, sarcoma, lymphoma, and neuroblastoma. Numerous specific markers of disease (an emerging field known as molecular epidemiology) are available for detection and mapping by nanorobots.

Another example is the chemographic mapping of organs. In the liver, hepatocytes located close to the portal vein (acinal Zone 1; Figure 8.26A) are bathed in blood that is comparatively rich in oxygen and nutrients, with only minimum exposure to metabolic waste products. In Zone 2, the blood is less fresh and wastes are more plentiful. In Zone 3, which extends to the central vein, the blood has little oxygen or nutrients but the highest concentration of metabolites. Metabolic and exocrine secretory activity creates distinct chemical and thermal gradients, clearly defining each lobule. This allows medical nanorobots to navigate between lobules by simply counting the number of lobules traversed, and to navigate within a chosen lobule by gradient averaging. Changes in measured gradients may evidence a recent ingestion of alcohol or other toxins, or may record the progression of hypoxia or cirrhosis which invariably begins in Zone 3.

Chemical localization to specific organs or tissues appears practical (Sections 8.2.5 and 8.5.2.2), as does chemical macrosensing of systemic states. For instance, a rise in serum adrenaline accompanied by surges in lactic acid production or oxygen consumption in the lower extremities might raise the reasonable inference that the patient is panicky and possibly fleeing from something that may be chasing him. Measurement of the concentration of thiocyanate in salivary secretions can distinguish smokers (~9 x 10^-3 gm/cm^3) from non-smokers (~2 x 10^-3 gm/cm^3). Measurement of telomere concentrations (a hormone secreted by fat cells) in the blood allows nanorobots to estimate the total amount of fat currently stored in the body, relative to the normal baseline of a given person.

Chemical tracking is also feasible. For example, a “bacterium moves through the water (or the liquid interior of your body) in a greasy cloud of its own waste products,” due to the excretion of metabolic byproducts through transmembrane pores. Such microbial exhaust plumes might plausibly be detectable for distances up to ~100-1000 microns downstream, depending upon bacterial emissions, local fluid flow rates and other factors (Section 7.2.1), possibly permitting tracking and interception by medical nanorobots. Interestingly, quorum sensing among bacteria enables some microbes to eavesdrop on the chemical communications of other species.

Developing embryos demonstrate a sophisticated “biological positioning system” by which each cell determines its location relative to other cells to eventually produce the appropriate tissue, organ, or nerve. For example, members of the “Wnt” and “Hedgehog” families of signaling proteins help cells (in developing limbs and organs) to distinguish up from down. As another example, neuron migrations within the developing embryo are thought to be guided by chemical signposts such as the proteins reelin (released near the destinations of migrating neurons), mDab1 (which may be a docking protein activated by reelin), and the homeodomain proteins DLX-1 and DLX-2. Neuroblast cells that divide in the subventricular zone of the lateral ventricle in the brain of adult mice migrate a distance of ~3-5 mm to the olfactory bulb; these migrating neuroblasts remain organized in chains until they reach the core of the olfactory bulb, after which they separate and migrate radially as individual cells to more peripheral layers, then halt and differentiate into neurons. During this process, the neural precursors are not guided by radial glial or axonal fibers, suggesting chemoanatomical guidance here too. Similarly, developing neurons shoot out axons whose tips are steered to their destinations by an array of guide molecules that are fixed on cell surfaces, or are located within the extracellular matrix, or are secreted by the axon’s targets—such as the netrins (a chemoattractant group of proteins) and the semaphorins such as collapsin (a chemorepellent protein). There is also a chemical “area code” that is used for cell localization during extravasation of leukocytes from blood vessels (Section 9.4.4.1). Medical nanorobots can be programmed to follow and to interpret any such positional chemical signposts normally used by motile cells.

Chemoanatomical may also be aided by the detection of isozymes—physically distinct forms displaying the same catalytic activity that may be present in different tissues of the same organism, in different cell types, or in different subcellular compartments in a human being. Isozymes are common in the sera and tissues of all vertebrates, insects, plants, and unicellular organisms. They are commonly used in clinical diagnosis, as, for instance, to distinguish normal serum and serum from patients with a myocardial infarct or with liver disease, using a lactate dehydrogenase (LDH) electrophrogram, although more recently the utility of isozymes in some clinical applications has been questioned.

8.4.4 Microbiotographies

While high-speed nanomedical prophylactic measures against pathogens will be readily available (e.g., Section 10.4 and Chapter 19), in some cases the physician may wish to obtain an infection map before beginning treatment. It may also be useful to prepare maps of particular symbiotic, commensal, or parasitic bacterial species,
or native mobile cellular elements, in the human body, whether for scientific or for diagnostic purposes.

For example, leukocytes and macrophages are readily recognized (Section 8.5.2.2) and could be statistically sampled to produce a crude whole-body leukographic map. A fleet of 2 trillion mapping nanorobots evenly distributed throughout a 0.1 m³ body volume gives each nanorobot a patrol volume of 50,000 microns³ (~1 capillary volume). At 310 K, a 1-micron nanorobot of normal density has an average thermal velocity of ~500 microns/sec (Eqn. 3.3). Assuming a cross-sectional area of ~1 micron², in 1 second a motile device traveling at this speed may trace out a nonrepeating zigzag path ~500 microns in length, or ~1% of its patrol volume, even while diffusing an additional radial distance of only ΔX ~ 1 micron (Eqn. 3.1). During its zigzag course, the device may collide at least once with ~1% of the motile body cells or microbiota present within the patrol volume. If -10 collisions are required to ensure a positive identification of a cellular coat, then ~0.1% of the entire target cell population present in the patient's body may be physically sampled, positively identified, affirmatively counted, and directly associated with a specific map voxel in ~1 sec—a sufficient statistical sample for most purposes, even given the likelihood of substantial double-counting. A whole-body white-cell-count map to ~1 cm³ voxel resolution requires ~10⁶ bits to describe and may be retrieved by the physician in ~5 sec using an in vivo mobile acoustic communication network; a ~1 mm³ resolution map (~10⁹ bits) fine enough to discover even the smallest macroscale infection site requires ~100-1000 seconds to outmessage in this manner. A sequence of readouts at regular intervals provides a clear strategic overview of the developing infection—some bacterial infections cause a selective increase in neutrophils, while infections with some protozoa and other parasites cause a selective increase in eosinophils.

Whole-body vascular (e.g., bloodstream only) surveys of red cells, white cells, platelets or pathogens will proceed ~10 times faster using a nanorobot fleet of similar number density, owing to the reduced search volume, though at the cost of slightly increased onboard computational complexity needed to compensate for the moving reference frame.

Discussion of possible protein coat marker modifications that bacteria or other pathogens might evolve in a nanomedical technology-rich treatment environment as a defensive tactic is deferred to Chapter 17.

8.5 Cytonavigation

Cytonavigation is the study of navigation in and around the individual cell. The topic is large and the present work can only scratch the surface. After a brief cytometric overview (Section 8.5.1), the discussion here addresses four important issues—distinguishing whether an encountered cell is self or nonself (Section 8.5.2.1), determining the exact cell type of an encountered cell (Section 8.5.2.2), the significant features and landmarks inside living cells, called cytophography (Section 8.5.3), and special considerations involved in navigating the cellular nucleus, called nucleography (Section 8.5.4). Issues relating to membrane penetration and cell entry by medical nanorobots are largely deferred to Section 9.4.5.

8.5.1 Cytometrics

How many cells are there in the human body? Serious estimates vary from a low of 10 trillion²¹,⁸⁶ to up to 50 trillion,²¹,⁶⁵ even 100 trillion.²¹,⁶⁵ The numbers show considerable spread because some authors refer only to tissue cells while others may refer to all cells physically present, both native and foreign.

The single most common native cell in the human body is the red blood cell (RBC, or erythrocyte). The average adult male carries ~5.2 × 10¹⁰ RBCs/cm³ (Appendix B) in his ~5400 cm³ blood volume (Table 8.1), giving ~28.1 trillion RBCs. The spleen has an additional ~70 cm³ of noncirculating RBC storage volume, giving another 0.4 trillion RBCs for a total of ~28.5 trillion RBCs in the body. Also confined to the bloodstream are a median ~2.5 × 10⁹ platelets/cm³, ~7.0 × 10⁹/cm³ leukocytes including neutrophils, eosinophils, basophils, lymphocytes and monocytes (Appendix B), plus another ~0.7 trillion platelet sequesters in a circulation-accessible pool in the spleen,²¹,²² giving ~2.1 trillion platelet cells in the human body and a white cell count of 0.4 trillion in the blood. There are also ~0.7 trillion lymphocytes in the lymphatic system (Table 8.5) and ~0.2 trillion macrophages and other reticuloendothelial (mono-nuclear phagocyte) cells throughout the human tissues.⁷⁴ Thus there are ~31.5 trillion native nontissue cells in the human body. There are also ~40 trillion foreign single-celled bacteria inhabiting the human colon, assuming ~35% of the ~340 cm³ colorectal contents are bacteria each averaging ~1.8 microns in diameter.

As for tissue cells, a 70 kg adult male reference body includes ~28.8 kg of body cell mass (BCM), defined as the amount of body tissue calculated to contain potassium at a concentration of 120 mEq (millimole)/kg.⁸¹ This figure is favored over the LBW (lean body mass) because BCM defines the component of body composition that contains the oxygen-containing, potassium-rich, glucose-oxidizing, work-performing tissue and contains all cellular elements concerned with respiration and chemical work, and mitotic activity.

Most human cells fall within a size range of 2-120 microns.⁸⁶ Platelets are ~2 microns, red cells ~3 microns x ~8 microns, neutrophils ~8-10 microns, lymphocytes ~6-12 microns in size, exocrine cells ~10 microns, fibroblasts 10-15 microns, osteocytes ~10-20 microns including processes, chondrocytes and liver cells ~20 microns, goblet and ciliated cells ~50 microns long and 5-10 microns wide, macrophages at ~20-80 microns, hematopoietic stem cells ~30-40 microns, and adipocytes filled with stored lipid are typically 70-120 microns in diameter (but may be up to five times larger in very obese people).⁸⁵-⁸⁶ Neurons vary enormously in size and shape, their bodies ranging from 4-120 microns in diameter with axonal processes varying between ~0.1-20 microns in diameter and ranging in length from a few microns up to ~1 meter;⁷⁹⁹ muscle cells (~30% of all tissue cells) also vary from 10-100 microns in diameter and may run up to ~50 cm in length. However, the “average” tissue cell, if such a thing may be said to exist, is probably ~20 microns in diameter,³¹³ weighing ~10 nanograms.

Since the reference human has a mean density of 1064 kg/m³ excluding bone minerals,⁸¹ a BCM of 28.8 kg makes ~0.027 m³ of cells; taking 8000 micron³ per cell gives a total of ~3.4 trillion tissue cells. Thus the human body contains ~35 trillion native cells, of which only ~10% are tissue cells, plus ~40 trillion foreign (mostly bacterial) cells in the colon for a grand total of 75 trillion cells. The number of tissue cells comprising each organ listed in Table 8.9 may be crudely estimated by multiplying tissue volume by the mean tissue cellular number density, or ~125 million cells/cm³.

8.5.2 Cytoidentification

Before entering and navigating a cell, a medical nanorobot must first determine whether the cell belongs to the patient or is a foreign cell (Section 8.5.2.1). The nanorobot must also determine which type of cell it has encountered, either to verify its location in a desired target organ or tissue, or to ensure that a targeted cell type has been found (Section 8.5.2.2). Fortunately, the plasma membrane
(the cell’s outermost coat; Section 8.5.3.2) is an open book for nanodevices equipped to read it.

8.5.2.1 Identification of Self

The cytochemical distinction between self and nonself is mediated by the histocompatibility molecules. These antigens are genetically coded in the major histocompatibility complex (MHC), a cluster of -28 genes located on chromosome 6 and comprising ~0.1% of the entire human genome. There are three broad classes (I, II, and III) of MHC genes that encode self antigens. The major function of the MHC-I and MHC-II molecules is the presentation of antigenic peptide fragments, derived from internal (MHC-I) or external (MHC-II) foreign proteins, to the immune system. It is the combination of an MHC molecule plus a foreign peptide fragment that appears antigenic to the immune system. (See Chapter 15.)

Most important are the MHC Class I molecules—glycoproteins encoded by several separate genetic loci which comprise the HLA (histocompatibility locus antigens, formerly “human leukocyte antigens”) complex. The three classical loci within the HLA complex are called HLA-A, HLA-B, and HLA-C. Essentially all adult nucleated cells in the human body (including white cells but excluding cells in the nervous system and cells of some tumors) display the classical MHC molecules of the A, B, and C regions on their surfaces in varying amounts.* Class I glycoproteins account for up to 1% of the total protein of the plasma membrane. Figure 8.33 shows the structure and orientation of a Class I MHC glycoprotein. Each MHC Class I molecule has two parts. The first part is a long folded glycosylated polypeptide chain of molecular weight -45,000 daltons (~340 residues). This chain has a short (~30 residue) hydrophilic tail inside the cell and a ~40 residue hydrophobic transmembrane segment. The extracellular portion has three ~90-residue segments designated α1, α2, and α3 with intrasegment disulfide bridges. Alloantigenic sites (carrying determinants specific to each individual) are located primarily in the α1 domain and to a lesser extent in the α2 domain, and there is a carbohydrate unit attached to the α2 domain. The α3 domain is relatively invariant. The second part of the MHC Class I molecule is a nonglycosylated 96-residue peptide called β2-microglobulin, of molecular weight ~12,000 daltons, which is noncovalently bound to the α3 domain of the longer glycosylated chain nearest the outer surface of the cell membrane. β2-microglobulin is not part of the active antigenic site of the HLA molecule but is essential for the expression of specificity.

MHC Class II molecules are also glycosylated integral membrane proteins (Fig. 8.34). These proteins are coded by at least six expressed genes located in the D region of the MHC cluster, each of which encodes either a ~34,000-dalton α chain or a ~28,000-dalton β chain. Each MHC Class II polypeptide consists of two external 90-residue domains, a 30-residue transmembrane domain, and a cytoplasmic domain with 10-15 amino acids. There are four serologically-determined Class II subtypes: HLA-DR and HLA-D (one α and several β chains), HLA-DQ (one α and one β chain), and HLA-DP (at least one α and one β chain). Each Class II glycoprotein molecule consists of at least two of these transmembrane polypeptide chains bound together noncovalently. Only the shorter β chains contain the alloantigenic sites, and both chains have carbohydrate units. MHC Class II molecules are less widely distributed than those of Class I, being found primarily on dendritic cells, glial cells in the brain, some epithelial and endothelial cells (e.g., Langerhans cells in the epidermis), monocytes, macrophages (including tissue macrophages such as the Kupffer cells of the liver), melanoma cells, activated (but not resting) T cells, and most B lymphocytes. Other cells may be induced to express MHC Class II proteins when exposed to γ interferon. Most parenchymal cells do not express MHC Class II molecules.

* Classical HLA tissue typing relied solely upon observations of cytotoxic reactions of human alloantisera with live lymphocytes. By 1998 it was apparent that this approach fails to distinguish all HLA alleles, and typing based on nucleotide sequences had begun to replace serological methods.
MHC Class III genes encode the ~20 proteins that comprise the complement system, including Bf, C2, C4A and C4B. These are functionally related in that they are each involved in the activation of the C3 component (Chapter 15). They allow human cell membranes to be distinguished from nonhuman (e.g., bacterial) membranes, a crude form of self identification, but are nonspecific and thus play no role in distinguishing the cells of one patient from those of another.

The >3.8 megabase gene locus of the MHC is the most polymorphic (having alternative types) in the human genome; the number of observed MHC (HLA) types appears to be near the optimum number based on evolutionary simulation experiments using a cellular automaton model of antigen-lymphocyte interactions. The HLA-uniqueness of an individual patient can only be estimated because the specificities do not occur with equal probability. For example, 16% of the human population has HLA-A1 but only 10% have HLA-B8. As an additional complication, the specificities are not strictly independent—given the above probabilities, the combination A1B8 should occur with a frequency of 1.6%, but the actual observed frequency is 8.8%. Members of the same genealogical family are also far more likely to have matching MHC Class II proteins than randomly selected unrelated individuals. Subject to these provisos, the existence of ~500 million distinct classical HLA combinations originally seemed to imply that only ~10 people in a worldwide population of 5 billion people shared exactly the same self-molecules, making each person literally “one in a billion.” However, based on nucleotide sequencing, as of 1998 there were estimated to be >10^{16} known combinations, making each person quite histocompatibility-unique.

MHC Class I and Class II self-molecules are manufactured in the rough endoplasmic reticulum (Section 8.5.3.5) and are then delivered to the plasma membrane (Section 8.5.3.2) for presentation to the extracellular environment. A typical cell may have 15,000-30,000 Class I molecules (~10-20/micron^2) at its surface; a

### Table 8.13. Recognized HLA Specificities

#### A. Specificities Based on Classical Serological Alloantisera Typing

<table>
<thead>
<tr>
<th>Class I Proteins</th>
<th>Class II Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HLA-A</strong></td>
<td><strong>HLA-B</strong></td>
</tr>
<tr>
<td>A1</td>
<td>A28</td>
</tr>
<tr>
<td>A2</td>
<td>A29</td>
</tr>
<tr>
<td>A203</td>
<td>A30</td>
</tr>
<tr>
<td>A210</td>
<td>A31</td>
</tr>
<tr>
<td>A3</td>
<td>A32</td>
</tr>
<tr>
<td>A9</td>
<td>A33</td>
</tr>
<tr>
<td>A10</td>
<td>A34</td>
</tr>
<tr>
<td>A11</td>
<td>A36</td>
</tr>
<tr>
<td>A19</td>
<td>A43</td>
</tr>
<tr>
<td>A23</td>
<td>A66</td>
</tr>
<tr>
<td>A24</td>
<td>A68</td>
</tr>
<tr>
<td>A25</td>
<td>A69</td>
</tr>
<tr>
<td>A26</td>
<td>A74</td>
</tr>
<tr>
<td>B35</td>
<td>B50</td>
</tr>
<tr>
<td>B37</td>
<td>B51</td>
</tr>
</tbody>
</table>

#### B. Specificities Based on Modern Nucleotide Sequence Typing

<table>
<thead>
<tr>
<th>Locus</th>
<th>Number of Alleles</th>
<th>Locus</th>
<th>Number of Alleles</th>
<th>Locus</th>
<th>Number of Alleles</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-A</td>
<td>67</td>
<td>HLA-F</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-B</td>
<td>149</td>
<td>HLA-G</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-C</td>
<td>39</td>
<td>HLA-DRA</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DRE</td>
<td>5</td>
<td>HLA-DPA</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DRB</td>
<td>179</td>
<td>HLA-DBP</td>
<td>69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQA</td>
<td>18</td>
<td>HLA-DMA</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HLA-DQB</td>
<td>29</td>
<td>HLA-DMB</td>
<td>5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total: 266
Total: 314
cell expressing Class II proteins may display 50,000–500,000 Class II molecules (~40–400/micron²) at its surface. A nanorobot chemotactic sensor pad (with appropriate reversible binding sites; Section 4.2.8) measuring 50–300 nm in diameter pressed against a cell surface will overlay at least one of each class of self-molecule. Vesicle-transported proteins (like MHC proteins) have an estimated lateral diffusion coefficient in the cell plasma membrane of ~2 x 10⁻¹⁴ m²/sec, so an MHC molecule (assumed diameter ~10 nm) experiences an effective absolute viscosity of η = 2.3 kg/m-sec (Eqn. 3.5) and thus migrates across the plasma membrane surface under the sensor pad from center to periphery of the sensor pad in 20–60 millisecc (vs. 0.01-0.5 millisec for phospholipid molecules in the membrane; Eqn. 3.1), allowing a single binding event and hence a detection in much less than a second.

Consider a chemotactic sensor pad of area Aₚad overlaying Nₐntigen antigen molecules, comprising an array of Nₗntogen chemosensors each of area Aₗntogen. If a single detection event requires t = 3 / η Aₗntogen Rₗntogen Aₘₐₓ² / kT seconds (Eqn. 3.1) where Aₘₐₓ is the mean antigen migration distance in the cell membrane, then the time needed to measure Nₗntogen specificities assuming a minimum of Nₐntogen contact events to ensure a binding event is tₘₐₓ = t Nₐntogen Nₗntogen / Nₐntogen (Eqn. 8.5) where N₃ = Aₚad / Aₗntogen and Aₘₐₓ² = Aₗntogen Nₗntogen / Nₐntogen, then:

\[
tₘₐₓ = \frac{3 \pi \eta Nₐntogen Rₐntogen N_encounter Nₗntogen² Aₗntogen}{Aₚad Nₐntogen kT}
\]

(Eqn. 8.5)

Taking Rₐntogen = 5 nm, Nₐntogen = 10, Aₗntogen = 100 nm², Aₚad = 90,000 nm², Nₐntogen = 1, and T = 310 K, then tₘₐₓ ~ (3 x 10⁻⁵) Nₗntogen. A nanorobot searching for a particular set of 7 HLA proteins (Nₗntogen = 7) thus requires tₘₐₓ ~ 0.001 sec to make the self/nonself determination for a particular cell membrane it has encountered. A nanorobot seeking to determine the HLA type of the membrane (e.g., in mapping mode) must in the worst case search all 254 known blood antigen types (Nₗntogen = 254), requiring at most tₘₐₓ ~ 2 sec.

Direct detection of blood group antigens, or of antibodies to blood group antigens in body fluids, permits at least partial self-recognition by bloodborne nanorobots without the need for any direct cell contact, which may be useful in establishing theater protocols (Chapter 12). ABH, Lewis, I and P blood group antigens are found in blood plasma, and serum IgM-class antibodies associated with the carbohydrate antigens of the ABO, Lewis, and P blood group systems are almost universal. Anti-M and anti-N are common, anti-Sd⁺ is found in 1-2% of normal people, and anti-Vw or anti-Wra is found in ~1% of patients. In persons who previously have been

\[
\text{Fig. 8.35. Structure of ABO blood system red cell surface carbohydrate antigens (modified from Cunningham).}
\]

<table>
<thead>
<tr>
<th>Conventional Name</th>
<th>Total # of Antigens</th>
<th>Common Antigens</th>
<th>Common Phenotypes</th>
<th>% Frequency</th>
<th># of Antigen Sites/RBC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blood group systems:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABO</td>
<td>4</td>
<td>A</td>
<td>A</td>
<td>40%</td>
<td>ABO: 1-2 x 10^6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>B</td>
<td>B</td>
<td>11%</td>
<td>H: 1.7 x 10^6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB</td>
<td></td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>Rh</td>
<td>47</td>
<td>D</td>
<td>D, DcE</td>
<td>42%</td>
<td>D: 1 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C, C_w</td>
<td>r, dce</td>
<td>37%</td>
<td>R_r, R_s: 15-33 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>E, e</td>
<td>r, D_e</td>
<td>14%</td>
<td>R_r, R_s: 14-19 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>f(ce)</td>
<td>r^* dce</td>
<td>4%</td>
<td>R_r: 12-20 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V(ce^*)</td>
<td>r^* dce</td>
<td>2%</td>
<td>R_r: 9-14 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HO</td>
<td>45%</td>
<td>cc: 70-85 x 10^3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AB</td>
<td>4%</td>
<td>Cc: 37-53 x 10^3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ee: 18-24 x 10^3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ee: 13-14 x 10^3</td>
<td></td>
</tr>
<tr>
<td>MNSs</td>
<td>37</td>
<td>M</td>
<td>M+N-</td>
<td>28%</td>
<td>MN: 5 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>M+N+</td>
<td>50%</td>
<td>Ss: 2.5 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S</td>
<td>M-N+</td>
<td>22%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>s</td>
<td>S+s-</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>S+s+</td>
<td>44%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S+s-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S-s-U</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>1</td>
<td>P_1</td>
<td>P^2+P_1+P_2+(P_1^3)</td>
<td>79%</td>
<td>more rare</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>P^2+P_1+P_2+(P_1^3)</td>
<td>79%</td>
<td>more rare</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P_2</td>
<td>P^2+P_1+P_2+(P_1^3)</td>
<td>21%</td>
<td>more rare</td>
</tr>
<tr>
<td></td>
<td></td>
<td>P_3</td>
<td>P^2+P_1+P_2+(P_1^3)</td>
<td>21%</td>
<td>more rare</td>
</tr>
<tr>
<td>Lewis</td>
<td>3</td>
<td>Le^a</td>
<td>Le(a+b-)</td>
<td>22%</td>
<td>Le^a: 3 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Le^b</td>
<td>Le(a+b+)</td>
<td>72%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Le(b+)</td>
<td>Le(a-b+)</td>
<td>6%</td>
<td></td>
</tr>
<tr>
<td>Kell</td>
<td>21</td>
<td>K</td>
<td>K+k+</td>
<td>91%</td>
<td>2-6 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>k</td>
<td>K+k-</td>
<td>8.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kp^+</td>
<td>K+k-</td>
<td>0.2%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kp^b</td>
<td>Kp(a-b+)</td>
<td>97.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Js^a</td>
<td>Kp(a+b+)</td>
<td>2.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Js^b</td>
<td>Kp(a+b-)</td>
<td>rare</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Js(a+b-)</td>
<td>100%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Js(a+b+)</td>
<td>rare</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Duffy</td>
<td>6</td>
<td>Fy^a</td>
<td>Fy(a+b)-Fy3+</td>
<td>17%</td>
<td>Fy^a: 7-13 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fy^b</td>
<td>Fy(a+b)-Fy3+</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fy^3</td>
<td>Fy(a+b)-Fy3+</td>
<td>34%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fy^4</td>
<td>Fy(a+b)-Fy3-4+</td>
<td>rare</td>
<td></td>
</tr>
<tr>
<td>Kidd</td>
<td>3</td>
<td>Jk^a</td>
<td>Jk(a+b)-Jk3+</td>
<td>28%</td>
<td>Jk^aJk^b: 14 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jk^b</td>
<td>Jk(a+b)-Jk3+</td>
<td>49%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Jk^3</td>
<td>Jk(a+b)-Jk3-</td>
<td>23%</td>
<td></td>
</tr>
<tr>
<td>Lutheran</td>
<td>18</td>
<td>Lu^a</td>
<td>Lu(a+b)-Lu3+</td>
<td>0.15%</td>
<td>Lu^a: 0.8-4 x 10^3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lu^b</td>
<td>Lu(a+b)-Lu3+</td>
<td>7.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lu^3</td>
<td>Lu(a+b)-Lu3+</td>
<td>92.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lu(a+b)-Lu3-</td>
<td>very rare</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diego</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cartwright</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xg</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sclanna</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dombrock</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coulton</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LW</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chido/Rogers</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kx</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gerbich</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cromer</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knops</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
pregnant or transfused, 0.16-0.56% have anti-D (Rh group) and 0.14-0.60% have anti-E and anti-C (Rh system), anti-K and anti-Fya, and several other antibodies in serum.960 Body secretions contain ABH, I and Lewis antigen but no P system antigens; Sd* antigen is found in most body secretions, with the greatest concentration in the urine.960

Other partial exceptions to the MHC system are the keratinocytes, smooth muscle cells, and fibroblasts, which do not constitutively express HLA Class II molecules.967 As a result, foreign fibroblasts cannot induce the generation of required helper T cells and thus stimulate no rejection response when transplanted between hosts.

8.5.2.2 Identification of Cell Type

How many different cell types are there in an adult human body? The usual estimate based on histological studies is that there are ~200 distinct kinds of cells that show alternate structures and functions.312,531,866 These represent discrete categories of cell types of markedly different character, not arbitrary subdivisions along a morphological continuum. Traditional classification is based on microscopic shape and structure, and on crude chemical nature (e.g., affinity for various stains), but newer immunological techniques have revealed, for instance, that there are more than 10 distinct types of lymphocytes. Pharmacological and physiological tests have revealed many different varieties of smooth muscle cells—for example, uterine wall smooth muscle cells are highly sensitive to estrogen and (in late pregnancy) oxytocin, while gut wall smooth muscle cells are not. Appendix C presents a catalog of the cells of the human body, slightly modified from an original compilation by Alberts et al.531 The catalog is organized by cellular function and omits subdivisions of smooth muscle cells, neuron classes in the CNS, various related connective tissue and fibroblast types, and intermediate stages of maturing cells such as keratinocytes (only the stem cell and differentiated cell types are given). Otherwise, the catalog is said to represent an exhaustive listing of the ~219 cell varieties found in the adult human phenotype. (Complexity theory and phylogenetic comparisons suggest that the maximum number of cell types \(N_{cell} \sim N_{gene}^{1/2} = 370\) cell types for humans with \(N_{gene} \sim 10^5\) genes).1266,766

Medical nanorobots can probably distinguish all of these cell types by surface chemical assay using chemosensor pads (Section 4.2.8). Antigenic specificities exist for species (xenotype), organ, tissue or cell type for almost all cells—possibly involving as many as ~10^4 distinct antigens. A comprehensive analysis of the cytoimmunography of all cell types is beyond the scope of this book, but a few examples of cell type-specific antigenic markers can be given to illustrate the tremendous power of this approach.

In the case of red blood cells, antigens in the Rh, Kell, Duffy, and Kidd blood group systems are found exclusively on the plasma membranes of erythrocytes and have not been detected on platelets, lymphocytes, granulocytes, in plasma, or in other body secretions such as saliva, milk, or amniotic fluid.960 Thus detection of any member of this four-antigen set establishes a unique marker for red cell identification. MNSs and Lutheran antigens are also limited to erythrocytes with two exceptions: GPA glycoprotein (MN activity) also found on renal capillary endothelium,961 and Lu^b-like glycoprotein which appears on kidney endothelial cells and liver hepatocytes.962 In contrast, ABH antigens are found on many non-RBC tissue cells such as kidney and salivary glands.955 In young embryos ABH can be found on all endothelial and epithelial cells except those of the central nervous system.963 ABH, Lewis, I and P blood group antigens are found on platelets and lymphocytes, at least in part due to adsorption from the plasma onto the cell membrane. Granulocytes have I antigen but no ABH.960

Platelets also express platelet-specific alloantigens on their plasma membranes, in addition to the HLA antigens they already share with body tissue cells. Currently there are five recognized human platelet alloantigen (HPA) systems that have been defined at the molecular level (Table 8.15). The phenotype frequencies given are for the Caucasian population; frequencies in African and Asian populations may vary substantially. For instance, HPA-1b is expressed on the platelets of 28% of Caucasians but only 4% of the Japanese population.964 A medical nanorobot that detects a
Table 8.15. Complete Listing of Recognized Human Platelet Alloantigen (HPA) Systems

<table>
<thead>
<tr>
<th>Alloantigen System</th>
<th>Allelic Forms</th>
<th>Former Names</th>
<th>Phenotype Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPA-1</td>
<td>HPA-1a</td>
<td>P^a_1 or Zw^a</td>
<td>a/a = 72%</td>
</tr>
<tr>
<td></td>
<td>HPA-1b</td>
<td>P^a_2 or Zw^a</td>
<td>a/b = 26%</td>
</tr>
<tr>
<td></td>
<td>HPA-2a</td>
<td>Ko^a</td>
<td>b/b = 2%</td>
</tr>
<tr>
<td></td>
<td>HPA-2b</td>
<td>Ko^a</td>
<td>a/a = 85%</td>
</tr>
<tr>
<td></td>
<td>HPA-3a</td>
<td>Bak^a or Lek^a</td>
<td>a/b = 46%</td>
</tr>
<tr>
<td></td>
<td>HPA-3b</td>
<td>Bak^b or Lek^b</td>
<td>b/b = 1%</td>
</tr>
<tr>
<td></td>
<td>HPA-4a</td>
<td>Pen^a or Yuk^a</td>
<td>a/a = 99%</td>
</tr>
<tr>
<td></td>
<td>HPA-4b</td>
<td>Pen^b or Yuk^b</td>
<td>a/b &lt; 0.1%</td>
</tr>
<tr>
<td></td>
<td>HPA-5a</td>
<td>Bi^a or Zav^a</td>
<td>a/a = 80%</td>
</tr>
<tr>
<td></td>
<td>HPA-5b</td>
<td>Br^a, Zav^a, or Hc^a</td>
<td>a/b = 19%</td>
</tr>
</tbody>
</table>

There are neutrophil-specific antigens and various receptor-specific immunoglobulin binding specificities for leukocytes. For instance, monocyte FcRI receptors display the measured binding specificity IgG1^+^IgG2^+^IgG3^+^IgG4^−^, monocyte FcRIII receptors have IgG1^+^IgG2^+^IgG3^+^IgG4^−^, and FcRII receptors on neutrophils and eosinophils show IgG1^+^IgG2^+^IgG3^+^IgG4^−^. Neutrophils also have β-glucan receptors on their surfaces.

Tissue cells display specific sets of distinguishing markers on their surfaces as well. Thyroid microsomal-microvillous antigen is unique to the thyroid gland. Glial fibrillary acidic protein (GFAP) is an immunocytochemical marker of astrocytes, and syntaxin 1A and 1B are phosphoproteins found only in the plasma membrane of neuronal cells.

Alpha-fodrin is an organ-specific autoantigenic marker of astrocytes,947 and syntaxin 1A and 1B are phosphoproteins found only in the plasma membrane of neuronal cells.

Alloantigen Allelic Former Phenotype

<table>
<thead>
<tr>
<th>System</th>
<th>Allele</th>
<th>Name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPA-1</td>
<td>a/a</td>
<td>P^a_1</td>
<td>72%</td>
</tr>
<tr>
<td>HPA-2</td>
<td>a/b</td>
<td>Ko^a</td>
<td>26%</td>
</tr>
<tr>
<td>HPA-3</td>
<td>b/b</td>
<td>Bak^a</td>
<td>1%</td>
</tr>
<tr>
<td>HPA-4</td>
<td>a/a</td>
<td>Pen^a</td>
<td>99%</td>
</tr>
<tr>
<td>HPA-5</td>
<td>a/b</td>
<td>Br^a</td>
<td>19%</td>
</tr>
</tbody>
</table>

Table 8.16. Unique and Shared Antigenic Surface Markers of Hepatopoietic and Hemopoietic Human Cells

<table>
<thead>
<tr>
<th>Hepatopoietic Cells (e.g., Hepatoblasts)</th>
<th>Hemopoietic Cells (e.g., Erythroid Progenitors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNIQUE Antigenic Markers:</td>
<td>OX63 (MCA 276), OX44 (MCA 371, CD 37)</td>
</tr>
<tr>
<td>α-fetoprotein, albumin, stem cell factor, hepatic</td>
<td>OX42 (MCA 275, CD118)</td>
</tr>
<tr>
<td>heparan sulfate-PGs (syndecans/perlecans)</td>
<td>c-KIT, stem cell factor</td>
</tr>
<tr>
<td>IGf I, IGf II, TGF-α, TGF-β receptor, αi integrin, αi integrin, connexin 26 and connexin 32</td>
<td>SHARED all known oval cell antigens, transferrins, γ glutamyl transpeptidase (GdT), glucuronol transferases (some isoforms), glutathione-S-transferases, ligandins, cEBP isoforms, IL-1 receptor, IL-6 receptor, interferon receptor, HGF receptor, insulin receptor, MDR1 and MDR2 (multidrug resistance genes), and connexin 43</td>
</tr>
</tbody>
</table>
The cadherin molecular family of 723-748-residue transmembrane proteins provides yet another avenue of cell-cell adhesion that is cell-specific. Cadherins are linked to the cytoskeleton. The classical cadherins include E- (epithelial), N- (neural or A-CAM), and P- (placental) cadherin, but in 1998 at least 12 different members of the family were known. They are concentrated (though not exclusively found) at cell-cell junctions on the cell surface and appear to be crucial for maintaining multicellular architecture. Cells adhere preferentially to other cells that express the identical cadherin type. Liver hepatocytes express only E-; mesenchymal lung cells, optic axons and neuroepithelial cells express only N-; epithelial lung cells express both E- and P-cadherins. Members of the cadherin family also are distributed in different spatiotemporal patterns in embryos, with the expression of cadherin types changing dynamically as the cells differentiate.

Carbohydrates are crucial in cell recognition. All cells have a thin sugar coating (the glyocalyx; Section 8.5.3.2) consisting of glycoproteins and glycolipids, of which ~3000 different motifs had been identified by 1998. The repertoire of carbohydrate cell surface structures changes characteristic as the cell develops, differentiates, or sickens. For example:

1. The array of carbohydrates on cancer cells is strikingly different from that on normal cells.

2. A unique trisaccharide (SSEA-1 or Le\(^a\)) appears on the surfaces of cells of the developing embryo exactly at the 8- to 16-cell stage when the embryo compacts from a group of loose cells into a smooth ball.

3. Bacterial pathogens often use cell-specific sugars to guide them to their preferred targets—*E. coli* are abundant in tissues surrounding the ureters leading from kidneys to bladder but are rarely found in the upper respiratory tract; group A *Streptococci*, which colonize only the upper respiratory tract and skin, rarely cause urinary tract infections; the gonorrhea organism *Neisseria gonorrhoeae* adheres only to cells of the genital and oral epithelia but not to cells of other organs. These bacterial carbohydrate specificities (e.g., carbohydrate-specific bacterial adhesins, lectins, and glycoconjugates) have been at least partially catalogued and molecular dynamics studies have begun.

Carbohydrate motifs are in theory more combinatorically diverse than nucleotide- or protein-based structures. While nucleotides and amino acids can interconnect in only one way, the monosaccharide units in oligosaccharides and polysaccharides can attach at multiple points. Thus two amino acids can make only two distinct dipeptides, but two identical monosaccharides can bond to form 11 different disaccharides because each monosaccharide has 6 carbons, giving each unit 6 different attachment points for a total of 6 + 5 = 11 possible combinations. Four different nucleotides can make only 24 distinct tetranucleotides, but four different monosaccharides can make 35,560 unique tetrascarbohydrates, including many with branching structures. A single hexasaccharide can make ~10\(^2\) distinct structures, vs. only 6.4 x 10\(^5\) structures for a heptapeptide; a 9-mer carbohydrate has a mole of isomers.

The coded sugar coating of red cells has already been described (Fig. 8.35). As an example, the CD44 family of transmembrane glycoproteins are 80-95 kilodalton cell adhesion receptors that mediate ECM binding, cell migration and lymphocyte homing. CD44 antigen shows a wide variety of cell-specific and tissue-specific glycosylation patterns, with each cell type decorating the CD44 core protein with its own unique array of carbohydrate structures. Distinct CD44 cell surface molecules have been found in lymphocytes, macrophages, fibroblasts, epithelial cells, and keratinocytes. CD44 expression in the nervous system is restricted to the white matter (including astrocytes and glial cells) in healthy young people, but appears in gray matter accompanying age or disease. A few tissues are CD44 negative, including liver hepatocytes, kidney tubular epithelium, cardiac muscle, the testes, and portions of the skin.

The selectin family of ~50 kilodalton cell adhesion receptor glycoprotein molecules can recognize diverse cell-surface antigen carbohydrates and help localize leukocytes to regions of inflammation (leukocyte trafficking). Selectins are not attached to the cytoskeleton. Leukocytes display L-selectin, platelets display P-selectin, and endothelial cells display E-selectin (as well as L and P) receptors. Cell-specific molecules recognized by selectins include tumor mucin oligosaccharides (recognized by L, P, and E), brain glycolipids (P and L), neutrophil glycoproteins (E and P), leukocyte sialoglycoproteins (E and P), and endothelial proteoglycans (P and L). The related MEL-14 glycoprotein homing receptor family allows lymphocyte homing to specific lymphatic tissues coded with "vascular addressin"—cell-specific surface antigens found on cells in the intestinal Peyers' patches, the mesenteric lymph nodes, lung-associated lymph nodes, synovial cells and lactating breast endothelium. Homing receptors also allow some lymphocytes to distinguish between colon and jejunum and select-related interactions, along with chemotactic agents and receptors and with integrin-Ig, regulate leukocyte extravasation (Section 9.4.4.1) in series, establishing a three-digit "area code" for cell localization in the body.

Viral capsid proteins are readily recognized, permitting identification of many virus species. Enveloped viruses, which have acquired a lipid membrane coating borrowed from the host cell upon release, will be more difficult to detect by simple surface marker sensing. A lipid-penetrating sensor or a host-cell "decoy" sensor (activating membrane fusion protein release (Section 9.4.5.4) by the virus, thus revealing its true identity) may be more useful.

Finally, cells may be typed according to their indigenous transmembrane cytoskeleton-related proteins. For example, erythrocyte membranes contain glycoporphin C (~25 kilodaltons, ~3000 molecules/micron\(^2\)) and band 3 ion exchanger (90-100 kilodaltons, ~10,000 molecules/micron\(^2\)). Platelet membranes incorporate the GP Ib-IX glycoprotein complex (186 kilodaltons); cell membrane extensions in neutrophils require the transmembrane protein ponticulin (17 kilodaltons); and striated muscle cell membranes contain a specific laminin-binding glycoprotein (156 kilodaltons) at the outermost part of the transmembrane dystrophin-glycoprotein complex. There are also a variety of carbohydrate-binding proteins (lectins) that appear frequently on cell surfaces, and can distinguish different monosaccharides and oligosaccharides. Cell-specific lectins include the galactose (asialoglycoprotein)-binding and fucose-binding lectins of hepatocytes, the mannosyl-6-phosphate (M6P) lectin of fibroblasts, the mannosyl-N-acetylglucosamine-binding lectin of alveolar macrophages, the galabiose-binding lectins of uroepithelial cells, and several galactose-binding lectins in heart, brain and lung.

Each cell expresses a different set of genes from the genome, and each gene normally represents a different protein (although in some cases alternate splicing can generate several proteins from the same gene; Chapter 20). Thus it seems plausible to conclude that each different cell type uses a unique constellation of proteins in its construction which can be detected as a set by nanorobot chemo sensors, thus unambiguously identifying the cell type.
(Of course, possibly up to ~50-70% of the proteins in different cells may be the same, serving common "housekeeping" functions.) If there are ~400 cell types (including all varieties of smooth muscle cells, neuron cells, and other cytchemically distinct cell subcategories not fully enumerated in Appendix C), then every cell type in the catalog may be uniquely identified using as few as \(\log_2(400) \approx 9\) binary antigenic markers. Given that Nature has employed ~580 markers among the HLA specificities and ~254 markers in the blood group systems to distinguish self from nonself, it is clear that biological systems may employ considerable multifunctionality and redundancy. By 1998 a complete catalog of all cell-specific antigens had not yet been compiled. However, the completion of sequencing of ~90% of the human genome by the spring of 2000\(^{3186}\) (with the complete version finished by 2002-2003\(^{3186,3187}\)), and foreseeable advances in DNA chip technology (used to sample gene expression patterns), should allow the rapid determination of unique sets of cell-specific antigens by circa 2002-2005.

In the worst case, 9 unique antigenic markers would be required for each of at most ~400 cell types, giving a maximum of 3600 antigenic markers needed to positively identify all cell types. From Eqn. 8.5, a \(-300\text{ nm}\)^2 chemotactic pad with \(N_{\text{spec}} = 3600\) implies a maximum cell typing measurement time \(t_{\text{meas}} \approx 400\text{ sec}\) in mapping mode. If a more plausible working set of ~300 key antigenic markers will suffice, then \(t_{\text{meas}} \approx 3\text{ sec}\) in mapping mode. On the other hand, a nanorobot searching for a particular set of \(N_{\text{spec}} = 9\) binary antigenic markers (e.g., seeking only liver cells, rejecting all nonliver cells) requires just \(t_{\text{meas}} \approx 0.002\text{ sec}\) to make the cell-type determination.

8.5.3 Cytography

Cytophory is concerned with the "geography" of the cell and directing navigation through the intracellular spaces. Discussion of the nucleus, the largest and most important cellular organelle, is deferred to Section 8.5.4. Methods of cell entry are described in Section 9.4.5.

8.5.3.1 Overall Cellular Structure

As recently as the late 1970s, the cell was often incorrectly described as a water-filled membranous sac enclosing a loosely structured population of discrete organelles. In reality, the interior of a cell is extremely compact, only a few times more open than a hydrated protein crystal.\(^{941}\) The cell interior is crisscrossed by many tens of thousands of cytoskeletal filaments of various gauges, attaching organelles to the nucleus, the plasma membrane, the ECM, and to each other.\(^{942}\)

Cells also have various shapes. Most cells are surrounded by and are anchored to neighboring cells. Such immobilized cells usually assume a polyhedral shape. Specialized cells adopt shapes related to the specific functions they perform. A dramatic example is the nerve cell, which has one or more cylindrical processes extending from the cell like the branches of a tree, which processes allow the cell to receive and to transmit electrical signals.

From the nanomedical perspective, the cell (Fig. 8.36) may be regarded as a large machine constructed of many smaller machines.\(^{182}\) These smaller machines are the organelles. Organelles are typically 0.5-3 microns in diameter, roughly the same size as the largest bloodborne medical nanorobots. Table 8.17 is a very approximate quantification of the known classes of cellular organelles and other major cellular components. Cytoskeletal number densities are crude but self-consistent estimates. Actual data may differ widely from the values given in the table depending upon:

1. size, age, and type of cell;
2. global and local respiration, nutrition and energy demand;
3. tissue and organ location;
4. environmental factors such as temperature, pressure, salinity, ECM activity, and extracytosolic toxicology; and
5. cellular secretory and mitotic status.

Organelles represent the principal metabolic and structural machinery of the cell. Each organelle type or cellular component is engineered to carry out specific functions while maintaining a specific structure, as described briefly below.

The small size of most organelles makes intra-organelle locomotion by whole nanorobots difficult or impossible, with the possible exception of the nucleus (Sections 8.5.4 and 9.4.6). This does not
256

Nanomedicine • Volume I

Table 8.17. Approximate Quantification of the Components of a Typical 20-µm Human Tissue Cell531,997

Cellular Component
or Organelle
Glycocalyx
(external to cell)
Plasma membrane
(~900 kg/m3)
Ribosomes
Rough endoplasmic
reticulum
Cisternae
Membrane
Smooth endoplasmic
reticulum
Cisternae
Membrane
Golgi complex
(~1090 kg/m3)
Cisternae
Membrane
Golgi vesicles
(~1120 kg/m3)
Secretory vesicles
(~1120 kg/m3)
Glycogen granules
(1560 kg/m3)
Lipid droplets
(850 kg/m3)
Vaults
Lysosomes
(1120 kg/m3)
Body
Membrane
Proteasomes
Peroxisomes
(1250 kg/m3)
Body
Membrane
Mitochondria
(1190 kg/m3)
Outer membrane
Inner membrane
Cytoskeleton, total
Superfine filaments
Microfilaments
Intermed. filaments
Thick filaments
Microtubules
Centrioles

Typical
Diameter

Typical
Length or
Thickness

Typical
Surface
Area

Typical
Volume

Typical Total Item
Number
Volume
per Cell
in Cell

~20µm

10-50nm

2400µm2

12-120µm3

1

12-120µm3

20µm
25nm

6-10nm
15-30nm

2400µm2
0.002µm2

14-24µm3
8000nm3

1
~107

~20µm3
80µm3

~µm
---

--5-6nm

--70,000µm2

1120µm3
720µm3
400µm3

1
---

~µm
---

--5-6nm

--32,000µm2

420µm3
240µm3
180µm3

1
--3

Total Item Total Item
Total Item
Volume
Surface
Surface
as a %
Area in Cell Area as a %

2400µm2

0.5%

0.3%
1.0%

2400µm2
20,000µm2

0.5%
4.3%

1120µm3
720µm3
400µm3

14.0%

--70,000µm2

14.9%

420µm3
240µm3
180µm3

5.3%

--32,000µm2

6.8%

3

--14,000µm2

3.0%

~µm
---

--avg. 8nm

--14,000µm2

410-510µm
300-400µm3
110µm3

1-300 stks
---

450µm
350µm3
110µm3

5.6%

30-80nm

--

0.008µm2

65,000nm3

~200,000

13µm3

0.2%

1400µm2

0.3%

0.1-1µm

--

0.13µm2

0.004µm3

~50,000

210µm3

2.6%

6000µm2

1.3%

2

3

3

2

10-40nm

--

0.002µm

10,000nm

~100,000

1µm

0.01%

200µm

0.2-5µm
55nm

-30nm

~1µm2
0.003µm2

0.1µm3
50,000nm3

~100?
~3000

10µm3
0.16µm3

0.1%
0.002%

100µm2
10µm2

0.02%
0.002%

0.5-1µm
--~11nm

--5-6nm
~36nm

--2.2µm2
0.002µm2

0.31µm3
0.30µm3
0.01µm3
3400nm3

~300
--~107

~90µm3
~90µm3
0.017µm3
~20µm3

1.1%

--600µm2
8000µm2

0.1%
1.7%

0.5-1µm
---

--5-6nm

--2.2µm2

0.31µm3
0.30µm3
0.01µm3

~300
---

~90µm3
~90µm3
0.017µm3

1.1%

--600µm2

0.1%

0.5-1µm

2-3µm
5-6nm
5-6nm

6.3µm2
6.3µm2
30µm2

~1µm3
0.035µm3
0.17µm3

300-3000
---

1000µm3
---

12.5%

-10-100nm
~µm
~µm
1-10µm
0.2-25µm
300-500nm

-400nm2
0.02µm2
0.03µm2
0.2µm2
0.016-2µm2
0.22µm2

-450nm3
36,000nm3
100,000nm3
900,000nm3
0.0001-0.02µm3
0.007µm3

1
~108
~107
~106
<104
~50,000
2

875µm3
45µm3
360µm3
100µm3
-370µm3
0.014µm3

10.9%

--7-8nm
7-8nm
20-40nm
~100nm
30-40nm
-6nm
65mm
---~20µm

--200µm2
200µm2
-~40,000nm2
196µm2
<78µm2
350nm2
390µm2
---2400µm2

268µm3
-1.5µm3
1.5µm3
6µm3
0.0005µm3
6.9µm3
25-65µm3
400nm3
0.18µm3
~206µm3
~3348µm3
~7712µm3
8000µm3

1
1
--1
2000-4000
1
1
2.5 x 106
46
1
1
1
1

268µm3
-1.5µm3
1.5µm3
6µm3
1.5µm3
6.9µm3
45µm3
10µm3
8.3µm3
~206µm3
~3348µm3
~7712µm3
8000µm3

-2-4nm
5-7nm
8-12nm
15nm
25nm
150nm

Nucleus, total
8µm
Nuclear envelope
8µm
(Outer layer)
~8µm
(Inner layer)
~8µm
Perinuclear space
-Nuclear pores
70-90nm
Nuclear cortex
~7.9µm
Nucleolus
<5µm
Nucleosomes
10nm
Chromatin
1.9nm
Nucleoplasm
-Cytosol (~1020 kg/m3)
-Cytoplasm
-Cell, total(~1064 kg/m3) ~20µm

0.2%

0.0002%
3.4%

41.9%
96.4%
100%

-6000µm2
(30,000µm2)

0.04%

1.3%

280,000µm2
40,000µm2
200,000µm2
30,000µm2
-10,000µm2
0.44µm2

59.2%

27,500µm2
400µm2
200µm2
200µm2
-120µm2
196µm2
-8750µm2
18,000µm2
---469,000µm2

5.9%

100%


preclude insertion of specialized sensory tools or small manipulatory devices into organelle interiors.

8.5.3.2 Cell Membrane

The cell and most organelles are individually surrounded by their own thin envelope. The envelope surrounding the entire cell is called the plasma membrane. According to the fluid mosaic model, all membranes are composed of a double layer of lipid molecules, called the lipid bilayer, in which proteins are embedded (Fig. 8.37). The lipid bilayer acts as a barrier to the diffusion of polar solutes, whereas the embedded proteins provide the pathways for

1. the selective transfer of certain molecular substances through the lipid barrier, and
2. the mechanical transfer of information from the ECM into the interior of the cell.

The plasma membrane actually represents only a tiny fraction of the total membrane surface in the cell. For example, the combined membrane surface area of the endoplasmic reticulum (Section 8.5.3.5) is 44 times larger than the plasma membrane surface area for a typical human cell (Table 8.17).

Lipid bilayer plasma membranes are 6-10 nm thick. The major membrane lipids are phospholipids, fatty acid chains in the range of 16-18 carbons long; chains with fewer than 12 carbons cannot form a stable bilayer. Phospholipids are amphipathic molecules—one end, the head, has a negatively-charged (polar) region, while the remainder of the molecule, the tail, consists of two (nonpolar) long fatty acid chains. The phospholipids in cell membranes self-organize into a bimolecular layer, with the nonpolar fatty acid chains in the middle. The polar regions are oriented toward the membrane surfaces due to their attraction to the polar water molecules in the extracellular and cytosolic fluids (Fig. 8.37).

The plasma membrane also contains other lipids (Table 8.18). For example, cholesterol, a steroid lipid, acts as a “mortar” that fills in small gaps in the phospholipid structure, thus improving membrane impermeability to small water-soluble molecules like glucose by a factor of ten. Cholesterol also acts as a membrane antifreeze agent, decreasing bilayer fluidity at higher temperatures (e.g., raising lipid bilayer “melting point”) and preventing hydrocarbon chains of phospholipids from aggregating at lower temperatures (e.g., lowering membrane “freezing point”). Plasma membranes may contain up to ~1 cholesterol molecule for each phospholipid molecule. The precise lipid composition of plasma membranes varies from one cell type to another, and also varies among the membranes of organelles within each cell type (Table 8.18). Medical nanorobots equipped with suitable chemosensors may access this information, both for cell type identification during extracellular navigation and for organelle type identification during intracellular navigation.

There are ~5 x 10^6 lipid molecules in a 1 micron^2 area of lipid bilayer or ~2.5 bilayer lipid pairs/nm^2 of cell membrane surface. Thus the plasma membrane of a typical 20-micron human tissue cell contains ~10 billion lipid molecules. Phospholipids are not covalently bound to each other, so each lipid molecule is free to move independently, resulting in considerable random lateral movement parallel to the bilayer surfaces. The long fatty acid chains each include one unsaturated bond, producing a kink in the otherwise straight chain that prevents close packing (and solidification). The chains also wiggle back and forth, so the lipid bilayer has fluid-like characteristics much like a layer of oil on a water surface. Movement of hydrophilic head groups through the hydrophobic interior of the membrane is thermodynamically unfavorable. Such flip-flopping, or transverse diffusion, does occur in membrane lipids but is relatively slow. For instance, a typical phospholipid molecule undergoes transverse diffusion (one flip flop between monolayers) once every several hours in a lipid bilayer. By contrast, lateral diffusion of phospholipids (movement within each monolayer) is so rapid that a lipid molecule can move 10 microns (the equivalent of ~12% of cell circumference) in a few seconds. (But see Section 9.4.3.3 regarding the membrane-skeleton fence model.)

Membrane proteins are embedded in the lipid bilayer plasma membrane. Indeed, it has been said that the lipid bilayer serves as a “solvent” for membrane proteins. The plasma membrane contains roughly equal masses of lipid and protein (Table 8.18). However, the mass of an individual protein molecule is much larger than the mass of any lipid molecule, so there are 10-100 times more lipid molecules than protein molecules. The plasma membrane of a typical 20-micron human tissue cell contains ~0.1 billion protein molecules.

There are two classes of membrane proteins: Integral (intrinsic) membrane proteins and peripheral (extrinsic) membrane proteins. Integral membrane proteins are closely associated with membrane lipids and cannot be extracted from the membrane without disrupting

### Table 8.18. Mass % Biochemical Composition of Cell and Organelle Membranes

<table>
<thead>
<tr>
<th>Type of Membrane Molecule</th>
<th>Liver Cell Plasma Membrane</th>
<th>Red Cell Plasma Membrane</th>
<th>Myelin Sheath</th>
<th>Mitochondrion Inner/Outer Membranes</th>
<th>Endoplasmic Reticulum Membrane</th>
<th>E. coli (Bacterial Membrane)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lipid</td>
<td>~17%</td>
<td>~23%</td>
<td>~22%</td>
<td>3%</td>
<td>~6%</td>
<td>~0%</td>
</tr>
<tr>
<td>Protein</td>
<td>~50%</td>
<td>~52%</td>
<td>~19%</td>
<td>~76%/~52%</td>
<td>~50%</td>
<td>~50%</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>~8%</td>
<td>~8%</td>
<td>~8%</td>
<td>~8%</td>
<td>~8%</td>
<td>~8%</td>
</tr>
<tr>
<td>Lipid Class:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td>~17%</td>
<td>~23%</td>
<td>~22%</td>
<td>3%</td>
<td>~6%</td>
<td>~0%</td>
</tr>
<tr>
<td>Phospholipids</td>
<td>~7%</td>
<td>~18%</td>
<td>15%</td>
<td>~35%</td>
<td>17%</td>
<td>70%</td>
</tr>
<tr>
<td>Phosphatidylethanolamine</td>
<td>~4%</td>
<td>7%</td>
<td>9%</td>
<td>~2%</td>
<td>5%</td>
<td>trace</td>
</tr>
<tr>
<td>Phosphatidylserine</td>
<td>24%</td>
<td>17%</td>
<td>10%</td>
<td>~39%</td>
<td>40%</td>
<td>0%</td>
</tr>
<tr>
<td>Phosphatidylcholine</td>
<td>19%</td>
<td>18%</td>
<td>8%</td>
<td>0%</td>
<td>5%</td>
<td>0%</td>
</tr>
<tr>
<td>Sphingomyelin</td>
<td>7%</td>
<td>3%</td>
<td>28%</td>
<td>trace</td>
<td>trace</td>
<td>0%</td>
</tr>
<tr>
<td>Glycolipids</td>
<td>22%</td>
<td>13%</td>
<td>8%</td>
<td>21%</td>
<td>27%</td>
<td>30%</td>
</tr>
<tr>
<td>Other lipids</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the lipid bilayer. Like phospholipids, integral proteins are amphipathic. Polar amino acid side chains lie in one region of the molecule and nonpolar side chains are in a separate region. Thus integral proteins vertically align with the amphipathic lipids in the plasma membrane—protein polar regions position themselves at the surfaces in association with polar water molecules, while the protein nonpolar regions are attracted to the interior in association with the nonpolar fatty acid chains at the center of the lipid bilayer membrane (Fig. 8.37; see also Figures 8.33 and 8.34). Many integral proteins can move laterally in the membrane; others are immobilized by links to a network of peripheral proteins located near the cytoplasmic surface of the membrane. How fast do embedded proteins laterally diffuse? If a single hybrid cell is created by fusing two cells having radiochemically-tagged membrane protein molecules, ~1 hour is needed for the two populations of transmembrane protein molecules to become thoroughly randomly intermixed.939

Most integral proteins are transmembrane proteins with polar regions at each end and a nonpolar region in the middle, spanning the entire membrane. These polar regions may extend up to 10-20 nm beyond the surface of the lipid bilayer, forming channels through which water, ions, or chemical signals can pass into the cell (Section 3.3.3). A few integral proteins do not cross the entire membrane and are found only in the outer or inner layer, performing functions localized to only one membrane surface. These proteins are also amphipathic and oriented parallel to the lipid molecules. Some are anchored to the membrane by covalent bonds with phospholipids. For example, in the red blood cell membrane, glycophorin spans the entire membrane, all glycolipids and most of the phosphatidylycholine are in the outer monolayer, and the majority of the phosphatidylethanolamine and phosphatidyserine molecules are in the inner monolayer where most of the proteins reside. (Cholesterol is distributed about equally between the two layers.) The number of different integral proteins in a membrane ranges from 6-8 in the sarcoplasmic reticulum to over 100 in the plasma membrane (including enzymes, transport and structural proteins, antigens and receptors), many of which are present in only a few copies per cell, although the 135 micron² red cell surface has ~1 x 10⁶ copies of the glycophorin A molecule and ~1 x 10⁵ copies of glycophorin B.1091

Peripheral membrane proteins are bound to the hydrophilic regions of integral membrane proteins or to the hydrophilic heads of membrane lipids by weak electrostatic forces.399 Most peripheral proteins are located near the cytoplasmic surface of the plasma membrane rather than on the extracellular surface and mediate such properties as cell shape and motility. Peripheral proteins are not amphipathic and do not associate with the hydrophobic regions of the lipids in the membrane interior. Both lipid and protein components of the plasma membrane are continually removed and replaced. Turnover allows the cell to continuously change out damaged components. This is a highly selective process, since the rate of turnover varies for different proteins and lipids. For instance, the half-life of some phospholipids in membranes is ~10,000 sec;399 the “off-rate” (half-life) for cholesterol from a lipid bilayer (e.g., the red cell surface) into the cytoplasm is ~7200 sec at 310 K.1113 Protein turnover half-lives may range from several minutes to several years, but the “typical” protein has a turnover half-life of ~200,000 sec531,939 or ~2 days. Protein replacement is carried out by protease enzymes located in the cytoplasm and in lysosomes. Replacement rates also depend upon cell type. For example, the plasma membrane surface of the macrophage has an unusually fast mean turnover time, ~1800 sec, vs. ~5400 sec for fibroblasts.996

The plasma membrane also contains small amounts of carbohydrate. This carbohydrate is covalently linked to some of the membrane lipids and proteins. Carbohydrate portions of the membrane glycoproteins (e.g., Section 8.5.2) are always located at the extracellular surface, forming the glyocalyx (together with collagen proteins and glycosaminoglycans, aka “mucopolysaccharides”). The red cell membrane, for instance, contains 52% protein, 40% lipid, and 8% carbohydrate by weight.939 A small proportion of membrane carbohydrate is glycolipids, but most is in the form of glycoproteins. The sugar units are usually short oligosaccharide chains attached to serine, threonine, or asparagine side chains.

The glyocalyx, or fuzzy coat, lies exterior to the plasma membrane. In most cell types, the glyocalyx is 10-100 nm thick consisting of tangled strands of up to ~10,000-atom glycoproteins each measuring 5-8 nm thick and up to 100-200 nm in length.531,939 The experimentally-measured thickness of the glyocalyx of various cells ranges from ~6 nm for human blood-group A erythrocytes,3163 to 13 nm in Eimeria microgametocytes,3588 20-30 nm for chick fibroblasts,3589 30-60 nm for human bladder cells,3590 40-70 nm for human lymphocytes,3591 ~50 nm for human myocardial cells,3592 56 nm for frog mesenteric microvessels,3593 >70 nm for rat vasculature,3594 ~81 nm for rabbit endothelial cells of the systemic arteries (e.g., carotid),3164 and 90 nm for human cochlear hair cells.3595 The most prominent glycaloci are found in intestinal epithelial cells, where the fuzzy coat may reach 150 nm in thickness and consists primarily of oligosaccharide chains 1.2-2.5 nm in diameter.399 (There is one report of rat venule endothelial cells with glycaloci up to 870 nm thick,3954 and a few macroscopic parasites such as the fork-tailed cercariae of the blood fluke Schistosoma mansoni have glycaloci 500-2000 nm thick.3596)

The glycoproteins of the glyocalyx provide a set of highly specific biological markers that are readily recognizable by suitably equipped medical nanorobots. These markers assist normal cellular interactions by allowing blood group recognition, bacterial and toxin binding sites, egg recognition by sperm, immune responses, guidance of embryonic development, and cellular lifespan determination (e.g., the red cell coat thins with age which may serve as an RBC removal signal for phagocytes and in the liver).960

The cell’s surface is also strewn with numerous pits and indentations. For example, one class of these is the “coated pits” whose inner surfaces are covered by a dense layer of the protein clathrin, important in receptor-mediated endocytosis wherein proteins and other large molecules are imported into the cytoplasm (Section 8.5.3.7). Another class of membrane indentation is the ~50 nm caveola ("tiny caves") that serve to draw substances such as vitamins and signal transduction molecules into the cell’s interior.1137,3576-3579 Caveolae are coated with a unique membrane marker protein called caveolin, making them easy for nanorobots to identify.

8.5.3.3 Cytosol

In traditional cell biology, the “cytoplasm” is the filling substance in the large space enclosed by the plasma membrane and surrounding the cell nucleus. It includes all non-nuclear cell organelles plus all fluid surrounding these organelles. Every point inside the cell, except the nucleus, is considered part of the cytoplasm. The “cytosol” is the liquid that fills all of the cytoplasmic region, except for fluids filling the interior of the cell organelles. Finally, the term “intracellular fluid” refers to all of the fluid inside a cell—the cytosol plus the fluid inside each organelle, including the nucleus. Internal hydrostatic pressure in the cytosol ranges from ~23 N/m² in the human erythrocyte cell1452 up to ~1.6 x 10⁵ N/m² in Amoeba proteus.1453
Fulton\textsuperscript{941} points out that in an average protein crystal ~40% of crystal mass is solvent water—half strongly absorbed by the protein in the hydration shell and half resembling bulk water in its properties. Protein crystals may range from 20-90% solvent, or 10-80% protein, by weight. Muscle cells are ~23% protein, red cells ~35%, and actively growing cells contain 17-26% protein by weight.\textsuperscript{938} Thus the cytosolic environment more closely resembles proteinaceous crystals than dilute solutions (e.g., <0.1% proteins).

Hydration water is not immobile water like ice, but it does have reduced mobility, different solvating characteristics, a higher heat capacity, and is generally more ordered than bulk water.\textsuperscript{1938} Water of hydration coats the macromolecular cell components such as carbohydrates, nucleic acids, and the proteins of membranes, and is required for enzyme activity. Mammalian cells can maintain glycolysis and normal respiration down to 60% dehydration, which suggests that in a typical cell the cytosol may consist of ~60% bulk water and ~40% water of hydration. Distribution of ordered water in the cell is a heterogeneous and dynamic process, possibly cytonavigationally useful. For instance, ~55% of the water of the vegetal pole region of the frog oocyte is bound water, with only ~25% bound near the animal pole cytoplasm and ~10% bound in the nucleus.\textsuperscript{1937}

Minton\textsuperscript{1010} showed that the volume occupied by proteins affects the activity of the other proteins in solution by “crowding” them into a smaller volume where they have less freedom of movement, forcing them into compact configurations that would occur far less frequently in dilute solutions. These effects can produce 50-1000-fold excesses in protein activity over the values predicted from dilute solutions. Crowding also reduces diffusion mobility—60%-70% of glycolytic enzymes are freely diffusing (30%-40% are immobilized on the matrix) in squid axoplasm that contains 2% protein by weight, but only 20% of cytoplasmic proteins are freely diffusing in oocyte cytoplasm that contains 30%-40% protein by weight.\textsuperscript{943} Crowding effects could be nonspecific and due simply to protein number density, or they could be specific and maintained by selection. McConnell\textsuperscript{1011} estimates that about half of the polypeptides in the cell may participate in specific associational structures. For instance, many enzymes exist in complexes that may involve a dozen or more proteins, potentially complicating required nanorobotic actions; according to the substrate channelling hypothesis, these complexes can help speed up reactions.

Even without these crowding effects, cells display localized biochemical gradients that are cytonavigationally significant. Specific molecules, organelles and physiological processes may be localized to defined zones within the cell. This regional cytoplasmic differentiation is regulated in part by cytosolic Ca\textsuperscript{2+} ion and by H\textsuperscript{+} ion (pH) spatial gradients.\textsuperscript{1076,1077} As another example, the cytosol immediately surrounding the Golgi apparatus is compositionally distinct from the cytosol that closely encircles the cell nucleus.\textsuperscript{551} Active chemical processing taking place within most cellular organelles produces persistent concentration gradients that may be used by in cyto nanorobots to establish proximity to specific organelles. Microsecond-cycle chemical nanosensors (Section 4.2.1) can sample the local environment far better than the typical time required for a molecule to diffuse a distance of ~1 micron, the mean separation between adjacent major organelles inside human cells. Taking cytoplasm absolute viscosity $\eta$ ~ $6 \times 10^{-3}$ kg/m-sec,\textsuperscript{562} then from Eqn. 3.1 the 1-micron diffusion time $\tau_{\text{diffuse}} >> \tau_{\text{meas}}$ (1 microsec) for small molecules like amino acids or glucose ($\tau_{\text{diffuse}}$ ~ 5000 microsec) or for 100,000-dalton macromolecules measuring ~10 nm in diameter ($\tau_{\text{diffuse}}$ ~ 130,000 microsec).

Control proteins such as the transcription factors that regulate gene expression are typically present in only ~300-3000 copies of each type per cell. More common cytosolic proteins may be present in numbers up to $10^6$-$10^7$ copies of each type per cell.

8.5.3.4 Ribosomes

Ribosomes are the smallest and most numerous “organelle” (macromolecular assembly) in the human cell. A typical liver cell contains $10^7$ ribosomes constituting ~5% of the total dry mass of the cell. Cells less actively involved in protein synthesis have correspondingly fewer ribosomes. Some ribosomes float freely in the cytosol, making proteins for intracellular use. Others are attached to membranes or to the cytoskeleton, synthesizing proteins destined for membranes or for export from the cell.

Ribosomes are ~25 nm in diameter. Each of the several types of ribosome (containing rRNA) is constructed of two subunits that fit snugly together. A typical ribosome might have a mass of 4.2 million daltons, comprised of 2.8 million daltons for the large ribonucleoprotein 60S subunit (~23 nm diameter) and 1.4 million daltons for the small ribonucleoprotein 40S subunit (~9 nm diameter).\textsuperscript{938,996} There is no intrinsic difference between free cytosolic ribosomes (50% or more of the total) and membrane-bound ribosomes—it is the “signal” sequence on the end of the protein being synthesized that directs a free ribosome to become a bound ribosome.

The ribosome is an ATP-powered protein-assembling machine (Fig. 2.2). Amino acids drawn from the cytosol are presented to the ribosome, which incorporates them one by one into polypeptides at ~20 Hz.\textsuperscript{531} The complete synthesis of an average-sized protein takes 20-60 seconds. Even during this short period, multiple initiations take place, with a new ribosome hopping onto the 5' starting end of an mRNA (messenger) molecule almost as soon as the preceding ribosome has translated enough of the amino acid sequence to get out of the way, thus allowing the assembly of many copies of the same protein to proceed almost in parallel.\textsuperscript{531} Thus, under physiological conditions, actively translated mRNA is found in polyribosomes, or polysomes, formed by gangs of multiple ribosomes spaced as close as 80 nucleotides apart along a single messenger molecule (Fig. 8.38). Further discussion of ribosomes may be found in Volume II.
8.5.3.5 Endoplasmic Reticulum

The endoplasmic reticulum (ER), the principal protein manufacturing facility of the cell (aside from free ribosomes), is the most spatially extensive cytoplasmic organelle (Fig. 8.39). The ER is an interconnecting membranous network of fluid-filled vesicles, branching tubules, and flattened sacs or cavities called cisternae, possessed by almost all eukaryotic cells. The 5-6 nm thick membrane of the endoplasmic reticulum typically constitutes more than half of the total membrane surface area present in a cell. The membranes enclose a space that is continuous throughout the network and connects with the perinuclear space between the two membranes of the nuclear envelope (Section 8.5.4.1), allowing transport of manufactured substances throughout the cell. The fluid occupying the luminal or cisternal space of the ER, which is typically 20-40 nm wide but ranges from 10-70 nm, is called the reticuloplasm. There are ER-specific luminal proteins called reticuloplasmins found only in the reticuloplasm—"if these proteins are detected, a nanorobot has localized its position uniquely to this space. For example, molecular chaperones that assist in protein folding and are unique to the ER lumen include immunoglobulin heavy-chain binding protein (BiP) and calnexin.

The Golgi complex (GC) is a multicisternal membranous structure crudely similar to the rough ER (minus the ribosomes). Substances synthesized by the ER usually pass via vesicles traversing the cytoplasm on microtubule tracks to the Golgi complex, where they are further processed, concentrated, sorted, and shipped out to various destinations. Some oligosaccharides are synthesized in the ER, while extensive portions of the same structures are degraded in the Golgi—"competing sets of reactions that are crucial in fixing the "address" of the intracellular compartment to which a newly synthesized macromolecule will be sent. Thus the GC serves as a production editor, allowing several successive sorting steps to take place while reducing the overall cytomanufacturing error rate. The Golgi acts as a "countercurrent" fractionation system to separate proteins destined for the plasma membrane from those to be retained in the ER.

8.5.3.6 Golgi Complex

The Golgi complex (GC) is a multicisternal membranous structure. Substances synthesized by the ER usually pass via vesicles traversing the cytoplasm on microtubule tracks to the Golgi complex, where they are further processed, concentrated, sorted, and shipped out to various destinations. Some oligosaccharides are synthesized in the ER, while extensive portions of the same structures are degraded in the Golgi—"competing sets of reactions that are crucial in fixing the "address" of the intracellular compartment to which a newly synthesized macromolecule will be sent. Thus the GC serves as a production editor, allowing several successive sorting steps to take place while reducing the overall cytomanufacturing error rate. The Golgi acts as a "countercurrent" fractionation system to separate proteins destined for the plasma membrane from those to be retained in the ER.

Morphologically, the Golgi complex consists of a series of flattened, membranous sacs, forming disk-shaped cisternae that are stacked together to form a cup-shaped structure. Cisternal boundaries are morphologically distinct and are only rarely, if at all, connected. A series of 5-8 sacs comprises a single Golgi stack or dictyosome, typically measuring ~1 micron in diameter and ~250 nm thick. As with the ER, the size and number of Golgi complexes vary according to cell type and metabolic activity. All eukaryotic cells have a GC. Some cells have just one stack; others, particularly those especially active in secretion, may have hundreds. One enzyme, thiamine pyrophosphatase, occurs only in Golgi membranes. Galactosyl transferase is found only in the GC. These may be used by nanorobots as unique cytochemical markers of the Golgi complex.
of the GC, vesicles bud off the tips of the saccules continuously and exit the complex, transported on cytoskeletal elements through an interaction with motor proteins. This is known as the "cisternal maturation model." Within the Golgi, net membrane flow can be quite rapid. In certain mucus-secreting cells of the intestinal mucosa, it takes only ~2000 sec for membrane components to move from the forming face of a Golgi stack to the maturing face. Vesicles exhibiting COPI (coat protein I) apparently transport some proteins backwards, both within the Golgi stack and also from the Golgi to the ER. One study counted ~400 individual vesicles in transit within ~0.2 microns of a ~4 micron region containing 2 stacks of 7 cisternae.

The two faces of a Golgi stack are biochemically distinct. Specific enzymes and receptor proteins are concentrated in the cisterna on the "cis" face of the stack (the "receiving" end), while other proteins are localized mainly in the cisterna on the "trans" face (the "shipping" end). For example, the receptor protein for the carbohydrate mannose-6-phosphate (Man6P) is present only in the forming cisterna of the Golgi stack, thus allowing the stack to recognize these proteins and target them to the lysosomes.

A 28,000-dalton protein called GS28 acts as one of the targeting receptors involved in the recognition of ER-derived vesicles that are destined for fusion with the cis-Golgi membrane. By 1998, more than two dozen different members of the small G protein superfamily (including the Rab and Arf protein families) had been implicated in the regulation of intracellular vesicular trafficking and membrane recognition. The Golgi complex also contains substantial amounts of cholesterol and sphingolipids, making a rising gradient from cis to trans. Various Golgi compartments are defined by the presence of particular enzymes (e.g., glycosidases, glycosyltransferases, and others involved in protein modification or sphingolipid synthesis) that are localized to one or more compartments, often in a graded manner.

The Golgi stack has other easily detectable asymmetries. Since the GC mediates the flow of secretory proteins from the ER to the cell exterior, the GC displays spatially variant protein and lipid compositional gradients along the Golgi polarity axis. Specifically, the saccul membranes at the forming face resemble those of the ER in morphology and composition—5-6 nm thick membranes defining a relatively narrow ~30-80 nm cisternal space, with ~20% phosphatidylcholine in the membrane. At the maturing pole, the saccl membranes more closely resemble the plasma membrane, ~10 nm thick, with a much wider cisternal space (~100 nm) and with ~10% phosphatidylcholine in the membrane. Variable membrane thickness is part of the control mechanism for spatially anisotropic budding.

Another asymmetry: Swarms of ~50 nm diameter membrane-bound vesicles always cluster on the side of the Golgi stack abutting the ER and along the circumference of a stack near the dilated rims of each cisterna.

8.5.3.7 Vesicles, Granules, and Vaults

For most membranes, such as those of the Golgi complex and lysosomes, the transfer of manufactured lipid and protein from the rough ER occurs by means of 30-80 nm transport vesicles that pinch off from the rough ER and fuse with the target membranes. Some of these vesicles are coated with a bristlelike polyhedral lattice of clathrin subunits. Clathrin, a 180,000-dalton protein, is a distinctive feature of vesicles involved in a variety of intracellular transport processes. Besides bringing secretory proteins from the ER to the GC, clathrin-coated vesicles also transport membrane proteins from the GC to lysosomes (Section 8.5.3.8), to the plasma membrane (Section 8.5.3.2), and to other cellular destinations. In each case the clathrin coat forms a basket or cage around the vesicle, and transport involves cytoskeletal elements. Clathrins are cell specific—for example, neuron clathrins have 12 pentagons and 20 hexagons, liver clathrins have 30 hexagons, and fibroblast clathrins have 60 hexagons and thus may be distinguished by medial nanorobots. In cultured cells, assembly of a free clathrin-coated vesicle takes ~1 minute, and >300-1000/min can be formed.

Cellular products are concentrated and packaged before being discharged to the outside of the cell, a process known as exocytosis. Concentration occurs as structures called condensing vacuoles located on the periphery of the Golgi complex fill with concentrated protein, then fuse with one another to form larger secretory granules and vesicles. Outbound vesicles are given unique membrane compositions, including high-specificity targeting molecules that can bind only to receptor molecules located on the appropriate surface of the target acceptor compartments. A well-known example is the synaptic vesicle targeting protein synaptobrevin (VAMP 1 and 2) which binds only to the neuron-specific plasma membrane proteins syntaxin 1A and 1B, thus ensuring proper vesicle docking and fusion at the neuron cell plasma membrane. In general, the fusion of two distinct lipid bilayers is energetically unfavorable in the absence of these specialized targeting proteins.

In some cells, the complementary process of endocytosis is also important. In endocytosis, the cell ingests extracellular materials by invaginating the plasma membrane, then budding off vesicles, vacuoles (phagocytic, pinocytic, etc.), or "endosomes" to the interior of the cell from these sites. Endocytosis and exocytosis produce opposite membrane flows, the former reducing plasma membrane mass and the latter increasing it via fusion. The net membrane flow can be rather large in cells that carry out both processes actively. For instance, in cultured macrophages an amount of membrane equivalent to the entire surface area of the cell is replaced in ~1800 sec, and macrophages may ingest ~25% of their volume per hour.

Storage granules are also important in the cell. For example, glycogen is the storage polysaccharide of the animal body, also known as animal starch. It is a polymer of glucose. In many cells, large individual molecules of polymerized glucose measuring 10-40 nm in diameter appear as granules. The enzymes needed to carry out the synthesis and degradation of glycogen (including the synthetic enzyme glycogen synthase and the degradative enzyme glycogen phosphorylase) are bound to the surface of these glycogen granules in a ~5 nm-thick shell. These granules constitute up to 4-6% by weight of liver cells (~5 million granules), up to 0.7-1% of muscle cells, with lesser amounts in other cells (typically <10^7 granules) and only very small amounts in the brain, and are absent in some cells.

Another storage vesicle found floating in the cytoplasm is the lipid droplet containing triglyceride, the main storage form of fatty acids in cells. In many cells, these insoluble triglycerides coalesce in the cytosol to form large, anhydrous droplets from 0.2-5 microns in diameter. In adipocytes, the cells specialized for fat storage, these droplets can be as large as 80 microns, occupying virtually the entire cytosol and constituting up to ~99% of the cell's organic matter.

Other intracellular storage vessels include melanin pigment granules found in certain cells of the skin and hairs; zymogen (enzyme-containing) granules synthesized in pancreatic cells, then
transported into the small intestine; inflammatory toxins stored as intracellular granules in eosinophil leukocytes; ferritin molecules containing cellular iron stores (each ~610-690 kilodalton molecule is comprised of 24 subunits of 18,500 daltons each, which surround in an 8-nm-diameter-cavity micellar form some 3000-4500 ferric atoms);\textsuperscript{396,3380} and water vacuoles, mucus vesicles, and crystals of various types. The smallest observed effective hydrodynamic radius of sonicated phospholipid vesicles is ~10.25 nm independent of hydrocarbon chain length for synthetic even-numbered 12-18 carbon-chain phosphatidylcholines, and ~10.7 nm for egg-yolk phosphatidylcholine vesicles.\textsuperscript{3949}

Finally, a related cell component is the vault, numbering in the thousands in human cells.\textsuperscript{3581} Vaults are barrel-shaped particles measuring ~55 nm x 30 nm, assembled from 96 copies of MVP (major vault protein) plus some integral RNA. Vaults look like pairs of unfolding flowers, each half having 8 petals attached to a central ring with a small hook. Vaults were first discovered in the mid-1980s and their exact function was not yet known in 1998. However, their structure suggested an ability to open and close as a natural part of their function in the cell,\textsuperscript{1410} and their size and shape would be an almost perfect match to dock at the nuclear pore complex (Section 8.5.4.2), which suggested to some investigators that vaults might serve to ferry mRNA around the cell.\textsuperscript{1001,3581}

In 1998, the chemical content of individual vesicles was analyzed regularly using a combination of optical trapping, capillary electrophoresis separation, and laser-induced fluorescence detection.\textsuperscript{1263}

### 8.5.3.8 Lysosomes and Proteasomes

Lysosomes are Golgi-budded organelles that contain ~40 digestive enzymes\textsuperscript{231} capable of degrading all major classes of biological molecules — including at least 5 phosphatases, 4 proteases, 2 nucleases, 6 lipases, 12 glycosidases, and an a-lysosulfatase.\textsuperscript{935} These enzymes are needed both to degrade materials brought into the cell from the outside and to degrade internal cellular structures that are damaged or are no longer needed. Lysosomal enzymes are acid hydrolases stored inside the lysosome as small granules ~8 nm in diameter—protein aggregates of the hydrolytic enzymes with a pH optimum around 5, representing over 60% of organelle mass.\textsuperscript{231}

The lysosome can release these enzymes after fusion with endosomes containing foreign matter, or to digest dead portions of the cell or malfunctioning organelles, or to destroy abnormal substances such as bacteria that enter the cell. An unusual oligosaccharide containing mannose-6-phosphate (M6P) serves as a recognition marker or address that targets all such enzymatic proteins to the lysosomes, and sorting nexins are proteins containing lysosomal targeting codes using either tyrosine- or di-leucine-based motifs.\textsuperscript{1032} These membrane markers are readily detectable by nanorobots suitably equipped with transmembrane sensory tools, although it is important not to destroy the membrane integrity or to allow leaks (Section 9.4.5.5) which may alter fluid pH, during any probing. Since intralysosomal pH is different from the rest of the cell, chemosensors intended to probe the lysosomal interior may require a different design (or different detection parameters) than cytosolic chemosensors. In general, nanorobots can exploit the same biochemical markers that the cell normally uses to transport organelles to specific locations within the cell. Organelle surface chemosensing is most valuable for navigation and identification, while chemosensing of the organelle interior is also important for diagnosis and treatment (Chapter 21).

Lysosomes are spherical or nearly spherical organelles, typically 0.5-1 microns (range 50 nm—3 microns) in diameter. They vary greatly in appearance and content according to cell type and also in relation to the physiological state of a particular cell. The lysosome has a unique 5-6 nm three-layered membrane\textsuperscript{938} with a special transmembrane transport protein that uses ATP to pump H\textsuperscript{+} into the organelle lumen, thereby maintaining the internal pH at 5. The lysosomal membrane also has special docking marker acceptor proteins that mark a lysosome as a target for fusion with specific transport vesicles in the cell. Soluble products of digestion can cross the membrane, exit the organelle, and enter the cytosol for recycling into the cellular metabolism. Indigestible material slowly accumulates in the cytoplasm as lipofuscin pigment granules and other residual bodies. Lysosomes are present in most cells except red cells.

Although it is normally quite stable, the lysosome membrane can become more fragile when the cell is injured or deprived of oxygen or when excessive amounts of vitamin A are present. Lysosomes were once called “suicide sacs” because lysosomal rupture can result in self-digestion of the cell, a process known as autolysis. However, it is now known that lysosomes are part of the normal cellular digestion apparatus relating the process of endocytosis to the processes of intracellular synthesis, storage, and transport,\textsuperscript{938} and structural deterioration of lysosomes does not occur rapidly in ischemic or post-mortem cells.\textsuperscript{2019,2020} Most of the digestive enzymes require the low pH of the lysosomal or peroxisomal vesicles for activation (just as some proteases require the low pH of the stomach).

In addition to the lysosomal degradation system, there is also a large number of small cytosolic proteasomes (each having a 700,000-dalton 20S core complex ~11 nm in diameter). This serves as an extra-lysosomal ATP-driven system for selectively degrading endogenous ubiquitinated proteins in virtually all human cells.\textsuperscript{1087,1088,3383} (Chapter 13). Indeed, the ~2 megadalton 26S proteasome complex (containing 30-40 different proteins) appears to be the major protease of the nuclear and cytoplasmic compartments of the eukaryotic cell—proteasomal degradation controls the lifetime of most cellular proteins, including many regulatory proteins, and generates peptide antigens for presentation by MHC Class I molecules to CD8\textsuperscript{+} cytotoxic T cells.\textsuperscript{2915} Proteasomes are ubiquitous and very abundant, comprising up to 1% of total cellular protein.\textsuperscript{1087,2016,2017} The entire 26S complex is 30-44 nm in length.

### 8.5.3.9 Peroxisomes

Peroxisomes (also called microbodies)\textsuperscript{1000} are rough ER-budded organelles typically 0.1-1 microns in diameter.\textsuperscript{531} Peroxisomes occur in most cells, but are most prominent in liver and kidney cells which are active in detoxification. The organelle destroys cytotoxic chemicals such as hydrogen peroxide, methanol, ethanol, formate, formaldehyde, nitrites and phenols, as well as D-amino acids which are not found in proteins and are not recognized by enzymes involved in the degradation of the more common L-amino acids. Peroxisomes also catalyze 25-50% of all fatty acid breakdown in cells (the remainder occurs in the mitochondria). The organelle is a concentrated source of three oxidative enzymes found in liver cells—D-amino acid oxidase, urate oxidase (usually present in a distinctive crystalline core), and catalase (which is up to 40% of total peroxisomal protein and most of the catalase present in a cell).

Peroxisomes may also protect cells from toxic oxygen levels\textsuperscript{3069-3072} since peroxisomal O\textsubscript{2} consumption is directly proportional to cellular oxygen concentration,\textsuperscript{939} unlike mitochondrial respiration. In case of excessive O\textsubscript{2} levels, peroxisomal respiration is greatly stimulated, reducing the intracellular oxygen tension.

### 8.5.3.10 Mitochondria

Mitochondria are the principal chemical energy transducers of the eukaryotic cell under aerobic conditions.\textsuperscript{3384} Except for the 10 reactions of the glycolytic pathway, all of the ATP-generating capacity
of eukaryotes lies within the mitochondria. Mitochondria are scattered throughout the cytoplasm in virtually all aerobic cells, with numbers ranging from ~300/cell for relatively inactive cells like lymphocytes up to 2000-3000/cell for very active cells such as liver, kidney tubule and cardiac muscle cells (where mitochondria may occupy up to 20% of total cell volume).\textsuperscript{531} Mitochondria are often clustered within the cell in regions of intense metabolic (hence ATP) demand. For instance, in muscle cells the mitochondria are organized in rows between adjacent contractive myofibrils in order to minimize the required diffusion distance for ATP molecules that are powering the activity. Similar localization appears in sperm tail flagella, in cilia, and at the base of kidney tubule cells where exchange with the blood is most rapid. Except for plant chloroplasts, mitochondria are unique among organelles in having their own DNA genomes, ribosomes and tRNAs that are quite different from those found in the cytoplasm.

After the nucleus, the mitochondrion is the largest organelle in most animal cells. The typical time-averaged dimensions are roughly cylindrical, with a 0.5-1.0 micron diameter and a ~3 micron length (and rarely, up to 10 microns)\textsuperscript{939}—about the size of a bacterium or modest-sized medical nanodevice. Mitochondria are usually depicted as tiny kidney-bean- or sausage-shaped organelles, but in living cells they squirm, flex, elongate, and change shape almost continuously (Fig. 8.40) in part due to cytoskeletal interactions. Mitochondria may be shuttled around a cell at up to ~10 microns/sec via dynein motors\textsuperscript{453} riding on microtubules (Section 8.5.3.11), bulging the plasma membrane as they travel.\textsuperscript{1249}

Mitochondria have four functional compartments: outer membrane, intermembrane space, inner membrane, and the matrix. The outer surface is a smooth, featureless, 5-6 nm thick membrane\textsuperscript{938} embedded with many copies of a transport protein that forms large aqueous channels through the lipid bilayer, so the membrane is permeable to all molecules of mass <10,000 daltons (which includes all metabolites pertinent to mitochondrial function). Since mitochondria do not grow by fusion with vesicles synthesized elsewhere in the cell, transfer of phospholipids from the endoplasmic reticulum to the mitochondrion requires special phospholipid transfer proteins in the outer membrane (e.g., glycerolphosphate acyltransferase and monoacyl glycerolphosphate acyltransferase) that can recognize a specific kind of phospholipid, remove it from a vesicle membrane and add it to the mitochondrial membrane. Special enzymes such as monoamine oxidase, acyl-CoA synthetase, and phospholipase A\textsubscript{2} are also present.\textsuperscript{996} The first 32 N-terminal amino acids of cytochrome c\textsubscript{1} constitute a fixed “leader sequence” that is recognized as a matrix targeting signal and allows the tagged molecule admission through the outer membrane.\textsuperscript{997} Thus, as with other organelles, the outer coat of the mitochondrion should be immediately and uniquely recognizable by in cyto nanorobots equipped with suitable chemosensors (Section 4.2).

The intermembrane space also contains numerous unique proteins, as for example the essential multispanning carrier (chaperone) proteins Tim10p and Tim12p.\textsuperscript{1142}

The 5.6 nm thick inner membrane is highly convoluted, forming a series of ~30 infoldings (each fold - 100 nm wide) known as cristae, which extend into the inner compartment or matrix, roughly quintupling the active surface area of the outer membrane (Fig. 8.41). The inner membrane is rich in cardiolipin, a phospholipid that accounts for 10% of the membrane lipid content and renders it unusually impermeable to most solutes. The inner membrane also contains a variety of special transport proteins that make it selectively permeable to those particular small molecules that are metabolized by mitochondrial enzymes concentrated in the matrix space,\textsuperscript{531} such as the integral membrane proteins Tim17, Tim 23, and Tim22p.\textsuperscript{1142} All of these membrane-specific proteins are readily detectable by nanorobots that are extending manipulators tipped with appropriately configured chemical sensors through the intermembrane space (Section 9.4.5).

The inner membrane is the locale of electron transport and ATP synthesis. A single mitochondrion has ~10,000 large protein F\textsubscript{1} (ATP synthase) complexes embedded in its inner membrane, randomly distributed and typically measuring ~10-20 nm in diameter.\textsuperscript{939} These integral proteins freely diffuse laterally within the inner membrane. Mobility is high—protein complexes collected at one end of a membrane by externally imposed electroophoretic forces return to a random distribution in just a few seconds.\textsuperscript{399}

The prominence of cristae is correlated with the relative metabolic activity of the cell or tissue. Heart, kidney, and muscle cells have high respiratory activity levels, and their mitochondria have correspondingly large numbers of prominent cristae.\textsuperscript{531} The number of cristae is three times greater in the mitochondria of cardiac muscle cells than in hepatic mitochondria, reflecting the greater demand for ATP in heart tissue. The cristae of mitochondria in different cell types are not only different in number and depth, but also in basic morphology (Fig. 8.42).\textsuperscript{531,3385-3388}

The interior of the mitochondrion is filled with a semifluid gel-like matrix that contains a concentrated mixture of hundreds of different enzymes related to aerobic cellular respiration and oxidation, plus several identical copies of the mitochondrial DNA genome, special mitochondrial ribosomes, tRNAs, and various enzymes required to express the mitochondrial genes. The mitochondrial genome\textsuperscript{5073} consists of a circular DNA molecule of ~11 million daltons, coding for about a dozen polypeptides. It has 16,569 base pairs and a contour length of ~5 microns. Mitochondria are self-replicating organelles (with the help of cell-supplied proteins

---

Fig. 8.40. Mitochondrial shape changes in living cells (redrawn from Alberts et al\textsuperscript{531}).
and lipids). When cellular requirements for ATP increase, mitochondria pinch in half (a process called fission) to increase their number, then both halves regrow to the former full size.\(^{940}\)

Mitochondria also are sometimes found as extended reticular networks\(^{3177-3179}\) which are extremely dynamic in growing cells (such as mammalian fibroblasts), with tubular sections dividing in half, branching, and fusing to create a fluid tubular web.\(^{3179,3180}\)

### 8.5.3.11 Cytoskeleton

In addition to the membrane-enclosed organelles described above, the cytoplasm of most cells is packed with filaments of various sizes collectively called the cytoskeleton. Much as the bony skeleton occupies ~11% of the human body volume (Section 8.2.4), the cytoskeleton likewise occupies ~11% of cell volume (Table 8.17). This elaborate three-dimensional webbing of filaments and tubules forms a highly structured yet dynamic matrix that extends throughout the cytoplasm, stretching from the nuclear envelope to the plasma membrane. The cytoskeleton helps to establish and maintain shape and plays important roles in cell movement, cell division, metabolism and growth,\(^{1426-1429}\) and even gene expression by accepting mechanical signals (e.g., through membrane proteins such as the integrins) originating in the extracellular matrix and transducing them into the nucleus.\(^{942}\) Cellular shape appears to be maintained by an architecture known as tensional integrity or “tensegrity”—a system that achieves mechanical stability because of the way compressive and tensional forces are distributed and balanced within the cell.\(^{1020,1021}\) Forcing cells to adopt different shapes causes the cells to switch between different active genetic programs.\(^{718,1425}\)

The cytoskeleton also serves as a framework for positioning, anchoring, and actively moving organelles and vesicles within the cytoplasm, for muscle contraction and the beating of cilia (Section 9.3.1.1) and flagella (Section 6.3.4.2), and for chromosomal movements. It has been estimated that up to 80% of unbound cytoplasmic proteins are not freely diffusible but are associated in some way with the cytoskeleton, and up to 20–40% of cytoplasmic water may be bound to the filaments and tubules of the cytoskeleton.\(^{939}\)

The most radical example of cytoskeletal dynamics may be the complete remodeling of the cellular microtubular array during replication—from a network radiating throughout the cell in interphase to the compact bipolar mitotic spindle during mitosis. There are five recognized classes of filaments, grouped according to their diameter and by the types of protein they contain. In order of size, starting with the thinnest, they are

1. superfine filaments,
2. microfilaments,
3. intermediate filaments,
4. muscle thick filaments, and
5. microtubules.

Microfilaments and microtubules can be rapidly assembled and disassembled, allowing a cell to modify its cytoskeletal framework according to changing requirements. The other filament types, once assembled, are less readily disassembled.

**A. Superfine Filaments** — Superfine filaments are short segments measuring 2-4 nm in diameter and up to 100 nm in length. Superfine filaments such as plectin\(^{1194,3389}\) serve to interconnect other filaments and microtubules, stabilize the nucleus,\(^{3391}\) regulate actin dynamics,\(^{3392}\) and also provide the means by which the membranes of organelles may be attached to stable or moving elements of the cytoskeleton.\(^{938,3389,3390}\)

**B. Microfilaments** — Microfilaments are the smallest of the major cytoskeletal components, measuring 5-7 nm in diameter and up to several microns in length. They are F-actin polymers of the contractile monomer protein G-actin, a single polypeptide consisting...
of 375 amino acid residues with a molecular weight of ~42,000. Actin is the most abundant protein in most cells, usually comprising more than 5% of the total cellular protein. Polymerization at the “plus” end and depolymerization at the “minus” end occur spontaneously, driven by local G-actin concentration. This produces a treadmilling action, wherein a given actin monomer is incorporated at the “plus” end, transfers slowly along the microfilament from one end to the other, and finally is lost by depolymerization at the “minus” end. Assembly speed for microfilaments of the actin cytoskeleton is 0.1-1 micron/sec. Tensile failure strength of a single strand is ~108 pN, while the compressive load force or “stall force” applied during polymerization by a single actin fiber is typically ~10 pN. Microfilaments make up a major portion of all cell cytoskeletons and are best known for their role in the contractile fibrils of muscle cells. They can form connections with the plasma membrane and thereby influence locomotion, amoeboid movement, and cytoplasmic streaming. They produce the cleavage furrows that divide the cytoplasm of cells after chromosomes have been separated by the spindle fibers during mitosis. Microfilaments also help develop and maintain cell shape by acting as tensile elements in the cellular tensegrity structure, pulling the plasma membrane and all of the cell’s internal constituents toward the nucleus at the core. Microfilaments also conduct mechanical signals throughout the cell at propagation speeds of 100-1200 m/sec, allowing 0.01-2 Hz signals to cross a cell in 2-20 nanoseconds.

Most cells have a dense network of subsurface microfilaments called the cell cortex, which includes actin-binding peripheral proteins such as spectrin (in erythrocytes) or filamin and vinculin (in fibroblasts) just below the plasma membrane. In the case of the red cell, the membrane skeleton (Fig. 8.43) is a highly organized two-dimensional triangulated network of actin oligomers tethered to spectrin tetramer filaments interconnecting ~70,000 nodes or microfilament junctional complexes comprising an equal number of triangular ~50 nm-wide meshes (mean mesh size 3000-4800 nm2). The cell cortex confers structural rigidity on the cell surface and facilitates shape changes and bulk movement. In some cells, microfilaments are ordered into long parallel bundles called stress fibers, which may span the entire length of the cell. Stress fibers also comprise the core of microvilli. Each cell has its own arrangement of microfilaments, thus no two cells are exactly alike.

Indeed, the microfilament network itself may form a complete tensegrity substructure within the larger cytoskeletal network.

C. Intermediate Filaments — Tough, insoluble intermediate filaments (IFs) have a diameter of 8-12 nm with a mean intertubule spacing of perhaps ~50-100 nm. They are the most stable of the cytoskeletal elements, not being subject to constant de-/re-polymerization, with a tensile failure strength of 20 nN per fiber (Table 9.3). With their high tensile strength and comparative positional stability, IFs act as internal guy wires to resist mechanical stress on the cell and thus are regarded as a scaffold supporting the entire cytoskeletal framework, with polysomal ribosomes located at IF junction control nodes (Fig. 8.44). IFs also have a tension-bearing role in some cells, as intermediate filaments are most extensively developed in those regions of cells most subject to mechanical stress. Another specific function of IFs is to maintain the position of the nucleus within the cell. Intermediate filaments form a ring around the nucleus with branches extending outward through the cytoplasm, and possibly extending downward into the pores of the nuclear envelope (Section 8.5.4.2) to connect with the nuclear cortex (Section 8.5.4.3). Intermediate filaments constitute ~1% of total protein, although in some cells (e.g., epidermal keratinocytes, neurons) IFs may represent up to 85% of the total protein of fully differentiated cells.

In contrast to microtubules and microfilaments, intermediate filaments differ in their composition from one tissue to another.
The known classes of IF protein (based on biochemical and immunological criteria) include the keratins of which there are at least 15 different varieties in each of the acidic and basic/neutral subclasses (in the tonofilaments of epithelial cells); vimentin, which is found in fibroblasts, connective tissue and other cells of mesenchymal origin; desmin, in muscle cells; glial fibrillary acidic (GFA) protein, in glial cells; neurofilament (NF) protein, present in the neurofilaments of nerve cells; and nuclear lamins A, B and C, which are found in the nuclear cortex of all cells. Most cells contain at least two different types of IF proteins: the three lamins in the nucleus, plus the cytoplasmic IF protein appropriate to the specific cell type. Because of this tissue specificity, cells from different tissues can be distinguished almost solely on the basis of the IF proteins present—intermediate filament typing via immunofluorescence microscopy is a useful diagnostic tool in detecting cancer and prenatal birth defects. In cyto nanorobots can use this same information to continuously verify the cell type in which they are resident.

D. Muscle Thick Filaments — Muscle thick filaments, ~15 nm in diameter and composed of the very large contractile protein myosin, are found mainly in striated muscle cells and occasionally in the cortex of nonmuscle cells. Myosin molecules not in filament form are also present in many, if not most, other cells where they interact with microfilaments to produce local forces and movements.

E. Microtubules — Microtubules are straight hollow cylinders with an outer diameter of 25 nm and an inner diameter of 15 nm. They vary greatly in length. Some are less than 200 nm long, while others, particularly in nerve cells (where they provide the framework that maintains the cell’s cylindrical shape), can be as long as 25 microns. Mean intertubule spacing is at least 200-300 nm, with a minimum radius of curvature of ~0.2 microns. Each microtubule is composed of a helical arrangement of the β-tubulin heterodimer protein, with 13 tubulin subunits in each rotation of the tight helix. The uniform orientation of the tubulin molecules confers an inherent polarity on the microtubule. Placed in a tubulin-rich environment, the “plus” end of the tubulin polymer (β-tubulin) will elongate much more rapidly than the “minus” end (α-tubulin). In the living cell, the ends of the microtubules farthest away from the center of the cell are always the “plus” ends. Microtubules radiate out as lacelike threads toward the periphery of the cell from a microtubule-organizing center (MTOC) near the nucleus. The best-known MTOC is the centrosome, which consists of granular material surrounding two centrioles (Fig. 8.36).

Microtubules are the most rigid of the cytoskeletal filaments and thus often serve as the principal cytoskeletal organizers and backbone elements (Fig. 8.44), providing (along with the ECM) the compressive elements of the cellular tensegrity structure. Tensile failure strength is ~1000 pN per fiber (Table 9.3). The flexural rigidity of individual microtubules has been directly measured as $K_{rigid} = 3.4 \times 10^{23}$ N·m², which is higher than the rigidity of actin filaments. What does this mean? For an elastic rod of length L attached to a wall by a pivot about which the rod is completely free to rotate, the magnitude of the critical buckling force of the rod is:

$$F_{buckle} = \frac{\pi^2 K_{rigid}}{L^2} \quad \text{[Eqn. 8.6]}$$

Hence a 1-micron long microtubule will not begin to buckle until an axial load of $F_{buckle} \approx 340$ pN is applied.

Microtubules define and maintain the overall shape and architecture of the cell, confer polarity on the cell, and determine the distribution of the microfilaments and intermediate filaments. By providing a radiating system of fibers to guide the movement of vesicles and other organelles, microtubules also contribute to the spatial disposition and directional movement of subcellular structures. For example, microtubules serve as “tracks” for the outward movement of the endoplasmic reticulum in growing cells, and microtubules transport membrane-bound vesicles in both directions along the axons connecting the body of the nerve cell with the synaptic knobs (axon transport) at a typical speed of ~2 microns/sec. Particles easily switch from one microtubule to another intersecting one during transport.

The structure immediately responsible for the separation of the chromosomes, the spindle fibers, is composed of microtubules, as are the organelles (the centrioles) that generate the spindle fibers at the time of cell division. Each cell has two centrioles, solid structures composed of microtubules. Prior to cell division, the centrioles replicate and then produce the intracellular architectural skeleton called the mitotic apparatus that guides the cell through mitosis. Spontaneous spatial pattern formation from oscillating microtubules has been observed, and cytoskeletal rearrangements appear to be biochemically regulated.

The microtubule elongation rate in cultured fibroblasts has been measured as ~0.06 microns/sec. The turnover time for the entire microtubular apparatus has been estimated as ~900 sec (0.25 hour) with a specific power output of ~0.6 watts/m³ averaged over the entire array. Not all cytoplasmic microtubules immediately participate in this rapid turnover during interphase—about 10% of the population seems to remain stable for at least 2 hours. Dynamic instability causes microtubule shortening in human monocytes, in the range of ~260 dimers/sec or ~0.15 microns/sec. Depolymerizing microtubules can drive kinetochore movements towards the minus end at ~0.5 microns/sec, exerting forces of up to ~100 pN.

Superfine filaments connect microtubules with each other, with intermediate filaments, and with adjacent organelles. The filaments extend from the microtubular surface and prevent direct contact with other cellular structures. This forms a narrow exclusion zone around microtubules that is devoid of other structures, and often appears as a clear halo with a radius from ~10 nm to ~50 nm in neurons. High molecular weight microtubule-associated proteins (MAPs) define zones of exclusion around microtubules and may help maintain the observed spacing between microtubules and cell organelles.

8.5.3.12 Cytonavigational Issues

Cytonavigational requirements are strongly mission-driven. A nonexhaustive list of mission classes might include:

1. examination or modification of the plasma membrane or cell cortex, including chemical testing for toxins or poisons in cell receptors and transport channels;
2. cytosolic chemical or pathogenic assay, chemical injection or extraction, or other selective cytosolic modification;
3. biocellular messaging or eavesdropping using chemical, mechanical, or other means;
4. organelle counting, dimensional measuring, and general cytocartography;
5. circumorganelle chemical assay;
6. organelle-specific surface membrane analysis or intraplasmonic chemical assay;
7. dynamic functional or structural testing of cellular components;
8. sampling, diagnosis, chemoinjection, replacement or repair operations to be performed upon an individual organelle or cytocomponent located at a specific cellular physical address;
9. cytoassembly or structural editing;
10. establishment of direct functional control over some or all of normal cellular functions, including metabolism, secretion, and mitotic cycling;
11. comprehensive cellular reconstruction;
12. long-term cytoplasmic materials or equipment storage;
13. sentinel or cytodefensive functions; or
14. activities involving the nucleus (Section 8.5.4). Each of these mission classes has very specific, and often quite different, navigational requirements.

Medical nanorobots can certainly undertake many useful tasks without physically entering the cell, relying solely upon diffusion, transport via cellular pumps, endocytosis and pinocytosis, or even nanoinjectors or manipulator appendages inserted through the plasma membrane (Section 9.4.5) while the nanorobot remains securely anchored outside the cell. However, it is well-known that motile entities are capable of entering and navigating the interiors of living cells for long periods of time without ill effect. For example, one description from early microscopic investigations included the following observation: “...lymphocytes entered the cells and circulated inside them for hours at a time. This odd relationship of lymphocytes to other cells, sometimes moving around them, sometimes entering them, they termed ‘emperiplois.’” (Emperiplois1286,1295 is a rarely observed and still poorly understood phenomenon that may only occur in pathological conditions.)

Micron-scale bacterial pathogens that invade nonphagocytic cells, securely anchored outside the cell. However, it is well-known that motile entities are capable of entering and navigating the interiors of living cells for long periods of time without ill effect. For example, one description from early microscopic investigations included the following observation: “...lymphocytes entered the cells and circulated inside them for hours at a time. This odd relationship of lymphocytes to other cells, sometimes moving around them, sometimes entering them, they termed ‘emperiplois.’” (Emperiplois1286,1295 is a rarely observed and still poorly understood phenomenon that may only occur in pathological conditions.)

Micron-scale bacterial pathogens that invade nonphagocytic cells, securely anchored outside the cell. However, it is well-known that motile entities are capable of entering and navigating the interiors of living cells for long periods of time without ill effect. For example, one description from early microscopic investigations included the following observation: “...lymphocytes entered the cells and circulated inside them for hours at a time. This odd relationship of lymphocytes to other cells, sometimes moving around them, sometimes entering them, they termed ‘emperiplois.’” (Emperiplois1286,1295 is a rarely observed and still poorly understood phenomenon that may only occur in pathological conditions.)

Micron-scale bacterial pathogens that invade nonphagocytic cells, securely anchored outside the cell. However, it is well-known that motile entities are capable of entering and navigating the interiors of living cells for long periods of time without ill effect. For example, one description from early microscopic investigations included the following observation: “...lymphocytes entered the cells and circulated inside them for hours at a time. This odd relationship of lymphocytes to other cells, sometimes moving around them, sometimes entering them, they termed ‘emperiplois.’” (Emperiplois1286,1295 is a rarely observed and still poorly understood phenomenon that may only occur in pathological conditions.)
even positional localization as crude as -1 micron resolution will uniquely locate all individual major organelles within the cytoplasm. This may be sufficient for most purposes, since random hydrodynamic flows induced by thermal fluctuations inside a cell have velocities of ~10 microns/sec in a time range of ~10 milliseconds with a characteristic length of ~1 micron.1069

A volumetric cytographic map using (100 nm)³ voxels requires 8 million 8-bit voxels or ~64 megabits of memory. A 1 micron³ nanorobot moving at ~1 micron/sec would require ~8000 sec to volumetrically survey the entire interior of a 20-micron cell using an efficient nonoverlapping scanning pattern, or ~300 sec to scan a region having the same volume as the nucleus. A volumetric map would allow computation of the shortest path to the plasma membrane in order to exit the cell quickly and with minimum disruption. Such maps might help the nanorobot to identify the axodendritic polarity of a neuron, locate the nucleus or a specific part of a membrane—for instance, a part of the membrane that is connected via intercellular junctions to adjacent cells, or which is adjacent to bone or digestive juices. Volumetric maps could also allow the nanorobot to return to the exact site of unwanted natural deposits (e.g., lipofuscin) or to sites where foreign objects are lodged, although long-term map stability is problematical.

The cytoplasmic membranes present inside a 20-micron tissue cell have a total surface area of ~180,000 micron², plus another ~280,000 micron² of cytoskeletal fiber surfaces (Table 8.17). Assuming 8-bit pixels, the totality of internal cellular surfaces could be instantaneously described using 1 micron² pixels with a 4 megabit map, or with (100 nm)² pixels using a 400 megabit map (requiring ~0.02 micron³ of hydrofluorocarbon memory tape). However, high-resolution cytographic surface survey maps are not as useful because of their short half-lives. Membrane lipids and transmembrane protein molecules have lateral diffusion speeds (Section 8.5.3.2) averaging ~3 microns/sec and ~0.02 micron/sec, respectively; giving a 0.3-50 sec half-life for a 1 micron² resolution surface map. Surface folds and other gross morphological features on the nuclear envelope, the endoplasmic reticulum and the Golgi complex may have half lives of ~10³-10⁵ sec, so maps of these features may have modest operational utility but still will have minimal archival utility. In any case, many malfunctions in these three organelles may require biochemical rather than mechanical interventions (Chapter 21).

For multiple-copy organelles, statistical sampling to acquire census data may provide the most useful information. Semipermanent structures such as mitochondria may be examined one by one to measure size and number density, to verify biochemical composition, and to test functionality. If a population of N cellular objects is independently and randomly sampled one by one by a population of Nnano in cyto nanorobots, taking a time t exam for each examination, then the probability that any one object has not yet been examined after n selections is px = (1 - N⁻¹)ⁿ [N nano].

However, the examination time T = n t exam / Nnano, and the average number of times the same object is examined npo = n N nano / N. For a cell containing N = 1000 mitochondria, mean travel distance between these organelles is Xtravel ~ 2 microns in a 20-micron cell, or travel ~ 2 sec travel time at a travel speed of ~1 micron/sec. Conservatively taking t exam ~ 10 sec, then reducing the probability of nonexamination of any mitochondrion to px ~ 0.1% requires n ~ 7,000 examinations or an average of npo = 7 examinations per mitochondrion, and requires T = 70,000 sec ~ 19 hours to complete all the examinations using a single in cyto nanorobot (Nnano = 1). A census sampling that reaches only px ~ 50% completeness requires just n ~ 700 examinations, npo = 0.7 examinations per mitochondrion and T = 7000 sec ~ 2 hours using one nanorobot. An npo ~ 1% sampling of a population of 10⁴ cytoplasmic free ribosomes requires n = 10,000 examinations and T ~ 10,000 sec ~ 3 hours assuming t exam ~ 1 sec (ttravel ~ 0.2 sec) for a single in cyto nanorobot. These times compare well even against the rapid multiplication rate of proliferating tissue cells (e.g., human embryonal cells, ~24 hours).

Post-examination individual target object tagging (e.g., using anchored messenger molecules) without physical sequestration reduces t exam only slightly but does not reduce t travel, hence has no significant impact on T. Note that successive mitochondrial samplings will generally be independent if Xtravel (~2 microns) >> ΔX_diffuse; from Eqn. 3.1, mitochondrial diffusion displacement is of order ΔX_diffuse ~ 20 nm, taking τ = t exam ~ 10 sec, R ~ 1 micron, and η ~ 10 kg/m·sec (in cyto, at 310 K; Table 9.4), hence the independence condition is usually satisfied.

8.5.4 Nucleography

Nucleography is the “geography” of the cell nucleus. The nucleus, 5-8 microns in diameter for a 20 micron tissue cell and up to 10 microns for a fibroblast cell, is the largest cellular organelle and the only one that is voluminous enough, in theory, to admit a micron-scale medical nanorobot into its interior. The nucleus is usually a large spherical or ovoid structure surrounded by its own nuclear membrane, although its shape generally conforms to the shape of the cell. For example, if a cell is elongated, the nucleus may be extended as well.
Almost all cells contain a single nucleus, whose primary function is the storage and expression of genetic information. However, a few cell types have multiple nuclei of similar size, such as skeletal muscle cells, osteoclasts, megakaryocytes, and some hepatocytes. A few cell types have no nucleus, such as red blood cells, platelets, keratinized squamous epidermal cells, and lens fibers.

In 1998 the finer details of nuclear structure were just beginning to be understood. Thus the following discussion of nucleography must be regarded as a very tentative work in progress.

### 8.5.4.1 Nuclear Envelope

The nuclear envelope enclosing the nucleus is a lipid bilayer similar in structure to the cell membrane, except that it is a double-layered membrane which is topologically more convenient for dissolution during mitosis and subsequent reassembly from vesicles. Each of the two lipid bilayer membranes is 7-8 nm thick. The outer nuclear membrane is occasionally continuous with the rough endoplasmic reticulum and is almost entirely surrounded by it (Fig. 8.45). Like membrane is occasionally continuous with the rough endoplasmic reticulum and is almost entirely surrounded by it (Fig. 8.45). Like membrane which is topologically more convenient for dissolution during mitosis and subsequent reassembly from vesicles. Each of the two lipid bilayer membranes is 7-8 nm thick. The outer nuclear membrane is occasionally continuous with the rough endoplasmic reticulum and is almost entirely surrounded by it (Fig. 8.45).

Intermediate filaments extend outward from the outer membrane into the cytoplasm, anchored on the other end to the plasma membrane or other organelles, thus positioning the nucleus firmly within the cell and increasing its mechanical stiffness almost tenfold.

The perinuclear space (or perinuclear cisterna) between the two membranes ranges in width from 10-70 nm but is usually a gap of 20-40 nm. This fluid-filled compartment is continuous with the cisternae of the rough ER (Fig. 8.46), thus providing one possible avenue for transporting substances between the nucleus and different parts of the cytoplasmic compartment.

The nuclear envelope disassembles at the onset of mitosis and is reassembled at the end of mitosis.

### 8.5.4.2 Nuclear Pore Complexes

The most distinctive feature of the nuclear envelope is the presence of numerous nuclear pores (Fig. 8.46), small cylindrical channels with eightfold symmetry that extend through both membranes and provide direct contact between cytoplasm and nucleoplasm. Each pore complex marks a point of fusion between the inner and outer membranes. Elements of the cytoskeleton appear to be attached to many pores, possibly allowing direct mechanical regulation of pore activity.

Each nuclear pore complex is a huge multimolecular assemblage measuring 70-90 nm in diameter, with a mass of 125 million daltons, ~34 times the size of a ribosome. Up to 100 different nucleoporin protein molecules make up the structure. Early experiments with passive gold particles showed that cytoplasmic particles with diameters of 5-6 nm passed into the nucleus in ~200 sec, those with diameters of 9-10 nm took ~10^4 sec, but particles larger than 15 nm didn't seem to enter at all. Closer examination has revealed that the pores are actually large enough to allow the passage of substrates as large as 23-26 nm, but this is still much too narrow for nanorobots or their flexible processes to pass through without damaging the mechanism. The nuclear localization sequence (NLS), a molecular tag consisting of 1-2 short sequences of amino acids, marks cytoplasmic proteins for active transport through the nuclear pores. Small (~40 nm) arm-like import receptors (cytoplasmic filaments) ringing the mouth of the pore bind to a cargo protein tagged with an NLS, then flex toward the pore to shove the cargo into the mouth.

The density of pores across the surface of the nuclear envelope varies greatly, depending mainly on cell type and the amount of RNA being exported to the cytoplasm. Values range from 3-4 pores/micron² in some white cells up to 50 pores/micron² in oocytes and a theoretical maximum density of 60 pores/micron². A typical 20 micron human cell has 2000-4000 pores embedded in its nuclear surface, a mean density of 10-20 pores/micron². Pore structures may protrude at most ~100 nm into the nucleoplasmic space.

### 8.5.4.3 Nuclear Cortex

The nuclear cortex is an electron-dense layer of intermediate filaments (composed of the nuclear lamins common to most cell types) on the nucleoplasmic side of the inner nuclear membrane. The cortex, also called the nuclear lamina or karyoskeleton, is up to 30-40 nm thick in some cells but is difficult to detect in others. Its proteinaceous fibers are arranged in whorls that may serve to funnel materials to the nuclear pores for export to the cytoplasm. These fibers may also be involved in pore formation. The nuclear cortex helps to determine nuclear shape, and also binds to specific sites on chromatin, thereby guiding the interactions of chromatin with the nuclear envelope.

### 8.5.4.4 Nucleoplasm and Chromatin

The nucleoplasm is the semifluid matrix in the interior of the nucleus. It contains some condensed but mostly extended chromatin (called heterochromatin and euchromatin, respectively), as well as a structural nuclear matrix of nonchromatin (mostly protein) material (Section 8.5.4.6). The chromatin represents chromosomes as they exist between cell divisions. Chromosomes assume a highly condensed (compact) state as the cell prepares to divide, but after mitosis most of the chromosomes relax into a highly extended state. The nucleus of the human cell contains 46 chromosomes of varying lengths, in 23 pairs. Each of these in turn are composed principally of a single deoxyribonucleic acid (DNA) molecule. The DNA contains the genes of the cell, and all ~100,000 genes are represented, though not expressed, in each nucleated cell. The nucleosol, or fluid com-

---

**Fig. 8.47. Schematic of chromatin distribution in the nucleus (redrawn from Alberts et al531).**
ponent of the nucleoplasm, contains salts, nutrients, and other needed biochemicals. A number of different granules are also present.938

During interphase (e.g., between cell divisions), individual chromosomes occupy compact, discrete territories within the nucleus that may range up to 4 microns in diameter (Fig. 8.47).2464-2467,3410-3417 The structure and location of these territories is specific for both cell type and mitotic stage,3412,3413 and may be arranged in the same spatial order as is found in the wheel-shaped ring aggregate known as the chromosome rosette at the time of mitotic prometaphase.1060 It has been proposed that active genes are preferentially localized to the periphery of the chromosome territories; RNA templates would be preferentially produced at the surfaces of these territories and then shed into interchromosomal domain channels for further processing and transport.1529 Others have argued against this.3414 But knowledge of any such spatial orderings, which was woefully incomplete in 1998, might permit intranuclear nanorobots to approximate their position inside the nucleus once the locations of three or more specific chromosomes have been definitively established. The diameter of a territory \( D_{\text{territory}} \sim D_{\text{nucleus}} \left( \frac{c_{\text{chromosome}}}{c_{\text{genome}}} \right)^{1/3} \sim 3 \) microns, where \( c_{\text{chromosome}} \) is chromosome size and \( c_{\text{genome}} \) is genome size, both measured in base pairs, and \( D_{\text{nucleus}} \) is the diameter of the nucleus.2464

Note, however, that these territories are not rigid. Changes in the relative positions of chromosomal territories often occur at ~0.3-0.4 nm/sec, and intraterritorial movement and flexing of subchromosomal foci measuring 400-800 nm in diameter have also been observed.1529,3415 Both cytoplasm and nucleoplasm may contain numerous as yet undiscovered intricate substructures that could provide many new navigational aids. An important task for early nanomedicine-oriented research will be to fully explore and elucidate this fine structure, and to determine whether or not it is stable enough to be relied upon in any way for intracellular navigation.

In its most relaxed state, chromatin resembles a network of bumpy threads weaving their way through the nucleoplasm. Chromatin is composed of roughly equal amounts of negatively charged DNA (comprising the chromosomes) and globular histone proteins (basic proteins which carry a positive charge at the normal pH found in the cell).938 Nucleosomes, the fundamental units of chromatin, are spherical clusters of eight histone proteins, connected like beads on a string by a DNA molecule that winds around each of them. The average cell nucleus contains 25 million nucleosomes, also called histone octamers. Each nucleosome is encircled by 146 base pairs of DNA. Nucleosomes have a mass of 206,000 daltons. About half of nucleosome mass is protein and half is DNA. Each human chromosome contains DNA with an average contour length of ~75 mm. (See Chapter 20 for more details.)

8.5.4.5 Nucleolus

The largest and most prominent “nuclear organelle” is the nucleolus, a highly coiled structure associated with numerous particles but not surrounded by a membrane.1141 The nucleolus is a ribosome-manufacturing machine. Assembly of precursor ribosomal subunits within the nucleolus requires ~1800 sec, while the complete assembly of a large ribosomal subunit (needing only protein to make a completed ribosome) takes ~3600 sec.531

The nucleolus is composed of DNA, RNA, and proteins. It also has a granular component (each granule ~150 nm thick) and a fibrillar component, and a variable internal structure.1141 The granular component consists of ~15-nm particles that are ribosomal subunits in the process of maturation. The fibrillar component consists of rRNA molecules that have already become associated with proteins to form fibrils with a thickness of ~5 nm. The size of the nucleolus correlates with its level of activity. In cells characterized by a high rate of protein synthesis and hence by the need for many ribosomes, the nucleolus can occupy 20-25% of nuclear volume (3-5 micron diameter in a 20-micron cell), mostly comprised of the granular component. In less active cells, the nucleolus is much smaller—as small as 0.5 micron in a mature lymphocyte.1141 Nucleoli are frequently located at or near the nuclear envelope, adhering directly to the
nuclear lamina or attaching to it by a pedicle. In nuclei having a centrally located nucleolus, the nuclear envelope is folded to form a nucleolar canal that is in direct contact with the nucleolus.\textsuperscript{1141}

Most human nuclei contain only one nucleolus, except for liver cell nuclei which may contain more than one nucleolus\textsuperscript{935} and cultured HeLa (cancer) cells which may have up to six.\textsuperscript{1141} The number of nucleoli in a eukaryotic cell nucleus normally is determined by the number of chromosomes with secondary constrictions, or nucleolus organizer regions (NORs). The human genome contains five NORs per haploid chromosome set, or 10 NORs per diploid nucleus, each located near the tip of a chromosome. However, instead of 10 separate nucleoli, the typical human nucleus contains a single large nucleolus representing the fusion of loops of chromatin from the 10 separate chromosomes; these enlarge, eventually fusing into the single large nucleolus when the cell enters mitosis.\textsuperscript{1141} As rRNA synthesis resumes, mitosis thus has no nucleolus and synthesizes no rRNA. Once mitosis is complete, the nucleolus reappears. As rRNA synthesis resumes, the nucleolus organizer regions (NORs) are activated, leading to the reappearance of the nucleolus.\textsuperscript{1141}

The nuclear matrix organizes thousands of DNA replication sites during mitosis. At left, RWGL model of chromosome 15 territory; at right, MLS model of chromosome 15 territory (courtesy of Tobias A. Knoch, Division of Biophysics of Macromolecules, German Cancer Research Centre, Heidelberg, Germany ©1999).\textsuperscript{2465}

8.5.4.6 Nuclear Matrix and Transcript Domains

Like the cytoskeleton of the cell, the nucleus is thought to possess a nuclear skeleton comprised of a network of insoluble protein fibers known as the nuclear matrix.\textsuperscript{1134,3418-3423} This finely branched meshwork of superfine filaments attaches to the nuclear lamina or cortex (Section 8.5.4.3) and extends to the nucleolus. Assuming \(10^6\) filaments each averaging \(50\) nm long gives a total fiber length \(L_{\text{fiber}} \approx 50,000\) microns, which for a \(D = 8\) micron diameter nucleus implies a mean grid spacing of:

\[
L_{\text{grid}} = \left( \frac{D^3}{L_{\text{fiber}}} \right)^{1/2} \approx 50 \text{ nm} \tag{Eqn. 8.7}
\]

Typical mesh spacings seen in electron micrographs do indeed range from \(50-100\) nm.\textsuperscript{1132} The filaments are constructed from a variety of nuclear matrix proteins (NMPs). More than a dozen NMPs have been identified, many of them cell-specific.\textsuperscript{3421,3422} Also, NMPs released by cancer cells differ from those released by normal cells of the same type.\textsuperscript{3422,3423}

The nuclear matrix organizes thousands of DNA replication sites during mitosis, and provides architectural order during interphase for RNA metabolism (wherein mRNA molecules are synthesized, spliced, and made ready for export to the protein-synthesizing machinery in the cytoplasm).\textsuperscript{1135} The regions where RNA metabolism occurs are called transcript domains.\textsuperscript{1133,1135,2021} A human nucleus typically has 20-40 transcript domains, each ranging in diameter from \(0.5\) to \(3\) microns and aligned in a planar horizontal array in the “lower” portion of the nucleus (Fig. 8.49); “lower” is toward the ventral surface of a cell that is adherent to the glass plate used in the microscopic study.\textsuperscript{1135} A single nucleolus occupies the “upper” or “lower” portions of the nucleus with equal frequency; its position is established by sequence-specific DNA binding proteins.\textsuperscript{1139} The outermost edge of a transcript domain lies an average of \(0.8\) microns (range \(0.5-1.2\) microns) from the inside surface of the nuclear cortex or nuclear envelope. Located in the core of each domain is a protein called SC-35 that is required for the assembly of the molecular machinery that splices the introns out of mRNA.\textsuperscript{3424} Each domain is surrounded by a discrete border of high-density chromatin—where a given chromosome always occupies the same general location in the nucleus at each mitotic phase. All 20-40 transcript domains collectively occupy \(5\%\) of the total nuclear volume, whereas the specific RNA accumulations occupy no more than \(1\%\).

However, nuclear matrix elements may not be necessary for physical support inside each chromosomal territory. Several theoretic descriptions of chromosome structure have been compared to experimental data, and one such description, the Random-Walk/Giant-Loop (RWGL) model,\textsuperscript{2466,3416} proposes that huge 3-5 million base-pair (Mbp) loops are bound to a nuclear matrix that permeates the region. More recently, polymer dynamics simulations of chromosomal territories performed by Munkel, Langowski and Knoch\textsuperscript{1013,2464,2465,3417} suggest that a Multi-Loop-Subcompartment (MLS) model is in better accord with experiment. In the MLS model, chromatin fiber is folded into \(120\) Kbp sized loops which are organized into rosettes of 1-2 Mbp. The rosettes (which may be structural units for replication, etc.) are interconnected by a piece of chromatin of similar base pair content so that no protein matrix is needed for structural support. The MLS model agrees with the metaphase organization proposed by Pienta and Coffey.\textsuperscript{2467} The MLS rosettes are clearly visible in Figure 8.50, which shows human chromosome 15 in its highly condensed metaphase state, as it might appear during mitosis. Figure 8.51 shows two polymer dynamics simulations of the 3-5 micron\textsuperscript{3} territory of human chromosome 15
in its relaxed state between cell divisions, using the RWGL model (left) and the MLS model (right). The numerous apparently matrix-free voids up to -0.5 micron in diameter that appear in the MLS simulation are presumably largely filled by another chromosome (not shown). (Sometimes the two copies of chromosome 15 are adjacent, but most of the time they are separated.) However, numerous -0.1 micron voids do exist (c.f. 100-nm keyhole passage; Section 5.3.1.2), and dynamic processes may temporarily open still larger spaces possibly large enough for a medical nanorobot to seek slow, relatively safe and unobstructed passage, but the theoretical pathways of spheres, diffusion and percolation through territories remain to be studied [T. Knoch, personal communication, 1999].

8.5.4.7 Nucleonavigational Issues

Cytoplasmic medical nanorobots can perform many useful tasks without entering the nucleus. Such tasks might include:

1. physical mapping and compositional analysis of the nuclear envelope;
2. monitoring of nuclear pore traffic;
3. near-complete regulation of nuclear pore traffic using multiple manipulators or other devices;
4. monitoring, initiating, or modifying cytoskeletally-mediated mechanical signal transduction into the nuclear interior;
5. injection of enzymes, RNA or DNA fragments, or other bioactive materials through nuclear pores using hollow nanoinjectors; or
6. partial or complete nucleoplasmic replacement using artificial chromatin detachment enzymes dispensed by multiple injection and extraction nanorobots positioned at the nuclear pores, operated simultaneously and located antipodally around the nucleus to establish flowthrough.

Entering the nucleus is somewhat more difficult than entering the cytoplasmic space through the plasma membrane (Section 9.4.5.7). One reason is that the nucleus is almost completely surrounded by the membranes of the endoplasmic reticulum (Fig. 8.45), and the nearest of these membranes may lie only a few hundred nanometers from the outermost surface of the nuclear envelope (Fig. 8.46). The nuclear envelope is also a double-walled membrane. Another reason why nuclear entry is difficult is that physical forces applied to cytoskeletal elements in the immediate vicinity of the nucleus may trigger unwanted transcriptional, structural, or metabolic responses from within the nucleus. Additionally, the 10-100 nm grid size of the nuclear matrix filaments and the DNA rosette loops may allow little free maneuvering room for nanorobots, which may find it difficult to avoid tearing the nucleoskeleton during passage.

Given that the nucleus of a 20 micron human cell is only ~8 microns in diameter and thus encloses a volume of at most ~268 micron^3 (Table 8.17), there is precious little maneuvering room for a medical nanorobot that may be 1-30 micron^3 in size. Once inside the nucleus, a principle safety concern must be avoiding damage to the relaxed euchromatin strands that permeate the nucleoplasmic space. Only ~0.1 pN of force is required to move chromosomes around the nucleus at ~0.1 microns/sec during mitosis. DNA base pair hydrogen bonds may be pulled apart with 70-75 pN/bond of force, although the DNA backbone itself can withstand up to ~10 nN in tension; most conservatively, nanorobots should never apply tensile or shear forces greater than ~50 pN to chromatin strands during nucleoplasmic locomotion. Care should also be taken to design nanorobot exterior surfaces that are free of sharp edges and which possess electrochemical characteristics that tend to be nonattractive to DNA and to nucleosomal components. Since DNA is negatively charged and histones are positively charged (to bind the DNA), and since the lipid membranes are internally hydrophobic but polar on the surface, the ideal nanorobot exterior might require a surface of alternating charges which is nonhydrophobic and neutral with regard to DNA, histones, and lipid surfaces.

How can a nanorobot determine its intranuclear position? Chemonavigation is one crude approach. For example, a large number of stable RNA species—most of the small nuclear ribonucleoproteins (snRNPs) essential for pre-messenger RNA splicing—are found in the nucleoplasm, the cytoplasm, or both. They range in size from 90-300 nucleotides and are present in 10^-10^-6 copies per cell (Table 8.19). snRNPs from HeLa (cancer line) cells contain ~40 proteins, 8 of which are shared by all snRNPs while the remaining ~32 proteins are snRNP-type specific. Nucleolus-specific proteins include UBF (94-97 kD), Ki-67 (345-390 kD), fibrillarin (34 kD), numatrin or B23 (38 kD), and nucleolin (100-110 kD). Different nucleolar markers appear during mitosis. Detection of these species via chemosensors allows a nanorobot to distinguish cytoplasmic, nucleoplasmic, and nucleolar spaces, and possibly may permit even finer localizations. Various nucleic and protein products are continuously being transcribed from various chromosomes, whose distinct territories within the nucleus are relatively fixed during interphase (Sections 8.5.4.4 and 8.5.4.6). Detection of specific proteins and intermediate RNA transcription products in the nucleolus thus may allow chromosomal segment localization without having to actually sequence the segment (e.g., in early-generation medical nanorobotic systems).

Nanorobots requiring a 100-nm three-dimensional Cartesian navigation grid inside the nucleus may exploit the relatively uniform acoustic characteristics of the extranucleolar nucleoplasm by tethering a minimum of four small acoustic beacons tetrahedrally placed near the inside surface of the nuclear cortex at the antipodes of three transnuclear orthogonal coordinate axes. Each beacon emits brief ~1 GHz chirps in a distinguishable format unique to each beacon, at a very low duty cycle to avoid excessive energy consumption. From Eqn. 4.52, each chirp attenuates ~10^-3 in amplitude per 100 nm of travel through the nucleoplasmic fluid. Hence an acoustic detector sensitive to 10^-6 atm pressure changes can establish nanorobot position relative to each beacon to ~100 nm accuracy if the beacon output amplitude during the brief chirp is ~10^-3 atm.

| Table 8.19. Some Species of Stable RNAs and Their Localizations in Human Cells |
|-----------------|---------------|----------------------|
| RNA Name | Length ([# of Nucleotides]) | Number of Molecules per Cell | Localization in Cell or Nucleus |
| U3 | 216 | 3 x 10^5 | Perichromatin granules |
| U6 | 106 | 3 x 10^5 | Nucleolus |
| U1 | 165 | 1 x 10^6 | Nucleoplasm (heterogeneous nuclear RNA) |
| U2 | 188 | 5 x 10^5 | Nucleoplasm |
| U4 | 139 | 1 x 10^5 | Nucleoplasm |
| U5 | 118 | 2 x 10^5 | Nucleoplasm |
| 7-3 | 300 | 2 x 10^5 | Nucleus |
| 4.55 | 91-95 | 3 x 10^7 | Nucleus and cytoplasm |
| 7S | 280 | 5 x 10^7 | Nucleus and cytoplasm |
| 7-2 | 290 | 1 x 10^7 | Nucleus and cytoplasm |
This should be low enough to avoid any possibility of cavitational damage to the chromatin or nuclear matrix. As a practical matter, additional beacons may be needed to reduce measurement uncertainties caused by nuclearplasmic nonuniformities and post-placement beacon positional shifts.

8.6 Ex Vivo Navigation

In this final Section of Chapter 8, we shall discuss ex vivo navigation—the task of navigating the spaces external to the human body. A comprehensive treatment is beyond the scope of this introductory text, so the following discourse is necessarily brief. From the nanomedical perspective, there are two distinct regimes in which ex vivo navigation may be necessary or useful: the epidermal regime and the exodermal regime.

8.6.1 Epidermal Navigation

Nanorobots traversing the epidermal regime (Section 9.5.2) must ascertain their position on the surface of the skin or in the hair. The epidermis averages ~100 microns thick but ranges up to 1000-1500 microns thick on the palms, soles, and other areas regularly subjected to rubbing or pressure. There are no blood vessels in the epidermis, so nanorobots may not be able to rely upon an internal microtransponder network or related navigational facilities.

The simplest solution to epidermal navigation is basic map-following. For example, the adult body has 2-4 million sweat glands (100-200/cm²), mean separation ~800 microns). Most of these are saline- and urea-emitting eccrine glands ~20 microns in diameter, prevalent on the back, chest, forehead, palms, and soles. A smaller number of scent-secreting apocrine glands ~200 microns in diameter are concentrated at the underarms, nipples, genitals, and anus. Apocrine secretions are promptly broken down by skin bacteria to make odoriferous androstenone (“stale urine”), androstenol (“musky”), and isovaleric acid (“sweaty goatlike”). Apocrine glands are persistent dermal landmarks that are readily mapped, along with their distinctive chemical secretions (including 0.028-0.4 mg/cm³ glucose). A positional or chemo-navigational sweat gland map of the ~2 m² human skin surface to ~800 micron resolution requires ~3 million pixels or 24 megabits of data storage assuming 8-bit pixels.

An alternative or supplementary system is a follicular or crinal map of 50-100 micron diameter hair shafts that are irregularly distributed over the surface of the skin. Follicular map elements are reasonably stable due to subdermal anchoring. Hair follicles normally are most numerous on the scalp (~125,000 hairs or ~200/cm², with a ~700 micron mean separation between shafts and a mean natural life of 0.4-4 years), on the axillary (underarm), perineal and pubic regions, on the eyebrows and eyelids of both genders (each eyebrow has ~600 hairs with an average life of 112 days), and in a highly variable number density on the face, chest, arms and legs, most notably in males. An epidermal follicular map recording the locations of ~200,000 individual hair shafts using (700 micron) 8-bit pixels requires 4 million pixels and 32 megabits of data storage. Crinal navigation is readily accommodated as the nanorobot uses dead reckoning to track its position while traversing individually addressed hair shafts.

Epidermal map reliability is enhanced by noting numerous landmarks that may be regarded as “permanent” features on the timescale of nanorobotic actions. Such features might include scars, birthmarks, tattoos, warts, corns, cysts, nevi, small skin flaps, or dematofibromas. Stretch marks are found over the abdomen and breasts during pregnancy, over the abdomen in cases of obesity or fluid retention, on the necks of goiter patients, over the loins of patients suffering from disorders affecting the sacroiliac region, and over any distended organ. Epidermal maps are further enhanced using dermatoglyphics, the study of the patterns of ridges and lines on the skin of the fingers, palms, wrists, toes, soles, and neck. Many of these patterns are unique to a given individual (thus may be used to verify patient identity, e.g., fingerprints) and are persistent enough over time to serve as the basis for stable epidermal maps. Dermatoglyphics already plays a role in traditional diagnosis—Down’s syndrome produces a characteristic mid-palm crease, and infant palm prints can reveal telltale signs of congenital heart defect. Dermal ridges are also a sexually dimorphic trait—males have more ridges than females, and both sexes usually have more ridges on the right hand than on the left hand. In 1998, mole (nevi) photographic surveillance was a preventative skin mapping procedure to detect the onset of melanomas.

Less permanent but still useful epidermal landmarks include wounds, blisters, sunburns, eczematous or psoriatic regions, or chemical stains. Epidermal microbiotigraphic maps (Section 8.4.4) are also possible, but microbial populations usually cluster near apocrine glands and thus are unlikely to provide significant additional navigational information.

Skin temperature varies greatly across the epidermis (Section 8.4.1). The nonuniform number density of sweat glands in the skin produces highly localized periglandular temperature, moisture, and chemical compositional variations. However, these variations probably add little new information and are more difficult to interpret because readings may be strongly influenced by numerous complicating factors such as clothing and cosmetics, general physical activity level, localized physical activity (e.g., clapping hands), nonuniform sunlight-mediated heating, airflow and winds, immersion of body parts in water, and so forth. Nanorobots attempting to use this kind of information will need a very sophisticated global knowledge of somatic status.

Epidermal maps more accurate than ~1 mm might not prove particularly useful because natural movement of highly elastic dermal tissues may inject measurement errors of this magnitude into relative positional data used to locate adjacent glands, follicles or other epidermal landmarks.

8.6.2 Exodermal Navigation

Exodermal navigation encompasses an extremely broad range of operating environments. Most nanomedically relevant are nanorobots resident on or in droplets or boluses of sweat, saliva, mucus, epidermal flakes, hairs, ejaculate, urine or feces that are discharged from the body. Exodermal navigation also encompasses all medically relevant environments such as:

1. clothing or bedding;
2. the surfaces and interiors of furniture (upon which fingerprints typically stand ~5 microns tall), domiciles, vehicles, food, cosmetics, and refuse; and
3. all possible locales including land (e.g., sidewalks, roads, public buildings), water (e.g., drinking glasses containing beverage, bathtubs, swimming pools, rivers), air (e.g., exhaled nanorobots or airborne elements of nanorobotic personal defensive systems) and even vacuum (e.g., spaceborne nanorobots). Each environment poses unique navigational challenges, a complete treatment of which is quite beyond the scope of this book.

The airborne nanorobot (Section 9.5.3) is an interesting example of these unique challenges. Airborne nanorobots can identify their host patient by chemical signature, much like a bloodhound or
mosquito following its quarry’s scent.3352 Such chemical signatures or “odortypes” may include:

1. naturally-produced “baseline” chemical scents;
2. behaviorally-related scents which may appear or intensify during specific events such as heavy exercise, fear reactions (e.g., emotional excitement alone can increase the sweat rate by ~50%), defecation or flatulence, sexual activity, intoxication, and the like;
3. artificial scents such as perfumes, colognes, cosmetics and deodorants; and
4. artificial molecular taggants specially designed to simplify the recognition task, as for instance an odorless, volatile, digitally-encoded messenger molecule emitted from an external facility that is controlled by the patient.

Airborne nanorobots can stationkeep in the vicinity of the host patient by acoustic homing on a coded ultrasonic beacon worn by the patient, all of whose emanations are inaudible to the human ear. From Eqn. 4.52 the amplitude of a ν = 100 KHz acoustic wave passing through STP air over a range of x = 1 meter only attenuates to ~25% of the original amplitude. Operational broadcast information is readily passed from patient to circumcorporeal aerial nanorobots using other similar acoustic channels (Section 7.4.8). Ex vivo nanorobots can also detect normal conversational speech at a range of ~2 meters using >2.4 micron³ pressure sensors (Section 4.9.1.6).

Airborne nanorobots can navigate and avoid no-fly zones (Section 9.5.3.6) by various methods. For instance, a flying nanorobot approaching an infinite planar sheet of 308 K (35˚C) human flesh would detect spatially anisotropic thermal emissions (peak wavelength λ_max = 9.35 microns, ν_max = 32 THz; Eqn. 6.20) in a response time of t_{meas} ~ 5 nanosec, assuming a photosensor area A_e = 1 micron², SNR = 2 (~7 photons per detection), and dermal radiative intensity I_d ~ 30 W/m² (see Eqn. 4.59). (Mosquitoes register a ~0.0006% temperature differential at a 1-cm distance from human skin in real time during flight; Section 4.6.1.)

As another example, all aerial nanorobots can continuously transmit relative skin-proximity data to their neighbors, allowing each device within a virtual “warning lattice” to estimate its rate of approach to the nearest prohibited surface. Consider a grid of 200 million nanorobotic acoustic buoys deposited on the ~2 m² dermal surface of a nude human body. The buoys are spaced 100 microns apart, and each device has a 10-micron-diameter acoustic radiator operating on a 16.7% duty cycle at 4.2 MHz, giving a ~10 microsec signal repeat time using 7-wave signal packets (SNR = 2). An aerial nanorobot with an acoustic receiver of equal size can detect a warning buoy signal at a distance of 100 microns from the skin-resident grid, assuming a continuous buoy broadcast power of ~35 pW/buoy through air at 298 K (25˚C), giving a total grid emission power of 7 milliwatts. Note that a nanorobot flying at 10 m/sec travels 100 microns in 10 microsec. Monitoring Doppler shifts in received frequency allows only the largest changes in relative velocity to be measured over multiple packet cycles in a usefully short period of time. At 4.2 MHz, summing ~50 packets (~500 microsec receipt time) allows a relative velocity change of ~1 meter/sec to be detected (Section 4.3.2).

Direct acoustic air ranging by individual flying nanorobots (i.e., echolocation) is also feasible but should be combined with a “warning lattice” system to allow neighbors and prohibited surfaces to be efficiently distinguished.

Power requirements may be estimated using Eqn. 7.22 by doubling X_path in the exponential, multiplying P_{comm} by (X_{path}/r_{antenna})² to account for the doubled path length, and dividing P_{comm} by the coefficient of reflection at the air/skin interface (0.9995; Section 4.9.1.6). A 23.1% duty cycle at 3 MHz (again giving a ~10 microsec repeat time at SNR = 2) allows aerial nanorobots to detect an approaching skin surface from a distance of 100 microns, while consuming a continuous power draw of ~29,000 pW for a 10-micron diameter receiver or ~230 pW for a 20-micron receiver, in air at 298 K.
9.1 Nanorobot Dexterity and Mobility

Manipulation and mobility are crucial basic capabilities in most classes of medical nanodevices. Manipulation includes handling fluids, biological objects such as tissue matrix fibers or cellular elements, and nanomachines or their components. Physicians must be able to direct tissue- or cell-repair nanorobots to travel to a specific site where treatment is required, and once there, to manipulate the local environment to achieve the desired results. Nanodevice mobility in vivo makes possible the rapid reconfiguration of nanomedical communication, navigation, and power systems, while ex vivo mobility allows the design of more robust diagnostic and personal defensive systems. In vivo locomotion also permits precise mapping of the internal regions of the human body across many size and time scales, both for diagnostic and for therapeutic purposes.

This Chapter opens with a discussion of adhesion forces at the molecular level and the performance of nanoscale fluid pumps and fluidic circuits (Section 9.2). Several useful classes of nanomanipulators are then presented, along with various tool tips and manipulator configurations such as massively parallel manipulator arrays (Section 9.3). Techniques of in vivo locomotion are next described in the context of biofluid and nanodevice rheology, including bloodstream streaming, cell walking and anchoring, tissue diving, cell penetration, intracellular mobility, and cytocarriage (Section 9.4). The Chapter concludes with a brief consideration of ex vivo locomotion (Section 9.5).

Our evaluation of the biocompatibility of various manipulation and propulsion systems for medical nanodevices is deferred to Chapter 15.

Finally, the reader should be aware that as of 1999, much of the available experimental data on cellular mechanics and forces in biomolecular systems was still very incomplete and ill-defined. There remained significant uncertainties in the reported numbers, many of which were of necessity compiled from large-area or bulk measurements, or for just one type of easily-studied cell, cell membrane, transmembrane protein, or receptor, and thus might prove inaccurate when applied to other, seemingly similar, cellular systems. Thus the calculations, estimates, and conclusions offered here are highly tentative and must be applied with great caution.

9.2 Adhesion and Fluid Transport

An understanding of surface forces is an essential preliminary to any study of mechanisms of manipulation or locomotion. Such adhesive forces become important at the submicron scale where nanorobotic parts are meshing with and sliding against one another. Surface adhesion forces may cause tools, parts or workpieces to stick together. Surface forces may cause airborne or solvent-immersed nanodevices to adhere to container walls, to other dry or humid surfaces, or to each other possibly causing clumping. For a 1-micron object, the adhesive forces may exceed gravitational and inertial forces by a factor of 10^6 or more. Capillary forces are also important in fluid transfers inside nanodevices, in fluid transfers between nanodevices or between nanodevices and their operating environment, and during nanodevice tasks requiring locomotion through, or limb/object manipulation in, a watery environment. Surface tension forces are especially relevant during locomotion missions requiring passage through air/water interfaces, as might occur in the lungs.

Bowling classifies adhesive forces between objects into three categories. The first category includes long-range attractive interactions that may bring a particle to a surface and establish the adhesion contact area, such as van der Waals forces (Section 9.2.1), electrostatic forces (Section 9.2.2), and magnetic forces (Sections 4.7.2 and 5.4.1). A second category of forces are the interfacial reactions that help to define the adhesion area, most importantly the capillary forces arising from the establishment of liquid or solid bridges between particle and surface (Section 9.2.3), but also including other effects such as sintering (diffusion and condensation), diffusive mixing, mutual dissolution and surface alloying, which will not be considered further here. In the third category are very short-range interactions that may strengthen adhesion after an adhesive contact area has already formed; such forces include chemical bonds of all types and various noncovalent bonds such as hydrogen bonds that have already been described in Section 3.5.1.

The rest of this Section is concerned with the transport and manipulation of fluids. After a brief discussion of capillarity theory and the unique problems associated with nanotubular fluid flow (Section 9.2.4), this Section presents basic aspects of continuum fluid flow (Section 9.2.5), describes the problems of effervescence and crystallescence during offloading (Section 9.2.6), and concludes by examining various design issues in submicron-scale fluid pumping, plumbing, mixing, and containerization (Section 9.2.7).

9.2.1 Van der Waals Adhesion Forces

In dry or vacuum environments, adhesion forces between <100 micron diameter particles and surfaces (nanorobot and other) separated by 100 nm or less are usually dominated by van der Waals interactions. Consider a spherical particle of radius r lying a small distance z_{sep} \leq 100 nm from a flat plate of surface area A >> \pi r^2. As a crude approximation, ignoring retardation effects that may become important in solution for z_{sep} > 5 nm, the van der Waals adhesive force is approximated by:

\[ F_{vdW} = \frac{H r}{6 z_{sep}^2} \quad \text{[Eqn. 9.1]} \]
for $z_{sep} \ll r$, where $H$ is the Hamaker constant for the interaction. (Eqn. 9.1 also applies to two crossed cylinders of equal radius.) The Hamaker constants for the materials comprising the sphere, the plate, and the surrounding medium are $H_s$, $H_p$, and $H_m$, respectively (Table 9.1), giving:

$$H = (H_s^{1/2} - H_m^{1/2}) (H_p^{1/2} - H_m^{1/2})$$  \text{(Eqn. 9.2)}

The adhesion separation distance $z_{sep}$ is often taken as -0.4 nm for particles in intimate (but chemically unbonded) contact with a surface. Hence the force required to overcome the van der Waals adhesive attraction of a perfectly rigid $r = 0.5$ micron particle adhered to a diamond plate in vacuo (with $z_{sep} = 0.4$ nm, $H = 340 \times 10^{-12}$) is $F_{vdW} = 180$ nN.

Real particles are not perfectly rigid but deform under the influence of this adhesive force. If the adhesion surface area is increased from point contact to a circle of radius $r_{adhesion}$, then the van der Waals adhesion force increases to:

$$F_{vdW} = \left( \frac{H}{6 \pi r_{sep}^2} + \frac{H r_{adhesion}}{6 \pi z_{sep}} \right)$$  \text{(Eqn. 9.3)}

According to the JKR theory of adhesion mechanics, for a sphere on a flat surface of the same material with a work of adhesion $W_{adhesion}$ (joined surface energy; Section 9.2.3) and elastic modulus $K_e$:

$$r_{adhesion} = \left( \frac{6 \pi r^2 W_{adhesion}}{K_e} \right)^{1/3}$$  \text{(Eqn. 9.4)}

Assuming $W_{adhesion} = 10$ J/m$^2$ and $K_e = 10^{12}$ N/m$^2$ for smooth flawless diamond and $r = 0.5$ micron gives $r_{adhesion} = 36$ nm. Taking $z_{sep} = 0.4$ nm contact with a flat diamond plate over a ~4100 nm$^2$ circular adhesion spot in vacuo, then $F_{vdW} = 1300$ nN, a mean surface pressure of $p_{mean} = F_{vdW} / \pi r_{adhesion}^2 = 3200$ atm over the contact circle and a peak pressure $p_{peak} = 1.5 p_{mean} = 4800$ atm at the center. As the sphere is pulled away, separation occurs abruptly when the contact radius falls to $r_{adhesion} / 4^{1/3} = (-63\%)$ $r_{adhesion} = 23$ nm.

Van der Waals forces also depend upon the roughness of the surface. If a smooth sphere approaches a surface having uniform rugosity or roughness $b_{rug}$ with roughness feature size $< r$, then the van der Waals force is approximately:

$$F_{rug} = \left( \frac{z_{sep}}{z_{sep} + 2 b_{rug}} \right) F_{vdW}$$  \text{(Eqn. 9.5)}

For $z_{sep} = 0.4$ nm and $b_{rug} = 10$ nm, $F_{rug} = 2\% F_{vdW} = 4$ nN for the rigid sphere (in vacuo) in the examples given above. Surface dimpling is a common practice in MEMS fabrication, for example, to reduce adhesion between polysilicon layers and the substrate.

A few other geometries of rigid bodies approaching contact are useful as well. For example, the van der Waals adhesive force between two flat plates of equal interfacial area $A$ and uniform separation $z_{sep}$ is:

$$F_{vdW} = \frac{H A}{24 \pi z_{sep}^3}$$  \text{(Eqn. 9.6)}

Two (1 nm)$^2$ diamond plates separated by $z_{sep} = 0.4$ nm in vacuo require $F_{vdW} = 0.07$ nN (7 atm) to overcome the van der Waals attractive force. Such plates are the same size as the contacting end surfaces of the logic rods employed in Drexler’s mechanical computer design (Section 10.2.1); the design assumes a rod realignment force of 1 nN in the exemplar calculations. This result suggests that forces on the order of 10-100 pN must be applied to mobile nanoscale components inside mechanical nanodevices in order to overcome static van der Waals adhesion that arises due to the contact proximity of parts within the structure.

For another example, two rigid spheres of radii $r_1$ and $r_2$, separated by a distance $z_{sep} = r_1, r_2$, have a van der Waals adhesive force of:

$$F_{vdW} = \frac{H r_{sep}^{1/2}}{6 z_{sep}^{3/2}}$$  \text{(Eqn. 9.7)}

where the reduced radius $r_{red} = (r_1 r_2) / (r_1 + r_2)$. A pair of 1-micron diameter rigid diamond spheres separated by $z_{sep} = 0.4$ nm in vacuo have a van der Waals mutual adhesive force $F_{vdW} = 90$ nN.

As one final example, the van der Waals force between two parallel cylinders of length $L$, radii $r_1$ and $r_2$, and separation $z_{sep}$ is (differentiating $V_{vdW}$ in Drexler’s):

$$F_{vdW} = \frac{H L}{128 z_{sep}^{3/2}}$$  \text{(Eqn. 9.8)}

| Table 9.1. Hamaker Constant for Various Materials$^{10,1146,1149,1152,1162}$ |
|---------------------------------|-----------------|
| Material                        | Hamaker constant($\zeta$J) |
| Vacuum or Air                   | 0               |
| Cadherin-II coated surface      | 0.029           |
| RBC plasma membrane             | 0.05-5          |
| $\alpha$-crystallin eye lens protein | 0.26          |
| Liquid helium                   | 0.57            |
| Methanol                        | 36              |
| Water                           | 37              |
| Polytetrafluoroethylene         | 38              |
| Acetone                         | 41              |
| Ethanol                         | 42              |
| Hydrocarbons                    | ~ 50            |
| Hydrogen peroxide               | 54              |
| Glycol                          | 56              |
| Fused silica (SiO$_2$ quartz)   | 65              |
| Polystyrene                     | 66              |
| Glycerol                        | 67              |
| KBr crystals                    | 76              |
| Polyethylene                    | 76              |
| Polyvinyl chloride              | 78              |
| Platelet surfaces               | ~100            |
| Mica                            | 135             |
| Alumina (Al$_2$O$_3$, e.g., sapphire) | 140           |
| Germanium                       | 252-290         |
| Silicon                         | 260-275         |
| Graphite                        | 275             |
| Copper                          | 325             |
| Diamond                        | 340             |
| Silver                          | 344             |
| Metals (misc.)                  | 300-500         |
| Silicon carbide                 | 440             |


A pair of 1-nm diameter, 10-nm long rigid diamondoid cylinders lying exactly \( z_{sep} = 0.4 \) nm apart along their entire length in vacuo feel a van der Waals force of \( F_{\text{GW}} = 1.5 \) nN. In all three examples, as before, the force rises slightly if the objects deform and declines significantly if surface rugosity is increased. (See also entropic packing, Section 9.4.2.3.)

### 9.2.2 Electrostatic Adhesion Forces

Two types of electrostatic forces may act to hold particles to surfaces. The first type of force is due to bulk excess charges present on the particle or surface which produce a classical Coulombic attraction known as the electrostatic image force. Consider two spheres of radii \( r_1 \) and \( r_2 \) located a distance \( z_{sep} \) apart, where \( z_{sep} \ll r_1, r_2 \). Then following Eqn. 4.40 the image force is given by:

\[
F_{\text{image}} = \frac{q_1 q_2}{4 \pi \varepsilon_0 \varepsilon_r z_{sep}^2} = \frac{4 \pi \sigma_1 \sigma_2 r_{sep}^2}{\varepsilon_0 \varepsilon_r} \quad \text{(Eqn. 9.9)}
\]

where \( q_1 \) and \( q_2 \) are charge on the two spheres (coul), \( \sigma_1 \) and \( \sigma_2 \) are the surface charge densities (coul/m²), \( \varepsilon_0 = 8.85 \times 10^{-12} \) farad/m (permittivity constant), \( \varepsilon_r \) is the dielectric constant of the medium, and \( z_{sep} \approx r_1 + r_2 \) is the distance between charge centers. Taking \( \sigma_1 - \sigma_2 = 10^{-5} \) coul/m² (~63 charges/micron²) (cf. ~1000 charges/micron² for “highly charged surfaces”\(^{1145}\)), \( r_1 - r_2 = 0.5 \) micron and \( \varepsilon_r = 1 \) in vacuo, then \( F_{\text{image}} = 0.009 \) nN.

The electrostatic image force between two flat plates of equal interfacial area \( A \), equal charge density \( \sigma \), and uniform separation \( z_{sep} \ll A/z_r^2 \) is:

\[
F_{\text{image}} = \frac{\sigma^2 A}{2 \varepsilon_0 \varepsilon_r} \quad \text{(Eqn. 9.10)}
\]

For \( \sigma = 10^{-5} \) coul/m² and \( A = 1 \) micron² in vacuo (\( \varepsilon_r = 1 \)), \( F_{\text{image}} = 0.006 \) nN (~0.00006 atm). At larger \( z_{sep} \) and a voltage differential \( V_{\text{plate}} \) between the plates with capacitance \( C_{\text{plate}} \), then:

\[
F_{\text{image}} = \frac{C_{\text{plate}} V_{\text{plate}}^2}{2 z_{sep}} = \frac{\varepsilon_0 \varepsilon_r A V_{\text{plate}}^2}{2 \pi z_{sep}^2} \quad \text{(Eqn. 9.11)}
\]

Thus if \( A = 1 \) micron², \( V_{\text{plate}} = 1 \) volt, and \( z_{sep} = 0.2 \) micron, then \( F_{\text{image}} = 0.1 \) nN.

The second and more important electrostatic force for very small particles is the electrostatic contact potentials, also known as induced charge density \( \sigma \), and uniform separation \( z_{sep} \ll A/z_r^2 \) is:

\[
F_{\text{contact}} = \frac{\pi \varepsilon_0 \varepsilon_r V_{\text{contact}}^2}{z_{sep}} \quad \text{(Eqn. 9.12)}
\]

Taking \( r = 0.5 \) microns, \( V_{\text{contact}} = 0.5 \) volts, and \( z_{sep} = 5 \) nm, then \( F_{\text{contact}} = 1 \) nN (~0.01 atm). Induced charge densities vary experimentally from 2 x 10⁻⁷ coul/m² (~2 x 10⁻³ atm contact pressure) for polystyrene on polystyrene\(^{1153}\) up to 2 x 10⁻² coul/m² (~200 atm) for SiO₂ on mica.\(^{1154}\) In this latter study, a pair of contact-adhered silica/silica sheets required a force of ~0.08 nN/nm to pull them apart and a mica/mica pair required ~ 0.11 nN/nm; however, due to contact electrification a silica/mica pair required 6.6-8.8 nN/nm to separate.\(^{1154}\) Proper nanomechanical design can avoid or enhance these effects, as required.\(^{10}\)

Double layer electrostatic contact forces usually predominate over electrostatic image forces for small particles; image forces are important only for materials that can carry high surface charges, such as polymers of poor conductivity or other extreme insulators. Because of their \( t/z_{sep} \) dependency, contact electrical forces are usually more important than van der Waals forces only for large particles and at separations wider than ~100 nm, although electrical forces can dominate in smaller particles if the surface asperities of the particles are significant enough to remove the bulk of the particle from the contact point, thus greatly reducing the van der Waals interaction.\(^{1146}\)

In either case, the total forces of adhesion far outweigh the force of gravity, given by:

\[
F_{\text{gravity}} = \frac{4}{3} \pi r^3 \rho g \quad \text{(Eqn. 9.13)}
\]

For density \( \rho = 3510 \) kg/m³ for diamond, acceleration of gravity \( g = 9.81 \) m/sec², and radius \( r = 0.5 \) micron, \( F_{\text{gravity}} = 10^{-5} \) nN.

Entropic and electrostatic effects can sometimes interact to produce an apparent repulsion between opposite charges—as, for example, between positively charged bilayer surface and negatively charged colloidal particles.\(^{3674}\)

### 9.2.3 Immersive Adhesion Forces

For hydrophilic particles exposed to high humidity, or immersed and then withdrawn from a fluid, a liquid film can form on the surface of the particle by capillary condensation or by capillary action between particle and surface. For a spherical particle of radius \( r \) linked by a liquid bridge to a planar surface of the same material with contact area \( A \), separation \( z_{sep} \) and liquid surface tension \( \gamma \), then the capillary force between sphere and plane is \( F_{\text{capillary}} = 2 \gamma A / z_{sep} \) for small contact angles between the liquid and the surfaces.\(^{1146}\) Assuming hydrophilic (wettable) surfaces with wetting coefficient \( \cos(\theta) \) as defined in Section 9.2.4 and \( z_{sep} \ll r \), the capillary force is given approximately by:

\[
F_{\text{capillary}} = 4 \pi r \gamma \cos(\theta) \quad \text{(Eqn. 9.14)}
\]

which has been confirmed experimentally by direct measurement.\(^{1149}\)

For the water-air interface, \( \cos(\theta) \approx 1 \) and surface tension \( \gamma = 75.6 \) x \( 10^{-3} \) N/m at the freezing point (273 K), 72.75 x \( 10^{-3} \) N/m at room temperature (293 K), 70.05 x \( 10^{-3} \) N/m at human body temperature (~310 K), 58.90 x \( 10^{-3} \) N/m at the boiling point at 1 atm (373 K), and 110 x \( 10^{-3} \) N/m for ice at 273 K.\(^{763,1149}\)

Thus at room temperature, the capillary force on an \( r = 0.5 \) micron particle is \( F_{\text{capillary}} \approx 460 \) nN; an \( r = 10 \) nm particle feels \( F_{\text{capillary}} \approx 0.9 \) nN, a rather substantial force. Liquid bridges containing dissolved substances may evaporate and create solid crystalline bridges. Hydrophobic surfaces are generally unaffected by capillary forces.\(^{1146}\)

Eqn. 9.14 also applies to nanorobots that are caught in an air-liquid interface such as the surface of a puddle of water, alveolar fluids in the lung, or an epidermal pool of sweat. However, adding capillary-active solute to a solvent may lower surface tension
considerably. For dilute surface concentrations of a solution of molecular area $A_{mol}$ in solvent of surface tension $\gamma_0$ at temperature $T$, the resulting surface tension of the solution drops to:  

$$
\gamma = \gamma_0 - \frac{kT}{A_{mol}} \quad \text{(N/m)}
$$

[Eqn. 9.15]

For palmitic or stearic acid (e.g., soap), $A_{mol} = 0.21 \text{ nm}^2$, added to water ($\gamma_0 = 70.05 \times 10^{-3} \text{ N/m}$) at 310 K, surface tension falls to $\gamma = 49.70 \times 10^{-3} \text{ N/m}$. More concentrated solutions further reduce surface tension, nonlinearly with concentration. For example, $\gamma = 25 \times 10^{-3} \text{ N/m}$ for a 0.1 M aqueous fatty-acid solution of hexanoic acid (CH$_3$(CH$_2$)$_5$COOH; MW = 116 daltons), a total of -6 x 10$^7$ molecules, -0.01 micron$^2$, or -0.01 picograms of hexanoic acid per micron$^3$ of solution. For the aqueous solution-air interface, inorganic electrolytes (e.g., NaCl), organic acid salts, low-MW bases, and certain nonvalolatile nonelectrolytes such as glucose and glycerin are capillary inactive.

Surface tension is reduced even more dramatically by biological surfactants. Lung surfactant, whose activity is largely attributed to the phospholipid palmitoylphosphatidylcholine, produces a nonlinear change in surface tension with area. As the lungs fill with air, more-inflated alveoli have higher $\gamma$’s than less-inflated alveoli, serving to stabilize alveoli of different sizes. Aqueous solutions of surfactants from mammalian lung extracts show a surface tension reduction to as low as $\gamma = 0.7 \times 10^{-3} \text{ N/m}$ at 20% of maximum film distension. Survanta, an intratracheal pulmonary surfactant commonly used in neonates, can reduce surface tension from 340 zJ to 153 zJ. Blood cells are subject to similar forces. It has been proposed that the ~15 nm gap frequently observed between the surfaces of aggregated red cells represents the position of the potential energy minimum where the forces of electrostatic repulsion (between negatively charged red cells) and the van der Waals attractive forces are equal. Taking $H = 5 \text{ zJ}$, $r_{red} = 3 \text{ microns}$ and $z_{sep} = 15 \text{ nm}$ in Eqn. 9.7, the net attractive force between red cells is $F_{vdW} \approx 0.01 \text{nN}$.

Surface force experiments have demonstrated molecular control of adhesion at a fine level. According to the commonly-used JKR theory, the mechanical force needed to separate two elastic spheres each of radius $r$ already adhered by molecular contact is:

$$
F_{adhesion} = \frac{3\pi}{4} W_{adhesion} r
$$

[Eqn. 9.17]

where $W_{adhesion}$ is the Dupre reversible work of adhesion (energy per unit area). For example, hydrocarbon rubber spheres display $W_{adhesion} = 71 \times 10^{-3} \text{ J/m}^2$ in air but only $6.8 \times 10^{-3} \text{ J/m}^2$ in water. When sodium dodecyl sulfate or dodecylammonium chloride is added to the water so the liquid can wet the surface, $W_{adhesion}$ actually becomes negative—the surfaces are pushed apart by the intervening fluid. Nonwetting liquids increase adhesion. For instance, polydimethyl siloxane rubber has $W_{adhesion} = 43.6 \times 10^{-3} \text{ J/m}^2$ to itself, in air. In water, a nonwetting liquid, adhesion nearly doubles to $74 \times 10^{-3} \text{ J/m}^2$; in methanol, a wetting liquid, adhesion is almost eliminated, falling to $6 \times 10^{-3} \text{ J/m}^2$. In nanomedical engineering applications where minimization of adhesive forces is an important design objective, adhesion of specific classes of materials may be virtually eliminated by surface modifications that reduce $W_{adhesion}$ as is commonplace in biological surfaces and in engineered capillary coatings designed to reduce electroosmotic drag force during capillary zone electrophoresis.

9.2.4 Capillarity and Nanoscale Fluid Flow

In vivo nanodevices will often find it necessary to inject or to extract small aliquots of fluid from organelles, cells, tissues, or the environment. Internal gas or liquid transfers will also be commonplace in fluidic tethers, hydraulic communication and power conduits, nanohydraulic pistons, manipulators and metamorphic bumper drivers, and in bulk transfers into sensor cavities or chemical reaction chambers inside nanofactories (Chapter 19). Nanopipes may also be used in biomimetic nanosystems such as artificial organs and artificial cell components such as tubulin substitutes (Chapter 21). Hence it is essential to examine the nature of fluid flows and “wicking” in nanocapillary vessels, and the likely behavior of fluids in nanomachines.

The theory of macroscale capillarity is well-studied. The concept of “wetting” is basic. Consider a fluid in a vessel. At the solid-liquid-gas interface (the vessel wall) there is a characteristic contact angle $\theta$, the angle at which the meniscus contacts the container wall. This angle is indicative of the balance between the forces of (1) adhesion between the liquid and the solid wall and (2) cohesion within the surface of the liquid. A contact angle $<90^\circ$ means that adhesion is strong and the liquid is “wetting” the tube; an angle $>90^\circ$ implies that adhesion is relatively weak and the liquid is “nonwetting.” In air against glass, the liquids alcohol, glycerol, and pure water have $\theta = 0^\circ$, while turpentine has $\theta = 17^\circ$ and impure water has $\theta = 25^\circ$. All are wetting. On the other hand, mercury on glass has $\theta = 140^\circ$ and water on paraffin has $\theta = 109^\circ$, both are nonwetting.

In the classical example of capillarity, one end of a narrow-bore open-ended tube is dipped vertically into a liquid. The liquid wets
the tube, and the liquid rises until the force due to surface tension \( F_{\text{cap}} \) pulling the liquid upward equals the force of gravity \( F_{\text{grav}} \) pulling the column of liquid downward at some height \( h_{\text{cap}} \) given by:\textsuperscript{[1163]}

\[
 h_{\text{cap}} = \frac{2 \gamma_{\text{lg}} \cos(\theta)}{g r_{\text{cap}} (\rho_l - \rho_g)} \tag{Eqn. 9.18}
\]

with wetting coefficient \( \cos(\theta) = (\gamma_{\text{lg}} - \gamma_{\text{lg}}) / \gamma_{\text{lg}} \) (the Young equation\textsuperscript{[1163]}) where \( \gamma \) is surface tension at the solid-gas (sg), solid-liquid (sl), or liquid-gas (lg) interfaces; \( \rho_l \) and \( \rho_g \) are liquid and gas density; \( g = 9.81 \text{ m/sec}^2 \) (acceleration of gravity) and \( r_{\text{cap}} \) is the inside diameter of the tube. Thus pure water with air at STP in a glass capillary with \( r_{\text{cap}} = 1 \text{ micron} \) can rise \( h_{\text{cap}} = 15 \text{ meters} \) against gravity, representing a mean pressure \( p_{\text{cap}} = h_{\text{cap}} g (\rho_l - \rho_g) = 1 \text{ atm} \). A nonwetting liquid does not rise up the tube, but falls instead.

In the more general case of a linear capillary tube whose flow vector makes an angle \( \theta \) with the ambient gravity field (e.g., \( \theta_{\text{grav}} = \pi \) is a tube pointing straight up), then the force on the fluid column is:

\[
 F_{\text{column}} = F_{\text{cap}} + F_{\text{grav}} \tag{Eqn. 9.19}
\]

\[
 F_{\text{cap}} = 2 \pi \gamma_{\text{lg}} \cos(\theta) r_{\text{cap}} \tag{Eqn. 9.20}
\]

\[
 F_{\text{grav}} = \pi h_{\text{cap}} (\rho_l - \rho_g) g \cos(\theta_{\text{grav}}) r_{\text{cap}}^2 \tag{Eqn. 9.21}
\]

Taking the largest reasonable tube that might fit inside an in vivo nanorobot, \( h_{\text{cap}} = 1 \text{ micron} \), \( r_{\text{cap}} = 0.1 \text{ micron} \), \( \rho_l - \rho_g = 1000 \text{ kg/m}^3 \), \( \gamma_{\text{lg}} = 72 \times 10^{-3} \text{ N/m} \) for pure water, and taking \( \cos(\theta) = \cos(\theta_{\text{grav}}) = 1 \), then \( F_{\text{cap}} / F_{\text{grav}} = 10^8 \), giving the familiar result that gravity can usually be ignored in nanoscale systems. For this example, \( F_{\text{column}} = F_{\text{cap}} = 44 \text{ nN} \) and column fluid pressure \( p_{\text{column}} = 2 \gamma_{\text{lg}} \cos(\theta) / r \approx 14 \text{ atm} \). Taking instead the smallest reasonable nanocapillary tube with \( h_{\text{cap}} = 20 \text{ nm} \) and \( r = 2 \text{ nm} \), then \( F_{\text{cap}} / F_{\text{grav}} = 10^{-18} \), \( F_{\text{column}} = 0.9 \text{ nN} \) and \( p_{\text{column}} = 700 \text{ atm} \), making a quite substantial wicking force. Of course, the frictional effect of the interface region requires that a finite pressure \( P_0 = 2 \alpha / r_{\text{cap}} \) must be applied before fluid motion will begin,\textsuperscript{[1188]} where \( \alpha = 0.01-0.05 \text{ N/m} \) (experimental); taking \( \alpha = 0.01 \text{ N/m} \), then \( P_0 = 2 \text{ atm} \) for \( r = 0.1 \text{ micron} \), or \( 100 \text{ atm} \) for \( r = 2 \text{ nm} \).

The capillary force may also be compared to the simple hydrostatic suction force:

\[
 F_{\text{suck}} = \pi r_{\text{cap}}^2 \Delta p \tag{Eqn. 9.22}
\]

A nanopipette with inside radius \( r_{\text{cap}} = 0.1 \text{ micron} \) which is pressed against a plasma membrane surface and is then evacuated behind the attachment point sufficiently to create a surfacial pressure differential of \( \Delta p = 0.3 \text{ atm} \) produces a purely hydrostatic suction force of \( F_{\text{suck}} = 1 \text{ nN} \).

Such classical continuum models assume, among other things, that the molecular graininess of the fluid can be ignored. This assumption fails when tube dimensions (e.g., \( r_{\text{cap}} \)) are comparable to the characteristic molecular length scale \( \lambda \) of the fluid.\textsuperscript{[10]} In a gas, \( \lambda_{\text{gas}} \) is the mean free path between collisions. For \( T_{\text{gas}} = 310 \text{ K} \) under ideal gas conditions, if we take the effective molecular diameter \( d_{\text{gas}} = 0.2 \text{ nm} \). At high pressure, van der Waals equation must be used (Section 10.3.2), but for \( p_{\text{gas}} = 1000 \text{ atm} \) with \( n_{\text{gas}} = 1.3 \times 10^{28} \text{ molecules/m}^3 \), then \( \lambda_{\text{gas}} = 0.4 \text{ nm} \). In a liquid, \( \lambda_{\text{liq}} = d_{\text{liq}} \) the molecular diameter; for water molecules, \( \lambda_{\text{liq}} = 0.3 \text{ nm} \).

A ratio of \( r_{\text{cap}} / \lambda_{\text{gas}} \approx 1 \) marks an important transition because it implies that molecules are colliding with the walls about as often as they are colliding with each other. If \( r_{\text{cap}} \ll \lambda_{\text{gas}} \), then intermolecular interaction is rare, wall collisions are relatively frequent, and molecular motion is largely ballistic between wall impacts. But if \( r_{\text{cap}} >> \lambda_{\text{gas}} \), then molecule/wall interactions are relatively rare and intermolecular collisions are relatively frequent. In the latter situation the continuum flow relations should become valid, for example when \( r_{\text{cap}} >> 200 \text{ nm} \) for gases at \( -1 \text{ atm} \) or when \( r_{\text{cap}} >> 0.4 \text{ nm} \) for gases at \( -1000 \text{ atm} \), or when \( r_{\text{cap}} >> 0.3 \text{ nm} \) for liquids.

The range of non-continuum skin layer effects appears quite narrow in many cases. For instance, two smooth mica surfaces that are slowly brought together while immersed in various nonaqueous fluids display oscillating attractive and repulsive forces as a function of separation when the gap is less than 6-10 molecular diameters; periodic jump distances are about equal to the molecular diameter of the fluid.\textsuperscript{[1165]} Oscillations become smaller between rough surfaces and in liquids that mix molecules of different sizes.\textsuperscript{[1143]} Similar experiments in water with electrolyte produce similar behavior, but with ionic double layer repulsions superimposed—oscillatory solvent forces become significant only at \( z_{\text{sep}} < 2 \text{ nm} \).\textsuperscript{[1166]} Elastohydrodynamic simulations suggest that flat gold surfaces separated by 2.3 nm with 0.9-nm asperities (the “near-overlap” case) would slide smoothly on a thin film of liquid lubricant molecules, albeit with nanometer-scale cavitated zones extending \( -3 \text{ nm} \) downstream persisting for \( -0.1 \text{ nanosec} \) at a sliding velocity of 10 m/sec.\textsuperscript{[1172]} Solvation forces may also appear at separations below a few molecular diameters.\textsuperscript{[1149]}

These results suggest that water-carrying nanocapillaries \( >2-4 \text{ nm} \) in internal diameter should function largely in line with continuum models. Early theoretical calculations predicted that open-ended carbon nanotubes as small as \( 0.8 \text{ nm} \) in diameter might act as “nanostraws” and couldwick in molecules from vapor or fluid phases.\textsuperscript{[1167]} Subsequent experiments\textsuperscript{[1168-1170]} confirmed that fluids with \( \gamma_{\text{lg}} \approx 190 \times 10^{-3} \text{ N/m} \), including liquid sulfur, liquid selenium and nitric acid, are readily drawn into the inner cavity of carbon nanotubes with inside diameters of \( 4-8 \text{ nm} \) through capillarity.\textsuperscript{[1169]} This limit is sufficiently high to allow nanotube wetting by water (\( \gamma_{\text{lg}} = 72 \times 10^{-3} \text{ N/m} \)), most organic solvents (\( \gamma_{\text{lg}} < 72 \times 10^{-3} \text{ N/m} \)), and liquid acids (e.g., \( \gamma_{\text{lg}} = 43 \times 10^{-3} \text{ N/m} \) for HNO\textsubscript{3}) which can then be used as low surface tension carriers to introduce dissolved solute into the nanotubes.\textsuperscript{[1170]} Bulk nanotubes are readily wetted by water.\textsuperscript{[1168]} Good wetting is favored when the polarizability of the tube material is higher than that of the liquid.\textsuperscript{[1171]} Pure metals like molten lead or mercury with \( \gamma_{\text{lg}} \approx 190 \times 10^{-3} \text{ N/m} \) do not wet carbon nanotubes and are not drawn in by capillarity. Such high-\( \gamma_{\text{lg}} \) liquids may be forced into the nanotubes by applying a hydrostatic pressure \( p_{\text{force}} \) given by the Laplace equation:\textsuperscript{[1168]}

\[
 \lambda_{\text{gas}} = (2^{1/2} \pi n_{\text{gas}} d_{\text{gas}}^2)^{-1} = (2^{1/2} \pi p_{\text{gas}} d_{\text{gas}}^2 / kT_{\text{gas}})^{-1}
\]

\( - 200 \text{ nm} \)
Thus, liquid mercury with γlg = 490 x 10^{-3} N/m can be pushed into an r_cap = 0.5 micron tube by applying P_force = 20 atm to the fluid; P_force = 2000 atm for r_cap = 5 nm.

By 1998, detailed molecular dynamics simulation of fluid flow inside 1.3-1.6 nm diameter carbon nanotubes had been a subject of active research.\textsuperscript{1173,1174} One major difference between macroscale and nanoscale fluidic systems is that in nanomachines the tube walls may be considerably less rigid\textsuperscript{*} and may flex and resonate. Nanomachine designers must take into account the effects of these vibrations as they are transmitted through nanoscale structural elements attached to the tubes. Furthermore, tube wall motions are sometimes strongly size-dependent. The simulations confirmed that fluids flow faster through rigid tubes than through floppy tubes, as predicted by classical hydrodynamics theory (Section 9.2.5). At high flow velocities we may expect nanotubes to exhibit buckling modes, variable patency, progressive wave propagation, slugging, flow-limiting flutter, and stall, by analogy to the fluid mechanics of compliant tube flow in human veins.\textsuperscript{361} Carbon nanotubes may be rigifidified by imposing a stretching tension, or possibly may be replaced by stiffer diamondoid tubes.\textsuperscript{1173} If the fluid has more than one kind of atom, size-mass effects and differing interaction ranges or strengths may cause one or more species to segregate at the walls or to flow at different rates.\textsuperscript{1174}

Material flow through nanotubes has been well-exploited by biology, including the bacterial sex pili\textsuperscript{3540} (Section 5.4.2), bacterial anal pores and cytoprocts,\textsuperscript{3550} cellular excretory canals (e.g., in the glandular cells of the pancreas), 10-100 nm intranuclear canals (Section 8.5.4.5), the 20-40 nm wide hollow tubular fluid transport network that forms the smooth intracellular endoplasmic reticulum (Section 8.5.3.5), the 75-nm diameter nutrient-transporting tubovesicular membrane network of the human malaria parasite,\textsuperscript{1183} the >100-nm wide fluid-carrying human bone canaliculi (Section 8.2.4), and prelymphatic tissue channels (Section 8.2.1.3).

The syringelike T4 bacteriophage tail assembly is perhaps the best-studied example of biological nanotube flow.\textsuperscript{1179,1180} The 100-nm long, 20-nm wide cylindrical T4 tail assembly, made of 15 different proteins joined in 24 annular segments with an 8-nm inside bore, has a set of small fibers near the tip that attach to the plasma membrane of the host cell. After a lysozyme-like enzyme opens a breach in the host cell wall, the tail sheath thickens and contracts,\textsuperscript{1181} inserting a -2.8 megadalton hollow core protein nanotube (80 nm long, 7 nm wide, 2.5 nm inside bore) through the host cell integument. The protein nanotube is then uncorked in the host cell, 2-nm wide DNA thread (~50% of head volume) through the 2.5-nm inside bore of the entire tube (t_flow) then follow directly from Eqn. 9.25 as:

\[ \eta \approx 72 \times 10^{-3} \text{ N/m} \text{ for water, } K_{\text{liquid}} \approx 300 \text{ atm.} \]

Taking η = 0.6915 x 10^{-3} kg/m-sec for pure water at T = 310 K, a nanotube with r Tube = 10 nm, l tube = 1 micron, and Δp = 1 atm passes fluid at V_{flow} = 0.6 micron^3/sec, v_{flow} = 2 mm/sec, with t_{flow} = 0.5 milliseconds. The incompressibility assumption generally holds in liquids and in gas flows where Δp << P_{tube} where P_{tube} is the head pressure at the tube entrance.

The above equations describe the resistance to Poiseuille (laminar) flow in a pipe, which is the minimum of resistance to all possible flows in a pipe.\textsuperscript{361} If the flow becomes turbulent, the resistance increases. The determinative parameter is a dimensionless quantity called the Reynolds number,\textsuperscript{1187} N_R, which is the ratio of the inertial pressure (\rho v_{flow}^2) to the viscous pressure (\eta v_{flow} / r_{tube}) in the flow of a fluid of density ρ, or:

\[ N_R = \frac{\rho v_{flow} r_{tube}}{\eta} \]  

For the example of the 10-nm nanotube in the previous paragraph, N_R = 10^{-3}. A large Reynolds number (Section 9.4.2.1) implies a large internal tensile strength K_{liquid} \approx 4 \gamma_{lg} / x_{molec},\textsuperscript{1163} where x_{molec} ~ 10 nm is the approximate maximum range of intermolecular forces. Taking γ_{lg} \approx 72 \times 10^{-3} N/m for water, K_{liquid} \approx 300 atm. There is experimental evidence that water saturated with air but denucleated by high pressures does exhibit a tensile strength on the order of ~300 atm.\textsuperscript{1175} This tensile strength, and not capillarity, is credited for the ability of tall trees to pull sap up to 115 meters high,\textsuperscript{1176} far exceeding the maximum 10.33 meters that water can be pulled vertically at the Earth’s surface using a vacuum pump. M. Zimmermann\textsuperscript{2033} reviews the sap transport system; the maximum pull recorded experimentally in trees is 120 atm.\textsuperscript{2032}
turbulent flow typically occurred at $N_R > 2000$, depending upon the smoothness of the entry conditions. The lowest value obtainable experimentally on a rough entrance appeared to be $N_R \approx 2000$. However, when extreme care was taken to establish smooth entry conditions (as is more likely in precisely structured nanotubes constructed using molecular manufacturing techniques) the transition could be delayed to Reynolds numbers as high as 40,000.

Even assuming the more conservative $N_R \approx 2000$ figure, it is clear that subsonic flow in nanoscale pipes will almost always be laminar. For a pipe with $r_{tube} = 1$ micron conveying water at 310 K ($\rho = 993.4$ kg/m$^3$), the onset of turbulent flow (taking $N_R > 2000$) occurs at $v_{turb} (= v_{flow}$ in Eqn. 9.29) $> 1400$ m/sec, very nearly the speed of sound in water. In gases or liquids of lower density, or in pipes of narrower bore, or if a less conservative transitional Reynolds number is available due to superior design, then $v_{turb}$ grows still larger. Thus turbulence is primarily a high-velocity, large-tube phenomenon.

In such turbulent flow situations with $N_R \approx 1000$, a well-known empirical formula for the volume flow rate is:

$$V_{turb} - \frac{V_{HP}}{Z} \{\text{m}^3/\text{sec}\} \quad \text{[Eqn. 9.30]}$$

where the turbulence factor $Z = 0.005 \ N_R^{0.64}$. Thus for $N_R > 3000$, $V_{turb} = 0.5 \ V_{HP}$. Nevertheless, even in turbulent conditions of the general flow, fluid motion nearest the tube wall remains laminar in a thin layer of thickness $x_{lam}$, often estimated as:

$$x_{lam} = \frac{50 \ \eta \ r_{tube}}{\rho \ \Delta p} \{\text{m}^3/\text{sec}\} \quad \text{[Eqn. 9.31]}$$

For flowing water at 310 K and $r_{tube} = 10$ micron, $l_{tube} = 100$ micron, and $\Delta p = 10$ atm giving $v_{flow} \approx 200$ m/sec and $N_R > 3000$, then $x_{lam} \approx 0.5$ microns.

Poiseuille's Law assumes a rigid pipe, an assumption that may not hold for human blood vessels, especially the veins. A complete treatment of fluid flow in elastic tubes is beyond the scope of this book. However, for steady laminar flow in an elastic tube with wall thickness $h_{wall}$, radius $r_{tube}$, length $l_{tube}$, $p_0$ and $p_1$ the pressures at the entry and exit ends of the tube, and $E_{wall} = Young's$ modulus of a Hookean (e.g., a linear force-distance relationship) wall material, then the volume flow rate through the tube may be approximated by:

$$V_{elastic} = \frac{\pi \ c_h r_{tube}^4}{24 \ \eta \ l_{tube}} \left\{1 - \frac{p_1}{c_h} \right\}^3 - \left\{1 - \frac{p_0}{c_h} \right\}^3 \{\text{m}^3/\text{sec}\} \quad \text{[Eqn. 9.32]}$$

where $c_h = E \ h_{wall} / r_{tube}$. Taking $p_0 = 1$ atm, $E = 100$ N/m$^2$ (typical for human skin; Table 9.3), then $V_{elastic} \approx 0.3$ micron$^3$/sec. Care must be taken in applying this formula because many biological tube materials such as blood vessel walls are non-Hookean and exhibit a linear pressure-radius relationship instead. For small elastic deformations, the volume flow rate of such non-Hookean tubes may be approximated by:

$$V_{limit} = \frac{\pi \ r_{tube}^4}{20 \ \eta \ l_{tube}} \bigg\{\frac{3}{4} \bigg(\frac{c_{solvated} - c_{ambient}}{3} \bigg) \bigg\}^{1/2} \{\text{molecules/sec}\} \quad \text{[Eqn. 9.33]}$$

where $\alpha$ is the compliance constant, measured experimentally in feline pulmonary veins as 1.98-2.79 m$^3$/N for $r_{tube} = 50-100$ microns down to 0.57-0.79 m$^3$/N for $r_{tube} = 400-600$ microns. Other geometric nonuniformities such as regular and irregular tube tapers, embedded resonance chambers, tube bifurcations or multitube confluences have important effects on flow rate but are beyond the scope of this book. Poiseuille’s Law also does not strictly hold for fluids seeded with polymers in concentrations as low as $10^{-2}-10^{-3}$ by weight, wherein drag may be reduced by a factor of 2-3.

It bears repeating here that as pipes get very small, they can clog more easily due to van der Waals adhesive forces, so it becomes increasingly important to engineer interior vessel surfaces for minimum adhesivity with respect to all materials likely to be transported through the pipes (Section 9.2.3). Bursting strength of fluid-filled pressure vessels is addressed in Section 10.3.1.

9.2.6 Effervescence and Crystallization

Diffusion-limited molecular inflow rates to nanodevices have already been discussed in connection with molecular transport (Section 3.2.2), chemical sensors (Section 4.2), chemical power sources (Sections 6.3.4.1 and 6.5.3), and chemical broadcast communication (Section 7.2.1). A related constraint applies when nanodevices attempt to offload endogenously produced or stored molecules into the local environment, whether by nozzled flow, reversible sorting rotors, or by simple bulk venting. If the excretion rate exceeds the solvation capacity of the surrounding solvent, offloaded gases may coalesce into bubbles and effervesce; offloaded solids may crystallize.

The inception of bubbles or crystals is a complex physical process whose detailed description is beyond the scope of this book. However, the diffusion-limited maximum offloading rate may be conservatively estimated by considering a point emission source releasing soluble gas molecules continuously into an aqueous environment at the constant rate $Q$ (molecules/sec). Once emitted, molecules with diffusion coefficient $D$ diffuse a distance $\Delta X \sim (2 \ D \ Q)^{1/2}$ (Eqs. 3.1 and 3.5), giving a maximum steady state concentration $c_{diffuse} \approx (3 / 4 \ \pi) (Q / 2 \ D)^{3/2}$ (molecules/m$^3$). As an approximation at pressures $\leq 100$ atm for gases that do not chemically unite with the solvent, Henry’s law gives the concentration of dissolved gas as $c_{solvated} = K_{Henry} p_{gas}$ (molecules/m$^3$), where $p_{gas}$ is the partial pressure of the solvated gas and $K_{Henry}$ is the Henry’s law constant for the gas (Table 9.2). If the ambient concentration of the gas in the solvent is $c_{ambient}$, then the maximum offloading rate $Q_{limit}$ necessary to enable the formation of bubbles occurs when $c_{diffuse} \geq (c_{solvated} - c_{ambient})$, or:

$$Q_{limit} \geq (4 \pi / 3) (c_{solvated} - c_{ambient})^{2/3} \{\text{molecules/sec}\} \quad \text{[Eqn. 9.34]}$$

Thus for oxygen dissolving in 0.15M saline (-human blood plasma), $D = 2.0 \times 10^{-9}$ m$^2$/sec (Table 3.3), $K_{Henry} = 7.4 \times 10^{23}$ molecules/m$^3$-atm (Table 9.2), and $c_{ambient} = 3 \times 10^{22}$ molecules/m$^3$ (venous plasma; Appendix B), we have $Q_{limit} \geq 8 \times 10^3$ molecules/sec at $p_{gas} = 1$ atm, or $Q_{limit} \geq 2 \times 10^3$ molecules/sec at $p_{gas} = 100$ atm. Faster offloading rates than $Q_{limit}$ may cause effervescence.
Table 9.2. Henry’s Law Constants (gas solubility in water at 298 K)\(^{390,326,2050}\)

<table>
<thead>
<tr>
<th>Solvated Gas</th>
<th>Henry’s Law Constant (molecules/m(^3)-atm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>He</td>
<td>~8 x 10(^{-22})</td>
</tr>
<tr>
<td>(N_2) (0.15M saline)</td>
<td>3.5 x 10(^{21})</td>
</tr>
<tr>
<td>(N_2) (pure water)</td>
<td>3.8 x 10(^{21})</td>
</tr>
<tr>
<td>(H_2)</td>
<td>4.7 x 10(^{21})</td>
</tr>
<tr>
<td>CO</td>
<td>5.8 x 10(^{21})</td>
</tr>
<tr>
<td>(O_2) (0.15M saline)</td>
<td>7.4 x 10(^{21})</td>
</tr>
<tr>
<td>(O_2) (pure water)</td>
<td>7.6 x 10(^{21})</td>
</tr>
<tr>
<td>(CH_4)</td>
<td>8.0 x 10(^{21})</td>
</tr>
<tr>
<td>(CH_4)</td>
<td>1.1 x 10(^{24})</td>
</tr>
<tr>
<td>(CO_2) (pure water)</td>
<td>2.9 x 10(^{24})</td>
</tr>
<tr>
<td>(CO_2) (0.15M saline)</td>
<td>1.9 x 10(^{25})</td>
</tr>
<tr>
<td>(C_2H_2)</td>
<td>2.0 x 10(^{25})</td>
</tr>
<tr>
<td>(C_2H_4)</td>
<td>2.5 x 10(^{25})</td>
</tr>
<tr>
<td>(NH_3)</td>
<td>1.5 x 10(^{28})</td>
</tr>
</tbody>
</table>

For carbon dioxide in saline (ambient arterial plasma concentration), \(Q_{\text{limit}} \approx 7 \times 10^9\) molecules/sec (1 atm) or 1.5 \times 10\(^{10}\) molecules/sec (100 atm). Once a gas bubble is formed and completely blocks a narrow pipe, the pressure required to dislodge it is approximately \(P_{\text{reservoir}}\) as given by Eqn. 9.24, ~3 atm for an air bubble caught in a 1 micron wide tube carrying water at 310 K.

Similar considerations apply to solute releases which, if too rapid, may result in crystallization of the solid. In this case, Eqn. 9.34 may be used if \(c_{\text{solvated}}\) is taken as the saturation concentration at the appropriate temperature and pressure. For example, a saturated 70% glucose solution has \(c_{\text{solvated}} = 7.8 \times 10^{27}\) molecules/m\(^3\); taking \(D = 7.1 \times 10^{-10}\) m\(^2\)/sec and \(c_{\text{ambient}} = 2.3 \times 10^{24}\) molecules/m\(^3\) (~blood plasma; Appendix B), \(Q_{\text{limit}} \approx 1.4 \times 10^{10}\) molecules/sec.

Gas solubility almost always declines at warmer temperatures because gas-liquid solvation is exothermic, whereas the solubility of most solids (including glucose and most important electrolytes) usually rises at higher temperatures. (See also Section 10.5.3.)

9.2.7 Fluid Pumping and Plumbing

After reviewing specific requirements in nanomechanical systems for containing fluids within walls and seals, Drexler\(^{10}\) concluded that nonbonded nanoscale interfaces could serve as seals to permit the relative motion of surfaces while hindering the flow of fluid between them. Such fluid-tight seals enable the construction of valves, pistons and cylinders, and hence fluid pumps. The following is a brief discussion of several useful classes of general-purpose nanopump designs. Molecule-specific nanopumps are described in Section 3.4; biocompatibility issues\(^{3234}\) are deferred to Chapter 15.

9.2.7.1 Pressure Release Pumps

Some form of motive force must be applied to cause fluid to flow through a tube. The simplest possible motive force is pressurization. Consider a valved cylindrical nozzle of radius \(r_{\text{tube}}\) and length \(l_{\text{tube}}\) attached to a reservoir of volume \(V_{\text{reservoir}}\) maintained at a constant pressurization \(P_{\text{reservoir}}\) (even while being emptied). When the valve is opened, the fluid vents through the nozzle at a maximum velocity \(v_{\text{max}}\) into an external fluid environment having static pressure \(P_{\text{external}}\), emptying the reservoir in a time \(t_{\text{empty}}\), both given by:

\[
\begin{align*}
  t_{\text{empty}} &= \frac{8 \pi l_{\text{tube}} V_{\text{reservoir}} \eta}{\nu r_{\text{tube}}^2 (P_{\text{reservoir}} - P_{\text{external}})} \quad \text{(Eqn. 9.35)} \\
  P_{\text{reservoir}} &\approx P_{\text{external}} + \frac{8 \pi l_{\text{tube}} V_{\text{max}}}{\nu r_{\text{tube}}^2} \quad \text{(Eqn. 9.36)}
\end{align*}
\]

Taking \(P_{\text{external}} \approx 1\) atm, \(\eta = 0.6915 \times 10^{-3}\) kg/m-sec for water at 310 K, \(r_{\text{tube}} = 10\) nm and \(l_{\text{tube}} = 100\) nm, then holding \(v_{\text{max}} \leq 10\) micron/sec (the velocity of random thermal intracellular hydrodynamic flows; Section 8.5.3.12) for intracellular discharges requires a driving pressure of \(P_{\text{reservoir}} \approx 0.5 \times 10^{-3}\) atm, which takes \(t_{\text{empty}} \approx 0.3\) sec to empty the reservoir at an energy cost (Eqn. 9.27) of \(P_{\text{flow}} = \frac{10}{\text{empty}} \approx 55\) J (~13 kT). As a practical matter, normal variations in physiological pressure due to ~1 Hz heartbeat pulse waves are 0.5-10 \times 10^{-3} atm in the tissues and veins and up to 50-70 \times 10^{-3} atm in the aorta (Section 4.9.1.2), so either valve aperture or \(P_{\text{reservoir}}\) must be dynamically controlled if a steady exit velocity is required at such low discharge velocities. In addition, a minimum pressurization of \(-P_{\text{force}}\) (Eqn. 9.24) must be used to overcome capillary forces and initiate the flow. Device cooling during rapid gas discharge is described in Section 5.3.3.

9.2.7.2 Positive Displacement Pumps

Mechanical energy may be converted to fluid flow using a side-mounted piston as a single-action reciprocating positive displacement pump. Consider a pipe of radius \(r_{\text{tube}}\) and length \(l_{\text{tube}}\), with valves at either end and a piston mounted in a stub cylinder of radius \(r_{\text{stub}}\) mounted perpendicular to the pipe axis at the midpoint of the pipe length (Fig. 9.1). The valves may be purely passive, or may be actively controlled in synchronization with piston motions.

During a single pump cycle, the valve at one end of the pipe is opened and the piston in the stub cylinder is drawn out, sucking a fluid volume \(V_{\text{cycle}}\) from the external environment into the pipe through the open end. The first valve is then closed, the second valve at the other end is opened, and the piston is pushed back into the stub cylinder, forcing an equal volume \(V_{\text{cycle}}\) of fluid out through the opposite end of the pipe which is now open. Cycling this pump at a frequency \(f_{\text{pump}}\) establishes a semicontinuous flow of \(V_{\text{pump}} = V_{\text{cycle}}\), with pumped-fluid velocity \(V_{\text{flow}}\) and pump power \(P_{\text{flow}}\) given by:

\[
\begin{align*}
  v_{\text{flow}} &= \frac{V_{\text{pump}}}{\frac{\pi l_{\text{stub}}^2}{4}} \quad \text{(m/sec)} \quad \text{(Eqn. 9.37)} \\
  P_{\text{flow}} &= \frac{P_{\text{dis}}}{\frac{8 \pi}{\nu r_{\text{tube}}^2} v_{\text{flow}}^2} \quad \text{(watts)} \quad \text{(Eqn. 9.38)}
\end{align*}
\]

where \(P_{\text{dis}}\) is the power dissipation due to piston wall friction, inertia, and fluid drag (Eqn. 6.13); for submicron pistons and \(V_{\text{pump}} \leq 1\) MHz, \(P_{\text{dis}} / V_{\text{pump}} \ll k\)T in good designs. From Eqn. 9.25, pumping pressure is given by \(P_{\text{external}} = 8 \pi l_{\text{tube}} V_{\text{flow}} / r_{\text{tube}}^2\).

Operating frequency is restricted by at least four factors:
1. The mechanical response time of the fluid defines a maximum operating frequency of \( v_{\text{max}} \leq \frac{v_{\text{sound}}}{(l_{\text{tube}}/2)} \); for \( r_{\text{tube}} = 100-1000 \) nm and \( v_{\text{sound}} \sim 1500 \) m/sec in water at 310 K, \( v_{\text{max}} = 30-3 \) GHz. (Note also the requirements that \( v_{\text{max}} << v_{\text{sound}} \) to avoid acoustic radiation losses, and that pressure and frequency must be sufficiently limited to avoid cavitation (Section 6.4.1).)

2. The maximum mechanical operating speed of nanoscale pistons is, conservatively, \( v_{\text{max}} \sim \) GHz.\(^{10}\)

3. The maximum acceptable flow velocity \( v_{\text{max}} \) defines a maximum operating frequency \( v_{\text{max}} = \pi r_{\text{tube}}^2 \frac{v_{\text{max}}}{V_{\text{cycle}}} = \frac{v_{\text{max}}}{4 r_{\text{tube}}} \) taking \( V_{\text{cycle}} = (2 r_{\text{tube}})^3 \) (e.g., one full piston throw opens a volume equal to one cubic pipe diameter); taking \( v_{\text{max}} = 4 \) m/sec, then \( v_{\text{max}} = 1-100 \) MHz for \( r_{\text{tube}} = r_{\text{stub}} = 1000-10 \) nm.

4. The maximum valving speed of \( v_{\text{valve}} = 0.01-1 \) m/sec (Sections 3.2.4 and 3.3.2) when applied to valves of radius \( r_{\text{tube}} \) defines a maximum valving frequency of \( v_{\text{max}} = v_{\text{valve}} / r_{\text{tube}} \); for \( r_{\text{tube}} = 1000-10 \) nm, \( v_{\text{max}} = 0.01-100 \) MHz.

These considerations suggest that as a conservative limit, \( v_{\text{max}} \sim 1 \) MHz may be appropriate in nanomedical piston pump designs that are intended to interact with an aqueous or protoplasmic external biological environment. Nanorobot piston pumps transferring durable fluids internally may safely operate at frequencies up to ~GHz and pressures >1000 atm.

From Eqsns. 9.37 and 9.38, and taking \( r_{\text{tube}} = 10 \) nm, \( l_{\text{tube}} = 100 \) nm, \( v_{\text{pump}} = 1 \) MHz, \( \eta = 0.6915 \times 10^3 \) kg/m-sec for 310 K water, and \( V_{\text{cycle}} = (2 r_{\text{tube}})^3 = 8000 \) nm\(^3\), then \( v_{\text{flow}} = 3 \) cm/sec, \( \Delta p = 1.6 \) atm, and \( P_{\text{flow}} = 1 \) pW assuming \( P_{\text{diss}} << P_{\text{flow}} \), which will normally be the case. The mean pressure amplitude developed during each piston cycle is \( \Delta p_{\text{piston}} = 8 \eta l_{\text{tube}} v_{\text{flow}} / r_{\text{tube}}^2 = 1 \) atm. Reynolds number \( R_{\text{R}} = 0.0004 \) (Eqn. 9.29), giving laminar flow throughout the cycle; power losses due to differential compression of the streamlines as fluids curve through an elbow to enter and exit the piston stub should be negligible. If precise control of flow volume is not necessary, the active valves described above may be replaced by passive one-way check valves or flap valves analogous to those found in human veins (Fig. 8.3).

Many other mechanical positive displacement pumps are readily imagined—including diaphragm pumps, rotary pumps (valveless), hydraulic vane pumps (valveless), lobe pumps (valveless), screw pumps (valveless), reciprocating flap-valve or plunger pumps, bellows pumps, multiple piston pumps, continuous ciliary or peristaltic pumps (valveless), and paddlewheel or waterwheel “gear” pumps (valveless)—all with similar scaling and performance parameters. The dominance of viscous forces over inertial forces in fluid flow at small scales favors the use of these positive displacement or static-type pumps over the use of rotdynamic or kinetic-type pumps such as centrifugal pumps (with impeller blades) and venturi or jet pumps (e.g., using a motive fluid, nozzle and diffuser).

9.2.7.3 Turbomolecular Gas Pumps

Drexler\(^{10}\) has briefly reviewed the general design parameters of nanoscale turbomolecular gas pumps. Pump effectiveness depends on the ratio of the blade speed to the characteristic thermal speed of the lightest gas molecule to be pumped. If the blades are constructed of diamondoid materials, then blades may be as thin as ~1 nm and blade speed can exceed the velocities of the fastest gas molecules (e.g., 1960 m/sec for hydrogen at 310 K; Eqn. 3.3), providing a compression ratio of \( \approx 10 \) per blade row.\(^{1204}\) Turbomolecular pumps are designed to operate under free-flow conditions involving essentially ballistic gas molecule trajectories of length \( l_{\text{gas}} = 10 \) nm at 20 atm pressure, or \( l_{\text{gas}} = 200 \) nm at 1 atm (Eqn. 9.23). Taking pump length per blade row ~10 nm for operating pressures up to ~20 atm, a pump assembly consisting of a stack of five blade disks achieves a compression ratio of ~10\(^3\) in a ~50 nm pump length.

9.2.7.4 Nonmechanical Pumps

Nonmechanical pumps may induce fluid flow without using any moving parts in direct contact with the fluid. Electrohydrodynamic (EHD) pumps are one important class of these devices. Several different types of EHD micropumps have been fabricated.

For example, in the DC-charge injection pump\(^{1205,1207}\) or ion drag pump,\(^{1207-1209}\) ions are injected into the liquid at an emitting electrode under high field conditions and move under the influence of Coulombic forces toward a second collecting electrode where they are absorbed. Collisions between ions and fluid molecules transfer momentum from the ions to the fluid, producing fluid movement in the direction of ion flow. A prototype micropump has achieved a maximum pumping pressure of 0.01 atm (static) for ethanol flowing between two (2500 micron\(^2\)) grids spaced 10 microns apart at 300 volts, producing a volumetric flow rate of \( V = 2.7 \times 10^{-11} \) m\(^3\)/sec at a flow velocity of ~4 microns/sec.\(^{1206}\) Static pump pressure \( P_{\text{static}} \) may be estimated\(^{1208,1209}\) as:

\[
P_{\text{static}} = c_{\text{geom}} \varepsilon_0 K_e \left( \frac{V}{d} \right)^2
\]

where \( c_{\text{geom}} \) is a correction factor for electrode geometry of order unity, \( \varepsilon_0 = 8.85 \times 10^{-12} \) farad/m (permittivity constant), \( K_e = \) dielectric constant, \( V \) is electric potential and \( d \) is electrode separation. Taking \( V/d = 10^7 \) volt/m and \( K_e = 74.31 \) for deionized water at 310 K, \( P_{\text{static}} = 0.7 \) atm. Pumping of different polar fluids with conductivities between 10\(^{-6}\)–10\(^{-12}\) (ohm-m)\(^{-1}\) such as propanol, acetone, deionized water, and many organic solvents including several oils has been demonstrated, but aqueous solutions of electrolytes could not be pumped due to their high ionic conductivity.
In the traveling wave electroconvection voltage pump,\textsuperscript{1210-1212} phase-shifted rectangular voltages are applied across a series of parallel electrodes. Temperature induced conductivity gradients induce free electric charges which can interact with the traveling field due to charge relaxation processes in the volume of the fluid, causing fluid flow in a direction transverse to the electrodes. A small monotonotic temperature gradient, which may be produced by the traveling waves themselves,\textsuperscript{2,1212} is also required. Early prototype devices could pump only nonconductive fluids,\textsuperscript{1211} but Fuhr et al.\textsuperscript{1212} have demonstrated EHD pumping of conductive fluids with conductivities between 0.0001-0.1 (ohm-m)\textsuperscript{-1} (e.g., up to -0.01M NaCl aqueous solutions) using high frequencies and a low voltage waveform. Typically, volumetric flow rates of $10^{10}$–$10^{12}$ m\textsuperscript{3}/sec and flow velocities of 50-1000 micron/sec were achieved using traveling wave frequencies of 0.1-30 MHz at ~40 volts across 30 micron-wide electrodes spaced 30 microns apart in 50 x 70 micron square channels ~4000 microns in length.\textsuperscript{1212} Flow velocity is a linear function of the square of the applied voltage and a complex function of the applied frequency. Further research is required to determine if unfavorable thermal scaling presents difficulties in submicron-size devices.

Other nonmechanical pumping methods have been studied. Electroosmotic pumping of electrolyte solutions has been observed in capillaries 50 microns in diameter or less, and separation of the components is also possible using electokinetic phenomena with applied voltages up to 10,000 volts.\textsuperscript{1213,1214} Electrowetting pumps using electrical control of the interfacial surface tension have produced -0.1 atm flow pressures in 10 micron radius channels.\textsuperscript{1211} Molecular flow can be induced by pulsed axial magnetic fields applied to carbon nanotubes placed in a transverse electric field. Ultrasonic pumps cause liquid to move in the direction of wave propagation with a speed proportional to the square of the acoustic amplitude;\textsuperscript{1214,1225} speeds of 130 micron/sec have been observed in water at 3.5 MHz.\textsuperscript{1214} Ferrofluids\textsuperscript{1216} containing ~10 nm particles in suspension and electro rheological fluids\textsuperscript{1227} allow electrostatic control of fluid flow via the Winslow effect,\textsuperscript{1229} although early reports that blood flow could be electrostatically regulated have not yet been confirmed experimentally.\textsuperscript{1240} Osmotic-pump motility across sugar gradients has been induced in micron-size phospholipid vesicles.\textsuperscript{1209}

Optofluidic pumping is extremely energy inefficient and is feasible only at dangerously high beam intensities. Consider a beam of light of radius $r_{\text{beam}}$ and intensity $I_{\text{beam}}$ that exerts a photon-pressure force of $F_{\text{beam}} = \pi \eta \tau_{\text{beam}}^2 k_\text{b} I_{\text{beam}} / c$, where $c = 3 \times 10^8$ m/sec (speed of light) and $k_\text{b} = 0$ for total transmission, 1 for total absorption, and 2 for total reflection. Setting this equal to the Stokes law force (Eqn. 9.73) on a spherical target of radius $r_{\text{target}}$ in a fluid of absolute viscosity $\eta$, gives the approximate magnitude of the fluid flow velocity as:

$$v_{\text{flow}} = \frac{k_\text{b} I_{\text{beam}}}{6 \eta c}$$  \hspace{1cm} \text{(Eqn. 9.40)}

if we conservatively assume that dimensionless $k_\text{b} = \tau_{\text{beam}} \alpha_{\text{light}}$, where $\alpha_{\text{light}} \approx 300$ m\textsuperscript{-1} (optical absorption coefficient; Section 4.9.4). Taking $I_{\text{beam}} \approx 10^{11}$ W/m\textsuperscript{2} (optical tweezer intensity; compare maximum “safe” in vivo intensity of $\lesssim 10^5$ watts/m\textsuperscript{2}; Section 6.4.3.2) and $r_{\text{beam}} = 200$ nm giving $k_\text{b} = 6 \times 10^{-5}$ (almost complete transmission), then $v_{\text{flow}} = 1$ micron/sec but with a huge power dissipation of $\pi \tau_{\text{beam}}^2 I_{\text{beam}} \approx 10$ milliwatts while developing a pumping pressure of only $P_{\text{pump}} \approx k_\text{b} I_{\text{beam}} / c = 2 \times 10^{-7}$ atm.

9.2.7.5 Fluid Mixing

In nanofluidic systems (e.g., pipes, chambers and junctions) of size $L \approx 100$ nm, even with strictly laminar flow, mixing of separate fluid streams is rapid because the small-molecule diffusion time $\tau \approx 10$ microsec (Eqns. 3.1 and Table 3.4). At the other extreme, systems of characteristic size $L \approx 1000$ microns with $v_{\text{flow}} \approx 1.4$ m/sec may readily establish a turbulent flow with $N_{\text{Lam}} \approx 2000$ in water at 310 K (Eqn. 9.29), and again mixing of multiple fluid streams can be quite rapid.

However, for systems of scale 1000 microns > $L > 100$ nm, diffusion times range from $10^{-3}$–$10^{-1}$ sec for small molecules, and even longer for large molecules. Nonlaminar flow would demand unreasonably high fluid velocities, and diffusive stirring (Section 3.2.3.1) yields only modest improvements in mixing. To achieve rapid mixing in this intermediate size range, a micromixer chamber is first filled with one liquid, and then the other liquid is injected into the chamber volume simultaneously through a large number of very small nozzles. In the simplest micromixer designs the mixing nozzles form a regular coplanar array.\textsuperscript{1215} In more complex designs, mixing nozzles will form a three-dimensional array possibly using a branching tree or fractal geometry with tip spacing ~1 micron to allow fully intimate mixing to be completed in ~1000 microsec; a ~100 nm tip spacing allows complete mixing in ~10 microsec.

9.2.7.6 Nanoplumbing and Fluidic Circuits

In fluid distribution systems that minimize total energy dissipation in laminar nonpulsatile flow, space-filling fractal networks of branching tubes are most efficient.\textsuperscript{3242} At each branch point in such a network, where a single large tube of radius $R$ bifurcates into $N$ branches with each branch tube having radii $r_1, r_2,\ldots, r_N$, Murray\textsuperscript{1220} found that shear stresses are equalized and flow impedance is minimized when:

$$R^3 = r_1^3 + r_2^3 + \ldots + r_N^3$$  \hspace{1cm} \text{(Eqn. 9.41)}

In the special case where $r_1 = r_2 = \ldots = r_N$, which is quite common in biological systems, Murray’s law reduces to $R^3 = N r_1^3$. For example, the human bronchial system typically has $N = 2$, so the ideal fractal network has $R/r - N^{1/3} = 1.26$; from Table 8.7, the actual value from trachea (generation 0) through the terminal bronchiole in generation 16 (after which alveoli begin to appear irregularly) is $R/r = 1.24$, in good agreement with Murray’s law. In turbulent flow regimes, the exponent on Murray’s law becomes 2.33, rather than 3.\textsuperscript{1361} And in pulsatile flow through elastic tubes, which dominates the aorta and major arteries in the human circulatory system, the energy minimization principle requires area-preserving branching, or $R^3 = N r^3$ for the ideal network.\textsuperscript{1369}

Complex nanoscale fluidic logic devices are readily imagined. While fractal geometries are likely, for simplicity consider a three-dimensional fluidic circuit of volume $V_{\text{circuit}}$. A volume fraction $f_{\text{tube}}$ consists of $N_{\text{tube}}$ fluidic gates and $N_{\text{valve}} = 1$ independent fluidic pathways each of length $l_{\text{tube}}$ and radius $r_{\text{tube}}$, and a second volume fraction $f_{\text{valve}}$ consists of $N_{\text{valve}}$ fluidic gates each of volume $V_{\text{valve}}$. A volume fraction $f_{\text{tube}}$ consists of $N_{\text{tube}}$ fluidic gates and $N_{\text{valve}} = 1$ independent fluidic pathways each of length $l_{\text{tube}} = 230$ nm. If typical flow velocity $v_{\text{flow}} = 1$ mm/sec,\textsuperscript{1228} then from Section 9.2.5 the total volume flow rate through the circuit is $N V_{\text{flow}} \approx 4000$ micron$^3$/sec at a pressure differential $\Delta P = 0.1$ atm. Circuit power dissipation (ignoring valve dissipation) is $N P_{\text{flow}} \approx 50$ pW and $t_{\text{flow}} \approx 0.2$ millisec assuming fully
parallel operation (e.g., a path length $l_{\text{tube}}$), allowing a circuit operating frequency of $f_{\text{low}} = 1 - 5000$ Hz.

Capillary networks are readily gated by applying appropriate voltages, allowing valveless switching of liquid flow among various fluidic pathways. Purely electrical valving of fluid flows has been common practice in neurobiological research for decades. For example, the flow of acetylcholine from the open end of a micron-sized pipette (as it is inserted into tissue) is prevented by applying a negative bias or braking current of 3 nanoamps at the tip; flow resumes when the braking current is removed.803 Constant-discharge flow-control valves using opposing polymer brushes with ~25 chains of ~40 monomers each498 and micron-scale electrorheological diodes498 have been investigated. Macroscale fluidic NOR logic elements that can be used to construct arbitrary Boolean logic circuits* (for controlling materials flow) are widely available commercially,127 although these often employ the Coanda effect, vortices, or turbulence effects, which generally don’t scale well to micron and submicron devices. In 1997, a microfluidic chip-based system for the integration of high-throughput drug discovery efforts was demonstrated by SmithKline Beecham and Orchid Biocomputer.1222 This microfluidic chip incorporated microfabricated components for valving and pumping of organic solvents using electrokinetic transport within a three-dimensional fluidic network. The pumping and valving mechanisms had no moving parts, “making large scale integration feasible and inherently reliable.” The study of microfluidic networks1228,2697,2698 using hundreds of micron-wide channels was an active research area in 1998, and integrated chemical systems were being widely discussed.121

9.2.7.7 Containerized Flow

Once a block of fluid has passed into the interior of a nanodevice, it may be containerized or divergently subdivided and then transported at velocities far higher than diffusive flow (Eqn. 3.1) or Poiseuille flow (Eqn. 9.25) would otherwise permit. The velocities available in containerized flows are largely independent of fluid viscosity and total transport time (unlike diffusive flow), and are nearly independent of fluid pressure and flow channel length (unlike Poiseuille flow).

For easy comparison, consider a disk-shaped container of radius $r_{\text{block}}$ and thickness $h_{\text{block}}$ with useful fractional storage volume $f_{\text{block}}$ (to account for container wall thickness and mechanism overhead). After loading with fluid, the container is transported through a channel of radius $r_{\text{tube}} = r_{\text{block}}$ and length $l_{\text{tube}}$ at a velocity $v_{\text{block}}$, giving a volumetric flow rate of $V_{\text{block}} = \pi r_{\text{tube}}^2 f_{\text{block}} v_{\text{block}}$ (m$^3$/sec) which is superior to the Poiseuille flow rate in a pressurized tube of equivalent size when:

$$v_{\text{block}} > \frac{r_{\text{tube}}^2 \Delta p}{8 \eta l_{\text{tube}} f_{\text{block}}} \quad \text{[Eqn. 9.42]}$$

Containerized transport thus may be preferred to Poiseuille flow in narrow, long transport channels, or in cases where fluid pressure is relatively low or fluid viscosity is relatively high, or where containers can be very thin-walled (e.g., high $f_{\text{block}}$) such as when strong diamondoid containers are used. For example, a containerized transport channel carrying water at 310 K with $r_{\text{tube}} = 10$ nm, $l_{\text{tube}} = 1$ micron, $\Delta p = 1$ atm, and $f_{\text{block}} = 0.9$ produces volume flow rates superior to Poiseuille flow if container transport velocity $v_{\text{block}} > 0.002$ m/sec. Container/channel sliding interfaces moving up to 1 m/sec have low frictional losses,10 and speeds up to ~100 m/sec can probably be tolerated in internal mechanisms if flow speed (entirely inside nanodevices) is a more important design criterion than minimizing power dissipation in a particular application. An additional consideration is that subdividing a volume of fluid into many smaller blocks of fluid allows these smaller blocks to be transported independently through numerous spatially independent channels of much smaller bore, greatly increasing the number of flow design options and greatly increasing the relative advantage over Poiseuille flow through channels of similar size.

9.3 Nanomanipulators

Numerous simple actuators have been developed for microelectromechanical systems or MEMS,1232,1253-1255,1267-1269 A few of these designs might possibly be useful in early nanomedical systems. All major components of an articulated micromanipulator having multiple degrees of freedom, workspaces on the order of 1 mm$^3$, positional resolution up to ~10 nm, operating frequencies up to 10 KHz, and slewing speeds up to 4 mm/sec have been demonstrated experimentally; complete systems have even been proposed.346,1255,1254

Actuation and manipulation may be very broadly defined. For example, self-assembling and self-disassembling -300 micron gold foil cubes have been fabricated and cycled repeatedly at -1 Hz;1251 “silicon origami” has also been described.1368 Nonmechanical actuation has also been explored. In the Magnetic Stereotaxis System,1246 a helmet with a cubic array of six superconducting coils is used to apply a 0.2 N force to a 3.2-mm wide, 4.7-mm long permanently magnetized cylindrical pellet that is embedded in brain material, causing stepped movement of the pellet through the brain material to a specified location with ~1 mm placement accuracy without any direct physical contact.

However, in the usual definition of a manipulator, an ergomechanical transducer (e.g., motors; Chapter 6) converts chemical, electrical, mechanical, acoustical, or some other form of energy into the mechanical energy of a manipulatory device which then exerts useful forces on objects or materials in the environment. In most cases a force-transmitting physical structure is required which may include both heavy load-bearing elements such as girders, struts, pneumatic tubes or rotary joints, and fine control elements such as stress cables, valves or switches (Section 9.3.1). An end-effector (e.g., gripper, cutter or molecular jig) is usually employed to focus and redirect the application of force at the distal end of the manipulation device (Section 9.3.2), or to perform specialized tasks. Sensors permit feedback control of manipulator motions as well as verification that the assigned manipulatory task has been properly executed (Section 9.3.3). Large numbers of manipulators can form ciliary arrays, allowing convenient mass transport of materials (Section 9.3.4); other classes of manipulators may be useful in bulk disassembly of materials (Section 9.3.5).

The diverse concepts outlined in this Section demonstrate the enormous range of possible nanomanipulator designs. These concepts are not intended as specific engineering proposals but rather as illustrations of certain useful classes of manipulators representing alternative design pathways that could provide the desired capability. No effort has been made to produce optimal systems that minimize energy dissipation, maximize speed, minimize parts count, or simplify manufacture. Biocompatibility issues will be addressed in Chapter 15.

* In the 1950s, Marvin Minsky and Rollo Silver built a “hydropip computer” using hydraulic logic elements consisting of millimeter-wide grooves and holes in multiple layers of plastic sheets with small rods and balls inserted in some of the grooves. When the assembly was pressed together and connected to a water supply, it became a hydraulic computer powered by a 3-inch high column of water, operating at ~30 Hz.
9.3.1 Nanoscale Manipulators

9.3.1.1 Biological Cilia

Biological cilia are motile, hairlike cylindrical processes typically 200-300 nm in diameter, 2-20 microns in length, and \(-10^{16}-10^{17}\) kg in mass. They are found in large numbers on some cell surfaces. Unicellular organisms use cilia both for locomotion and for food collection, while multicellular organisms use cilia primarily for mass transport of the environment past the cell.

Each cilium is bounded by an extension of the cellular plasma membrane (Fig. 9.2A), rising from a cytoskeleton-anchored centriole-like basal body consisting of an array of 9 triplets of common-walled microtubules -150 nm in diameter and 300-500 nm in length. Sprouting upward from the basal body is the primary ciliary structure, called the axoneme. The axoneme is a cylinder of microtubules (Section 8.5.3.11) with 9 outer doublets of microtubules and two additional complete microtubules in the center, often called the central pair (Fig. 9.2B). The 9 outer doublets are extensions of two of the three microtubules comprising each of the nine basal body triplets. Each outer doublet of the axoneme consists of the A tubule (a complete 25-nm outer-diameter microtubule with the usual 13 protofilaments made of tubulin dimers), the B tubule (an incomplete microtubule with only 10-11 protofilaments), and a common wall containing the protein tektin (an intermediate-filament-like protein with high tensile strength) (Fig. 9.2C).

A set of 1-2 megadalton dynein side arms project outward from each of the A tubules, reaching clockwise toward the B tubules of the adjacent doublet (Fig. 9.2D). The dynein arms occur in pairs, one inner arm and one outer arm, spaced at regular ~30 nm intervals along the microtubule. Nexin protein interdoublet links at wider intervals join the outer doublets making circumferential rings, and radial spokes project inward from each of the 9 microtubule doublets. Microtubules do not change length during ciliary movement. Rather, the stalk of each dynein arm attaches to and detaches from the microtubules along the microtubule. Nexin protein interdoublet links at wider one inner arm and one outer arm, spaced at regular ~30 nm intervals to the adjacent doublet (Fig. 9.2D). The dynein arms occur in pairs, reaching clockwise toward the B tubules of the adjacent doublet (Fig. 9.2D). The dynein arms occur in pairs, one inner arm and one outer arm, spaced at regular ~30 nm intervals along the microtubule. Nexin protein interdoublet links at wider intervals join the outer doublets making circumferential rings, and radial spokes project inward from each of the 9 microtubule doublets.

In early experiments, bullfrog gullet ciliary surfaces developed a beating frequency of ~30 Hz with an angular velocity of 12 deg/millisecond and a tip velocity of 2500 microns/sec in the effective stroke, while the bend in the recovery stroke propagates at ~350 microns/sec. Taking Pcilium = 0.05 pW and vcilium = 500 microns/sec, then Fcilium = Pcilium / vcilium ~ 0.1 nN.

A simple cilium 12 microns long may beat at a frequency νcilium ~ 30 Hz with an angular velocity of 12 deg/millisecond and a tip velocity of 2500 microns/sec in the effective stroke, while the bend in the recovery stroke propagates at ~350 microns/sec. Taking Pcilium = 0.05 pW and vcilium = 500 microns/sec, then Fcilium = Pcilium / vcilium ~ 0.1 nN.

9.3.1.2 Nanociliary Manipulators

Equally complex nanomechanical ciliary systems could probably be designed. But for simplicity, consider a nanociliary manipulator patterned more closely after the muscular trunk of the elephant or tentacle of the squid and based only loosely on the biological cilium. The model consists of a cylindrical central shaft of radius rshaft and length lshaft, the tip of which is bonded to one or more conductive shaft-attached control cables (radius rcable) whose lengths and tensions can be predictably altered, causing the shaft to flex. A cable pulled with a force Fcable applies a transverse deflection force to the nanocilium shaft of approximate magnitude Fdeflect ~ (rcable/lshaft) Fcable for small deflections. The force required to deflect the terminus of a simple cantilevered rod is Fdeflect = 3 Eshaft Ishaft / lshaft4, where Eshaft is Young’s modulus and Ishaft = π rshaft4 / 4 is the second moment of area (also known as the moment of inertia of the area) for a shaft of circular cross-section.

Solving for the deflection gives:

\[
d = \left(\frac{4}{3} \pi \right) \left(\frac{r_{\text{cable}}}{l_{\text{shaft}}} \frac{l_{\text{cable}}^2}{I_{\text{cable}}} \frac{F_{\text{cable}}}{F_{\text{shaft}} r_{\text{shaft}}} \right) \left(\frac{l_{\text{shaft}}}{r_{\text{shaft}}} \right)^4
\]

[Eqn. 9.43]

and a deflection angle \( \theta_{\text{shaft}} = \tan^{-1} \left( d / l_{\text{shaft}} \right) \). The shaft begins to buckle when Fcable exceeds the Euler force:

\[
F_{\text{buckle}} = \frac{\pi^2 E_{\text{shaft}} I_{\text{shaft}}}{l_{\text{cable}}^2} = \frac{\pi^3 E_{\text{shaft}} r_{\text{shaft}}^4}{4 l_{\text{shaft}}^3}
\]

[Eqn. 9.44]
The maximum force that can be transmitted through a diamondoid cable of circular cross-section and tensile strength $T_{\text{str}} = 1.9 \times 10^{11} \text{N/m}^2$ with $r_{\text{cable}} = 1 \text{ nm}$ is $F_{\text{max}} = \pi r_{\text{cable}}^2 T_{\text{str}} = 600 \text{ nN}$. (For hollow cylinders of inside radius $r$ and outside radius $R$, $r_{\text{cable}}$ should be replaced by $(R^2 - r^2)$ in Eqn. 9.44.) For a large nanocilium, we take $r_{\text{shaft}} = 50 \text{ nm}$, $l_{\text{shaft}} = 1000 \text{ nm}$, and $E_{\text{shaft}} = 10^8 \text{ N/m}^2$ (tendon strength). At maximum deflection where $F_{\text{cable}} = F_{\text{max}} = 600 \text{ nN}$, then $d = 400 \text{ nm}$, $\theta_{\text{shaft}} = 22^\circ$, and bending stiffness $k_{\text{shaft}} = F_{\text{deflect}} / d = (3 \pi / 4) (E_{\text{shaft}} r_{\text{shaft}}^3 / l_{\text{shaft}}^4) \approx 0.001 \text{ nN/nm}$. Maximum lateral force that may be applied by the nanocilium tip is $F_{\text{deflect}} = 0.6 \text{ nN}$; setting this force equal to the Stokes law force (Eqn. 9.73) gives a maximum nanocilium velocity through water of $-10 \text{ cm/sec}$. If cable tension is uniformly distributed over a shaft endplate of radius $r_{\text{shaft}}$, the maximum mechanical pressure is $p_{\text{shaft}} = F_{\text{max}} / (\pi r_{\text{shaft}}^2) = 8 \times 10^9 \text{ N/m}^2 < T_{\text{str}}$. Shaft buckling commences at $F_{\text{buckle}} = 5 \text{ nN}$ with $d = 3 \text{ nm}$ and $\theta_{\text{shaft}} = 0.2^\circ$; classical positional variance of the tip due to thermal noise is $\Delta x = (kT / \nu)_{\text{shaft}}^{1/2} \approx 0.1 \text{ nm}$, maximum lateral displacement. Setting this force equal to the Stokes law force (Eqn. 9.73) at the maximum $-10 \text{ cm/sec}$ tip velocity.

For a very small nanocilium, we take $r_{\text{shaft}} = 10 \text{ nm}$, $l_{\text{shaft}} = 100 \text{ nm}$, and $E_{\text{shaft}} = 10^9 \text{ N/m}^2$ where $F_{\text{cable}} = F_{\text{max}} = 600 \text{ nN}$, then $d = 25 \text{ nm}$, $k_{\text{shaft}} = 14^\circ$, $k_{\text{shaft}} = 6 \text{ nN/nm}$, $F_{\text{deflect}} = 6 \text{ nN}$, $p_{\text{shaft}} = 2 \times 10^9 \text{ N/m}^2 < F_{\text{cable}} = 78 \text{ N/N}$, maximum angular coordinate steering of tip position. Adding additional control cable triplets on successive connection rings spaced along the length of the shaft allows independent addressing of intermediate shaft segments to permit the shaft to flex in a variety of sinuous shapes with some additional control of total radial deflection. Two orthogonal pairs of tensioned cables attached on either side to ball joints separating each segment gives a more complex but highly redundant and highly controllable articulated manipulator (Fig. 9.3).

### 9.3.1.3 Pneumatic Manipulators

Differential tensions also can be applied to a nanomanipulator structure using gases (e.g., pneumatic actuators) or liquids (e.g., hydraulic actuators). Nanoscale pneumatic actuator elements must be of size $L >> r_{\text{gas}}$ (Eqn. 9.23) to ensure continuum flow. Thus for $L = 1 \text{ micron}$, minimum pneumatic operating pressure $p_{\text{gas}} >> 0.2$ atm; for $L = 10 \text{ nm}$, minimum $p_{\text{gas}} >> 24$ atm.

There are many examples of pressurized hydraulic tissue inflation in macroscale biology, including sea anemones, nematodes, echinoderm tube feet, spider legs, pupal butterfly wing extension, piscine swim bladders, and human sex organs. The pressurized pumping of DNA through a nanotube in the T4 bacteriophage has already been described (Section 9.2.4). As another possible example in nanobiology, the sperm of the sea cucumber *Thyone* rapidly elongates its ~700-nm wide spherical acrosomal vesicle into a 50-nm wide, 90-micron long tubular acrosomal process in less than 10 sec after a rapid influx of water doubles the volume of the 0.5 micron$^3$ acrosomal cup in which the vesicle is stored in 50-70 millisec (Fig. 9.4). This has led to the proposal that the elongation is at least partly hydraulically driven by a profilactin osmotic pressure of ~0.5 atm, because the speed of actin polymerization alone appears to be about an order of magnitude too slow.

Nanoscale manipulators may be constructed of simple orthotropic tubes with flexible ribbing and an elastic but nonextensible spine that allows manipulator flexure in only one direction when pressurized (Fig. 9.5). Experiments using compound manipulators constructed using centimeter-sized pneumatic orthotropic segments that were pressurized to 3-6 atm produced 90˚ flexures in 0.3-1 sec with 2-5 mm placement accuracy (0.4%-0.9%) and load-bearing strengths up to ~20 N at the tip.

Trying to construct highly articulated manipulators with such a complex arrangement may be unrealistic, but a simpler alternative is to use a flexible tube that expands and contracts at different rates. For a tube of radius $r$, the volume change $\Delta V = 2\pi r l \Delta r$ gives a pressure change $\Delta p = \pi r^2 l \rho \Delta V / V$ where $\rho$ is the density of the fluid. For a rapidly elongating vesicle, the volume change is $\Delta V = 2\pi r l \Delta r$ giving a rapid influx of water doubles the volume of the 0.5 micron$^3$ acrosomal cup in which the vesicle is stored in 50-70 millisec. This has led to the proposal that the elongation is at least partly hydraulically driven by a profilactin osmotic pressure of ~0.5 atm, because the speed of actin polymerization alone appears to be about an order of magnitude too slow.

![Fig. 9.3. Two-dimensional representation of segmented 3-D manipulator using ball joints (modified from Ma et al.)](image)

![Fig. 9.4. Schematic of the acrosomal process in *Thyone* (modified from Oster and Perelson).](image)
the three chambers. In a second configuration, spiral reinforcing fibers ($\alpha = 5^\circ$-$20^\circ$) force a circumaxial rotation of the cylinder when chamber pressure is applied, allowing useful screwing motions to be generated at the tip. Thus the manipulator can achieve pitch, yaw, stretch, and rotation, although these four movements are not fully independent because there are only three independent control parameters.

Consider the tri-chambered pneumatic manipulator of Figure 9.6 with one end fixed, $\alpha = 0^\circ$, tube radius $r_{\text{tube}}$, tube wall thickness $t_{\text{wall}}$, and chambers of equal relaxed length $L_0$ along the center axis that are then pressurized to $p_1$, $p_2$, and $p_3$, respectively. The fiber-reinforced tube wall material has Young's modulus $E_{\text{tube}}$ in the transverse direction. Then the $x$-$y$ projection of manipulator rotation around the vertical axis ($\theta$) is given by:

$$\theta = -\frac{\tan^{-1}\left(\frac{2(p_1 - p_2 - p_3)}{3^{1/2}(p_2 - p_3)}\right)}{\tan^{-1}\left(\frac{2(p_1 - p_2 - p_3)}{3^{1/2}(p_2 - p_3)}\right)}$$  \hspace{1cm} \text{(Eqn. 9.45)}$$

For each of the three chambers, $\theta_1 = \theta$, $\theta_2 = \theta + (2 \pi / 3)$, and $\theta_3 = \theta - (2 \pi / 3)$. The angular deflection of the manipulator down from the vertical axis ($\psi$) is the same for all three chambers, given as:

$$\psi = \frac{L_{\text{manip}}}{R_{\text{curve}}}$$  \hspace{1cm} \text{(Eqn. 9.46)}$$

$L_{\text{manip}}$ is the new stretched length of the manipulator at the center axis, $R_{\text{curve}}$ is the radius of curvature of the manipulator, and the linear deflection from the vertical axis is:

$$d = (1 - \cos(\psi)) R_{\text{curve}}$$  \hspace{1cm} \text{(Eqn. 9.47)}$$

where:

$$\begin{align*}
L^2_{\text{manip}} &= (L_0) L_{\text{manip}} = \left(\frac{\pi r_{\text{tube}} L_0}{36 E_{\text{tube}}}\right) \sum_i (p_i) = 0 \hspace{1cm} \text{(Eqn. 9.48)} \\
R_{\text{curve}} &= \left(\frac{9 E_{\text{tube}}}{2 \pi}\right) \left(\frac{\pi r_{\text{tube}} + 2 L_{\text{manip}}}{\sum_i (p_i \sin(\theta_i))}\right) \hspace{1cm} \text{(Eqn. 9.49)}
\end{align*}$$

and Eqn. 9.48 is readily evaluated using the quadratic formula. Making a few simplifying assumptions, the bending stiffness of the manipulator is given approximately as:

$$k_{\text{tube}} \approx \frac{3^{3/2} \pi E_{\text{tube}} r_{\text{wall}} t_{\text{tube}}^3}{L_{\text{manip}}^3} \hspace{1cm} \text{(Eqn. 9.50)}$$

and thus the maximum force that can be exerted at the tip is:

$$F_{\text{tip}} = k_{\text{tube}} d$$  \hspace{1cm} \text{(Eqn. 9.51)}$$

Finally, the lowest resonant frequency $\nu_{\text{res}}$ of a tri-chambered pneumatic manipulator of mean density $\rho_{\text{manip}}$, mass $M_{\text{manip}} = \pi r_{\text{tube}}^2 L_0 \rho_{\text{manip}}$, and carrying a tip load of $M_{\text{load}}$ is given by:

$$\nu_{\text{res}} = \left\{\frac{3 E_{\text{tube}} L_{\text{tube}}}{4 \pi^2 L_{\text{manip}}^{1/2} (2.36 M_{\text{manip}} + 0.236 M_{\text{load}})}\right\}^{1/2} \hspace{1cm} \text{(Hz)}$$  \hspace{1cm} \text{(Eqn. 9.52)}$$

where the second moment of area $I_{\text{tube}} = (1/4) \pi r_{\text{tube}}^4 + (1/2) L_{\text{manip}} r_{\text{tube}}^3$. Manipulator power requirement in the absence of a viscous operating medium is:
where $\nu_{\text{manip}}$ is manipulator operating frequency.

As an example of a nanomedically useful pneumatic manipulator, consider a tri-chamber design with $L_0 = 1200$ nm, $r_{\text{tube}} = 200$ nm, $v_{\text{wall}} = 50$ nm, $E_{\text{tube}} = 10^8$ N/m$^2$, and maximum operating pressures up to ~1000 atm (pressure/volume relationships for air calculated using the ideal gas law are ~5% in error at 100 atm, ~50% in error at 1000 atm; Table 10.2). Application of a $p_1 = 6$ atm pressure pulse to one chamber of the manipulator produces an $F_{\text{tip}} = 0.1$ nN lateral force and a $d = 1$ nm tip displacement at zero load, costing ~3 kT/cycle at $T = 310$ K. Application of a 60 atm pulse produces ~1 nN and a 10 nm displacement (~300 kT/cycle); a 600 atm pressure step for this design.

$\text{Manipulator volume is } 0.15$ micron$^3$ and manipulator mass $M_{\text{manip}} = 0.1$ picogram. At zero load in vacuo, $v_{\text{res}} = 30$ MHz up to ~1000 atm; $v_{\text{res}}$ falls to ~5 MHz for a 1 picogram load (roughly the mass of a 1 micron$^3$ nanorobot) and to ~1 MHz for a ~30 picogram load, so up to ~MHz operating frequencies seem reasonable at modest loads. (For comparison, a 1 nN lifting force supports a 100,000 picogram mass against gravity.) At maximum cyclical tip displacement $d = 100$ nm and $v_{\text{manip}} = 1$ MHz, mean tip velocity is $v_{\text{tip}} = d v_{\text{manip}} = 10$ cm/sec. Maximum viscous drag power in vivo is $P_{\text{drag}} = 140$ pW using the formula for Stokes frictional dissipation (Eqn. 9.74) at a 10 cm/sec tip velocity and taking $r = (2 r_{\text{tube}} L_0)^{1/2}$ for the manipulator and $\eta = 1.1 \times 10^{-3}$ kg/m-sec in tissue plasma at 310 K.

The pneumatic design outlined above is quite versatile. Manipulator tip displacement scales as $d = L_0 p / E_{\text{tube}}$. Manipulator tip force scales as $F_{\text{tip}} = r_{\text{tube}}^3 / L_0^2$; tip force is nearly independent of $E_{\text{tube}}$ down to $10^8$ N/m$^2$, below which $F_{\text{tip}}$ falls rapidly. Pneumatic nanomanipulators using diamondoid wall materials (e.g., $E_{\text{tube}} \sim 10^{11}$ N/m$^2$) can have a stiffness on the order of $k_{\text{tube}} \sim 10^{-10}$ nN/nm (comparable to the telescoping manipulator described in Section 9.3.1.4), although in such cases the maximum pneumatic tip displacement even at maximum operating pressures may be limited to ~1 nm or less. Longitudinal cables can be added to actively control the helical pitch (e.g., $\alpha$) of the reinforcing fibers; adding both circumferential and longitudinal reinforcing fibers allows deformation of compliant tube structures into multiply-coiled shapes. Adding tensioned control cables and additional layers of pneumatic chambers can give a full six degrees-of-freedom (DOF) manipulator, albeit with some increased complexity of control. The performance of individual axial segments (e.g., selective locking/unlocking or differing Young’s moduli of adjacent segments) allows arbitrary serpentine flexures through a three-dimensional work volume (Fig. 9.7). End effector control lines may be routed along the central axis.

Many interesting designs for “tentacle” or “snake” manipulators, also known as highly-redundant or hyper-redundant manipulators, have appeared in the literature. These may be fabricated as stacks of multiple pneumatic segments connected coaxially. A biological snake may have ~200 separate jointed segments, though each joint has a pitch range of only $\pm 10^\circ$; snake robots having artificial joints with pitch ranges up to $135^\circ$ have been built. Such designs are useful in nanomedical robotic systems because hyper-redundant manipulators can be made extremely fail-safe—if one joint fails (e.g., by locking tight, not falling limp, by design), the accessible work volume is not significantly diminished.

### 9.3.1.4 Telescoping Manipulators

Another useful class of manipulators is the telescoping design (Fig. 5.14), in which snugly fitted or threaded components extend or retract via sliding motions induced either pneumatically or by shaft-driven rotations of linked elements. The best-known telescoping nanomanipulator design has been proposed by Drexler in the context of a stiff mechanism that may be used for precision molecular positioning and assembly work. As depicted in Figure 9.8, the mechanism features a central telescoping joint whose extension and retraction is controlled by a 1.5-nm diameter drive shaft. The rapid rotation of this drive shaft (up to ~1 m/sec tangential velocity) forces a transmission gear to quickly execute a known number of turns, causing the telescoping joint to slowly unscrew or screw in the axial direction, thus lengthening or contracting the manipulator. Additionally, two pairs of canted rotary joints—one pair between the telescoping section and the base, the other pair between the telescoping section and the
working tip—are controlled by toroidal worm drives. These joints enable a wide variety of complex angular motions and give full 6-DOF access to the work envelope (Fig. 9.9: isolated planetary gear (Section 2.4.1) shown at lower left, for comparison). By engaging and disengaging various drive shafts using clutches, these shafts can be made to turn through a known number of rotations between locked states giving odometer-like control of manipulator joint rotations. The power supply and drive shaft control systems originate outside the manipulator, thus are not subject to tight geometric constraints.

Drexler’s telescoping manipulator is approximately cylindrical in shape with an outside diameter of ~35 nm and an extensible length from 90 nm to 100 nm measured from top of base to working tip. The manipulator includes a hollow circular channel 7 nm in diameter to allow tool tips and materials to be moved from below the manipulator through the base up to the working tip. At the tip, a slightly larger region is reserved for a mechanism to allow positioning and locking of tool tips. Device stiffness (which scales roughly with arm length) is \( k = 25 \text{ nN/nm} \), so that a force of 1 nN may be applied with an elastic deflection in the arm of only 0.04 nm. The classical positional variance of the manipulator at \( T = 310 \text{ K} \) is \( \Delta x = (kT / k)_{1/2} = 0.01 \text{ nm} \), although with a manipulator tube segment thread pitch of 0.5 nm and drive shafts that execute ~700 turns to produce a single tube segment rotation, segment extension may be altered in ~0.001 nm steps. The workspace includes a 180° hemispherical-shell volume ~100 nm in diameter and ~10 nm deep. Conveyance of the tip through a full 100-nm arc requires ~10 microsec at a conservative ~1 cm/sec arm speed. Efficient task planning permits much smaller motion arcs, allowing ~MHz operating frequencies. The manipulator dissipates 0.1 pW of power (power density ~105 watts/m3) while in motion under no-load conditions. Energy dissipated per unit mass of delivered payload is scale-invariant, with increasing size.10 Power dissipation scales roughly as the square of tip velocity; 11 power density scales inversely with arm length.

Like the ciliary and pneumatic designs, the telescoping manipulator is hermetically sealed, thus maintaining a controlled internal environment while allowing leakage proof operation in vivo.

### 9.3.1.5 Stewart Platform Manipulators

Another class of simple manipulators is the 6-DOF Stewart platform.12,13 The Stewart platform exploits the geometry of an octahedron—one triangular face serves as a mobile platform and the opposing triangular face serves as the fixed base. Six struts of the octahedron connect the base to the platform and can be varied in length, thus controlling platform position and orientation with respect to the base. The six struts carry only compressive or tensile, but no bending, forces. Figure 9.10 shows a two-dimensional analog with three struts and 3 DOF to illustrate a portion of the available range of motions. This \( 5 \times 10^{-20} \text{ kg} \) device has pressure-actuated vernier ratchets that allow strut lengths to be altered in ~0.1 nm increments across a 100-nm work envelope, providing excellent positional control.10 Merkle12 has examined a novel family of 6-DOF positional control devices—the double tripod configuration—which offers greater stiffness for a given structural mass than a robotic arm and significantly greater range of motion than a Stewart platform. Tradeoffs between stiffness and range of motion in the double tripod can be continuously adjusted by altering simple design parameters.

![Fig. 9.9. External shape and range of motion of a telescoping nanomanipulator (adapted from Drexler).](image1)

![Fig. 9.10. Two-dimensional schematic of the range of motion of a Stewart platform manipulator (adapted from Drexler).](image2)

Assembly of parts while immersed in a fluid eliminates electrostatic and surface tension effects.1150 In the earliest, most primitive Stewart platform type nanomanipulators operating in aqueous suspension, R. Merkle suggests that leg extension and retraction may be controlled by manufacturing the legs using a repeating sequence of binding sites on the surfaces of two opposing struts. Chemical binding molecules are designed that can selectively link one strut to another, but only between one specific pair of binding sites (one site on each strut). Varying the linking agent controls which part of a strut links to the adjacent strut; cycling of link molecule types in solution forces the linking and unlinking in such a way as to move the struts along.1083

### 9.3.1.6 Metamorphic Manipulators

Metamorphic manipulators (Section 5.3.1.3) are fully reconfigurable manipulative or locomotive appendages that can be deployed from compact storage volumes. Surface area, contour, volume, length, and stiffness of the protuberance may be varied at ~KHz frequencies (Section 5.3.1.4). Metamorphic manipulators exuded from a nanorobot surface may temporarily assume many forms including arms/legs, hands/fingers, barriers and surrounds,
grapples, bubbles, and a wide variety of geometric shapes, all under programmatic or stochastic control (Section 9.3.3).

Many examples of highly plastic protuberances are known in microbiology at many size scales, such as the pseudopods of fibroblasts and amoebas (an amoebic pseudopod can extend and engulf a 100-micron paramecium in ~10 sec, applying a stress of 1462 up to ~2900 N/m²), blebbing and microspikes, the ~40 nm import receptors (cytoplasmic filaments) ringing the mouth of the nuclear pore complex, cyclosis in giant green algae cells at velocities up to 100 microns/sec, and the flexible plasma membrane extensions that occur during amoeboid locomotion (Section 9.4.3.7). Human axons elongate at ~10 nm/sec. The green plastid tubules of chloroplasts in plant cells are yet another example. These tubular projections arise initially as protuberances from the chloroplast surface that elongate and extend. They are dynamic in the living cell, continuously changing their shape, moving around, shrinking back, and sometimes connecting with other chloroplasts. Their width ranges from 350-850 nm, with lengths up to 15 microns. Plastid protuberances may extend at a velocity of 10-100 nm/sec and encircle another body, much like a finger curling around a sphere. Metamorphic manipulators are expected to present significant, though probably not insurmountable, design challenges.

9.3.2 Nanoscale End-Effectors and Tool Tips

An end-effector is usually placed at the tip of a manipulator in order to focus and redirect the application of forces. Examples of simple force redirection include a screwdriver head or bolting socket, a cutting blade, a perforative needle, and levers, pry bars or wedges.

End-effectors can also be used to achieve very fine motion control at the tip. Perhaps the simplest task such an end-effector can perform is the grasping or gripping of target objects in the environment. Figure 9.11 offers a small sampling of the many possible schematic concepts for grasping end-effectors. Mechanical grippers as in (A) have been fabricated with a jawspan of ~10 microns, tines ~2 microns thick, and a ~5 KHz resonant frequency, and have been used to grasp individual 2.7-micron polystyrene spheres, dried red blood cells, and even various protozoa including a 7-micron diameter, 40-micron long Euglena. Pneumatic grippers (B) have been built and operated on a larger scale. Figure 9.11(C) illustrates how high-frequency acoustic point sources or ultrasonic resonance fields can be used to reversibly secure a small target object in a small space; electrodynamic fields, electrohydrostatic forces, and optical tweezers employ similar principles. The suction gripper (D) works on any target immersed in a gas or liquid environment (Eqn. 9.22), whereas the magnetic gripper (E) only works on magnetic materials. The electrostatic gripper (F) induces attractive Coulombic, image force, or dipole charge effects, and is helped by the increased breakdown field strength of very small gaps due to the Paschen effect. The van der Waals gripper shown in Figure 9.11(G) uses variable nanoscale surface roughness to regulate the van der Waals attractive force; fluid surface tension or related forces may be modulated to reversibly grip the target. Finally, any soft object in the environment can be impaled by a narrow articulating spike, which then drops anchor inside the object until the gripper is ready to withdraw (Fig. 9.11(H)).

Most of the end-effectors shown in Figure 9.11 involve direct physical contact with the target object. Undesired or accidental surface adhesivity (Section 9.2) of grippers and manipulators is a serious design issue that must be addressed with respect to all possible environments and target objects likely to be encountered by the manipulation mechanism during the performance of its mission. Complementary diamondoid surfaces in vacuo may adhere with a contact tensile strength up to ~10⁴ atm, representing ~1 nN of contact force per nm² of contact area and releasing ~5400 zJ/nm² of energy when the surfaces are joined. However, noncomplementary surfaces in fluid environments can in theory be engineered to display almost any interfacial adhesivity desired—including negative adhesivity or repulsion (Section 9.2.3). Gripper/object adhesivity may vary according to many factors that are subject to design control. For example, all else equal, passivated diamond is 5-10 times more “sticky” in the van der Waals regime than glass or plastic (Table 9.1). Manipulator mechanism surfaces should in general be highly noncomplementary to contain surfaces, biological surfaces, nanorobot surfaces, and the surfaces of adjacent manipulators. Contact electrification is minimized by using materials with a small contact potential difference between the gripper and the target object.

A wide variety of specialized complex nanomedical tool tips are readily imagined, although an exhaustive listing or description is beyond the scope of this book. Among the more interesting possible functions (offered without regard to specific implementation) are the following:

1. Injectors—pipettes, glue applicators, aerators/foamers, and compound grasper/liquid injectors.
2. Adhesion Antennae—partially selective binding tips that are swept through the environment, wherein desired moieties or particles adhere and can be removed from the environment or drawn into the nanodevice (i.e., selective biochemical “dust mops”); transmembrane in cyto sensory appendages.
3. Core Sampler—a rapidly-spinning hollow cylindrical cutting tool that is driven into nearby tissue, whereupon suction is applied, an irising diaphragm closes at the distal end which separates the sample from the main tissue mass, and then the
tool is withdrawn, extracting a full cylindrical core sample of neatly excised tissue that is ready for further analysis or transport. A similar tool may be required for ice burrowing (Section 10.5.2).

4. **Noncovalent Biochemical Welders**—rejoins noncovalently linked biological structures that have become detached (Section 9.4.4.3); closes gaps in lipid bilayer membranes after being breached by the passage of an nanomechanical appendage or entire nanorobot device (Section 9.4.5.6); reattaches vesicles and organelles to microtubular tracks; reseals intercellular gap junctions.

5. **Joining and Spooling Tools**—bobbins, spinning and spooling tools; fiber manipulators including stitching, stapling and threading devices; rivet pin inserters.

6. **Compression and Compaction Tools**—extruders, compactors, compression rollers, beam benders, crimpers, and compression fitters including shrink fitters and press fitters.

7. **Coarse Shearing Tools**—microscale slitting, scissoring, piercing, punching, trimming, shaving, notching, and perforating; “eggbeater” tools for rapid pulping or liquefaction of semifluid biological materials (Section 10.4.2.5.3).

8. **Coarse Machining Tools**—microscale drilling, lathing, sawing, filing, planing, grinding and polishing.

9. **Pattern Impression and Coating Tools**—transfer molding, pantographic tracing; surface laminators, abstractors, decorators and other surfacing tools.

10. **Membrane Manipulation Tools**—cell and organelle lipid membrane bending, stretching, tearing, and joining; electric field-induced lipid bilayer membrane demixers; transmembrane channel inserters.

11. **Molecular Tools and Jigs**—tool tips designed to allow manipulation of individual molecules or molecular structures, perhaps using enzyme-like functionalizations for grasping (e.g., binding sites), cutting (e.g., pepsin, trypsin, collagenases, proteases, lipases, nucleases), joining (e.g., ribozymes, ligases), splicing (e.g., spliceosomes), folding (e.g., molecular chaperones), copying (e.g., DNA polymerase, reverse transcriptases), appending (e.g., transferases), and packing or pumping (e.g., hexagonal pRNA packs coiled DNA into viral capsids). The engineering of “wet” enzymes to accurately and reliably function when attached to stiff manipulators will be challenging because any attachment methods (multiple bonds) which will give the tip a high positional accuracy might also interfere with enzyme function. Enzymes generally require nanoscale movement to effect catalysis—the more tightly the tip is held, the less likely the enzyme will retain functionality.

Tool exchange ports and tool tip garages should lie within the accessible work volume of the nanomanipulator system, to allow convenient and watertight tool tip changeout, or else tool tips must be transferrable internally within the manipulator (e.g., Section 9.3.1.4).

### 9.3.3 Sensors and Manipulator Control

In addition to end-effectors, it will often be convenient to position various classes of sensors along the length of a manipulator, or at the tip. Monitoring control cable lengths, pneumatic chamber pressures, drive shaft rotations or strut lengths can allow a controller to know the position of a manipulator tip to the accuracy of one control increment, assuming no external forces and assuming regular recalibrations by periodically returning the arm to a precisely known “home” position. If low-stress stretch-inelastic metrological cables are embedded in the manipulator structure, then physical displacements of those cables in response to an applied load may be compared to no-load force/distance profiles, revealing the external force vector. Thus actual tip displacements are recorded directly by the metrological cables, and force feedback is continuously available. End effectors may include a wide variety of additional useful sensors, including gripper force and adhesion sensors, chemotactic sensors, vibration and acoustic sensors, electric or magnetic sensors, thermal or optical sensors, and so forth (Chapter 4).

Techniques for precise nanomanipulator control had received insufficient formal attention as of 1998. Pure open loop control (e.g., dead reckoning with no sensory feedback) could be sufficient in some molecular manufacturing operations. In this regard the inkjet printer might be a more suitable conceptual model than the robot arm, since a printer can build up considerable three-dimensional complexity during repeated passes over the workpiece following a simple linear chain of instructions in a raster pattern. Adding the ability to sense a fiducial structure near the workpiece can provide periodic checks on registration and reduce placement error, and Drexler has described the use of sterically complementary probes or alignment pegs to help guide end-effectors precisely to their targets.

However, force reflection and other sensory feedback are likely to play important roles in nanomedical manipulator control. J.S. Hall, cited in Drexler, points out that the ratios of the times and energies consumed by a typical nanomechanical computation to the times and energies consumed by a typical nanomanipulator motion are enormously greater than the corresponding ratios for microprocessor computation and macromechanical robotic arm motion. In other words, computation is relatively cheap for macroscale robotic manipulators while arm motion is relatively cheap for nanoscale robotic manipulators. Thus the moment-by-moment computer control of arm trajectories is the appropriate paradigm for macroscale robots, but not for nanoscale robots. For nanoscale robots, the appropriate manipulator control paradigm is often trajectory trial-and-error, also known as sensor-based motion control or “groping.”

Using trajectory trial-and-error, the nanomanipulator is moved generally in the direction of a desired endpoint, then is stopped and sensory feedback from the tip is obtained and examined for a match with the target destination profile. Given only a partial match, the arm is next moved in some direction and new sensory feedback is sampled. If the profile match has improved, another increment of motion is applied in the same direction; if the match has deteriorated, another direction is tried. Much like a blind man groping for his knife and fork by feeling around on the surface of a dinner table until he finds the utensils, the nanomanipulator is directed by trial-and-error progressively toward its goal, until the target is unambiguously located.

Efficient search algorithms for detecting and avoiding local maxima, identifying saddle points, rejecting outliers, and minimizing search times are well-known to mathematicians. In some cases, chemonavigational sensory readings may be taken using sensor pads (Section 4.2.5) at the manipulator tip to determine proximity to specific cellular processes or organelles. Unknown objects may be examined by haptic (tactile) exploration or by other means. Outside the cell, extracellular matrix (ECM) elements may be repeatedly chemotactically sampled to allow a nanomanipulator to feel its way toward a particular structure or node. Deployment of multiple nanomanipulators from a single device permits the simultaneous...
coordinated manipulation of several objects in the environment. In many cases, Brownian transport may suffice to bring a molecular target to the nanorobot arm.

If a nanomanipulator can be moved in minimum increments of $\Delta L$ at mean velocity $v_{arm}$ throughout a typical work volume $V_{work}$, then an efficient search algorithm can locate any target volume of size $\Delta L^3$ that lies within $V_{work}$ in at most $N_{trial} \approx \log_2(V_{work}/\Delta L^3)$ trials, requiring a search time $t_{search} \sim N_{trial} \Delta t$ where the mean time per nanomanipulator motion is $\Delta t \approx V_{work}/v_{arm}^{1/3}$ assuming a reasonably smooth gradient in roughly cubical work volume. Taking $\Delta L = 1$ nm, $v_{arm} = 1$ cm/sec, and $V_{work} = 0.1$ microns$^3$ for a cell-repair type manipulator, then $N_{trial} \approx 27$ trials and $\Delta t \approx 46$ microsec, giving $t_{search} \approx 1$ millisecond for the nanomanipulator arm to search for (and find) its intracellular target using the trial-and-error paradigm. This is consistent with the estimated time to search cell surfaces for specific markers (Eqn. 8.5). For the telescoping nanomanipulator to be used in molecular manufacturing as described by Drexler,$^{10}$ and taking $\Delta L = 0.25$ nm, $v_{arm} = 1$ cm/sec, and $V_{work} = 3$ nm$^3$, then $N_{trial} \approx 8$ trials and $\Delta t \approx 0.1$ microsec, giving $t_{search} \approx 1$ microsec.

Other control paradigms might also be useful in molecular manufacturing applications requiring high placement accuracy, possibly including full closed-loop operation in special cases.

### 9.3.4 Manipulator Arrays

Arrays of micron- or submicron-scale manipulator elements can be used to create 1-, 2-, or 3-dimensional programmable motion fields. In the simple 1-dimensional case, a single row of manipulators alternately grasp, transfer, and release a load, passing the payload down the line like a bucket brigade. In the 2-dimensional case, surfaces may be coated with closely-packed manipulators, allowing rapid and continuous transfers of multiple payload streams along multiple surface vectors simultaneously. In biology, ciliated epithelium is found in respiratory, excretory, digestive, circulatory, genital, and nervous systems of various animal groups, to drive fluid flow.$^{16,17}$

Examples in human physiology include the ciliated air passages of the bronchi (Section 8.2.2) and the villated surfaces of the small intestine (Section 8.2.3). In these natural systems, payload transport speeds are typically 10-100 microns/sec. For instance, experiments with ciliated bullfrog gullet tissue show maximum transport efficiency when moving a load of ~5 grams/cm$^2$ at a speed of 27 microns/sec.$^{526}$

Ciliary arrays used by protozoa for locomotion also display time-synchronized wave patterns$^{558}$ (Section 9.4.2.5.1), which may be analogous to the operation of submicron motion fields.

There are many possible applications of manipulator arrays in nanomedicine. Ciliary arrays inside hollow nanofactory workspaces can regulate throughput of physical materials for manufacturing (Chapter 19). Arrays can establish specified flows of environment fluids such as blood plasma across sensors, or may be used in device locomotion (both crawling and swimming; Section 9.4) or in self-cleaning activities. Medical nanosystems may incorporate complex three-dimensional manipulator arrays to control the transport of biocomponents in cell mills and organ mills (Chapter 21), to move chyme across an artificial stomach lining (Chapter 26), or to move fluids through artificial organs designed for chemical processing or filtration such as an artificial liver or kidney (Chapter 26). Ciliary transport systems may also support unique and highly sophisticated artificial whole-body nutrient transport systems such as the vasculoid-class devices described in Chapter 30. These nanomechanical arrays can be operated at transport speeds (1-100 cm/sec), flexure frequencies (up to 1 MHz), and power levels (0.1-100 pW per array element) characteristic of individual nanoscale manipulators (Section 9.3.1).

In 1998, the prototyping and testing of ciliary arrays in MEMS work was an active research area. Ataka et al$^{1274}$ fabricated and operated a 1 cm$^2$ array of 512 cantilever-type bimorph manipulators (each one 500 microns long, 100 microns wide, and 6 microns thick) and operated it at 10 Hz in a coordinated fashion to transport a 2.4 mg piece of silicon wafer at 27-500 micron/sec using ~100 micron strokes and a 4 milliwatt input power to each actuator. Bohringer et al$^{1639}$ fabricated a prototype array of 1024 cilia called the “cilia chip” that moves silicon chips at speeds up to 200 microns/sec with a placement accuracy of ~3 microns and a lifting capacity of 250 N/m². Individual ciliary pixel control allows movement of multiple components on different trajectories, enabling the performance of various assembly tasks. A larger manipulator chip (M-chip) consisting of ~10,000 micro-actuator “resonators” (5-micron tall cilia) has also been fabricated on a few cm$^2$ of silicon substrate and tested at 5 KHz.$^{1645}$ A part dropped onto the array is propelled forward at a constant speed of ~800 microns/sec.$^{1646}$ Bohringer et al$^{1643,1644}$ also created a theory of programmable force fields to model the actions of a ciliary array; simulations have uncovered numerous important operational design issues.$^{1646}$ Analogously in experimental biology, an array of kinesin motors fastened to a fixed surface transported microtubules hand-over-hand when ATP was added.$^{2425,2426}$

Peter Will et al$^{1640}$ developed a software toolkit to allow researchers to program Intelligent Motion Surfaces,$^{1641,1642}$ to perform a variety of simple functions including centering parts, spiraling a part to center in a convergent field array, spiraling a part away from the center in a divergent field array, rotating a part about the center using four triangular fields, orienting a part in a programmable field array, and sorting two parts using a hole as a trap for a part being translated past it. Properly arrayed, such surfaces can form an assembly pipeline in which parts may enter at one end and be successively sorted, spaced (so parts flow through at a constant rate), centered, aligned and, ultimately, inserted into other parts.

### 9.3.5 Materials Disassembly for Disposal

From time to time, medical nanorobots will find it necessary to accept bulk materials of various sizes and types, and reduce those materials to smaller and simpler forms in preparation for analysis, transport, or disposal from the human body. For example, nanorobots may be required to mechanically excise unwanted tissue or mineral deposits from arterial walls. Disabled nanodevices or their component parts found freely floating in tissues may also require disposal. Both requirements are discussed below. A fuller treatment of the biocompatibility of in vivo refuse, and routes of final disposal thereof, is deferred to Chapters 15 and 16.

#### 9.3.5.1 Morcellation and Mincing

Morcellation is the fragmentation of biological materials as they are excised from the body; mincing is the progressive reduction in size of tissue segments by mechanical means. Chunks of biological material may be excised from the body using a tool such as the core sampler described in Section 9.3.2. Cutting is required because many connective tissues have strengths on the order of 10-100 atm and the body is only pressurized to ~1 atm, so vacuum suction alone is insufficient to pull coherent tissues apart.

In the simplest design, morcellated material is passed through successive banks of fixed diamondoid cutting blades or rotating diamondoid cutting disks positioned at various angles with progressively more closely-spaced blade edges. The target material is minced into nanometer-sized bits. Each blade is geared for maximum force application at relatively slow speed and turns in a tightly fitted sleeve, providing an automatic self-cleaning action. More complex
designs might include paddlewheeling-bucket or orthogonal-blade arrays whose interiors are cleared by pistons after completion of the mincing action on material compacted into the cutting chamber by a larger piston. Table 9.3 shows that diamond and other potential blade materials are far stronger than all forms of biological matter, including dental enamel which is the hardest natural substance found in the human body.

To quantify the mincing action, consider a blade of length \( L_{\text{blade}} \), cutting edge width \( W_{\text{edge}} \), and compressive strength \( T_{\text{blade}} \) which applies a force \( F_{\text{blade}} \) to a biological sample of length \( L_{\text{sample}} \) and thickness \( x_{\text{sample}} \) and tearing strength \( T_{\text{sample}} \). Assuming \( L_{\text{blade}} = L_{\text{sample}} \), then the minimum force required to cut the sample is:

\[
F_{\text{min}} = T_{\text{blade}} W_{\text{edge}} \leq F_{\text{blade}} \quad \text{(2) (Eqn. 9.54)}
\]

The energy required to make all cuts down to the final pieces is:

\[
E_{\text{minced}} = \left( \frac{T_{\text{sample}} W_{\text{edge}} V_{\text{sample}}^{2/3}}{(2^{1/3} - 1)} \right) \quad \text{(Eqn. 9.54)}
\]

For disposal thus must include provision of capabilities to allow the partial or complete dismantlement of disabled or uncooperative devices.

Finally, there is the problem of nonbiological waste disposal. Such wastes might include accidental releases into the in vivo environment (external to the nanorobot) of raw feedstock materials, specialized fuel elements, pressurized nanocontainers, or nanodevice components such as tool tips, sensors, integument plates, or hard-material work products in various stages of completion. All manner of nanorobot-related detritus might plausibly be found freely floating in the tissues or bloodstream, from the smallest diamondoid parts up to and including detached whole manipulator arms, major subassemblies following a major device breach, or debris resulting from the execution of a voluntary autodestruct protocol (Chapter 12). While reliable and clean operation is an essential design objective in nanomedical systems, it would be irresponsible to ignore the need for an explicit nonbiological materials waste scavenging and disposal capability (Chapter 16).

9.3.5.3 Diamondoid Decomposition

While most hard materials can be comminuted in approximately the same manner as biological materials (Section 9.3.5.1), the...
onboard disassembly of diamondoid materials is clearly the greatest challenge. The six general approaches proposed in this Section are randomizing processes that destroy any atomic-level positional or compositional information contained within the structure that is being dismantled. Controlled information-preserving disassembly processes are described in Chapter 19.

### 9.3.5.3.1 Grinding

Two substances of nearly equal hardness can scratch each other, so a set of corrugated sacrificial diamond grinding elements can be used to mechanically grind or saw the target mass into nanometer-scale particulate waste. This process consumes a volume of grinder mechanism which is approximately equal to the volume of the target, hence is relatively energy inefficient because, in addition to the power dissipated by grinding, the sacrificial grinders must then be rebuilt which requires another expenditure of a similar magnitude of energy for fabrication. At least twice the volume of the original target in particulates is produced by this process. Alternatively, two or more diamond targets may be rubbed against each other destructively, most efficiently after the manner of a macroscale ball mill. Note that diamond detritus may be ground, though less efficiently, by sacrificial grinders comprised of materials of lesser hardness or strength—for example, cubic boron nitride or borazon, first synthesized in 1957, scratches diamond with ease.

### Table 9.3. Mechanical Strength of Nanomedical Materials

<table>
<thead>
<tr>
<th>Material</th>
<th>References</th>
<th>Young’s Modulus E (N/m²)</th>
<th>Failure Strength* (N/m²)</th>
<th>Material</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human Body Materials:</td>
<td></td>
<td></td>
<td></td>
<td>Other Inorganic Materials:</td>
<td></td>
</tr>
<tr>
<td>Saphenous vein</td>
<td>362</td>
<td>1.6-4.8 × 10¹</td>
<td>~2 × 10⁵</td>
<td>Water-ice (268-271 K)</td>
<td>1608-1609</td>
</tr>
<tr>
<td>Carotid artery</td>
<td>362</td>
<td>&lt;5.9 × 10⁴</td>
<td>~2 × 10⁵</td>
<td>Lead</td>
<td>460,1287</td>
</tr>
<tr>
<td>Elastin</td>
<td>362,364</td>
<td>6 × 10⁴</td>
<td>~2 × 10⁵</td>
<td>Tin</td>
<td>460</td>
</tr>
<tr>
<td>Muscle tissue</td>
<td>362</td>
<td>--</td>
<td>1-6 × 10³</td>
<td>Concrete (unreinforced)</td>
<td>364</td>
</tr>
<tr>
<td>Elastic tissue</td>
<td>521</td>
<td>--</td>
<td>3 × 10³</td>
<td>Concrete with PMMA</td>
<td>364</td>
</tr>
<tr>
<td>Thoracic aorta</td>
<td>521</td>
<td>0.1-1.7 × 10⁴</td>
<td>0.1-1.7 × 10⁴</td>
<td>(reinforcing fibers)</td>
<td></td>
</tr>
<tr>
<td>Erhythroyte membrane</td>
<td>362,1325,1415</td>
<td>~0.001-1 × 10⁴</td>
<td>1-20 × 10⁴</td>
<td>Gold</td>
<td>460</td>
</tr>
<tr>
<td>Axenomal microtubules</td>
<td>1449</td>
<td>~5 × 10⁴</td>
<td>~1 × 10⁶</td>
<td>Silver</td>
<td>460</td>
</tr>
<tr>
<td>Actin microfilaments</td>
<td>362,3210</td>
<td>~10⁻¹⁰</td>
<td>~6 × 10⁴</td>
<td>Glass</td>
<td>1164,1280,1286</td>
</tr>
<tr>
<td>Soft tissues (est.)</td>
<td>--</td>
<td>--</td>
<td>~0.0-1.0 × 10⁷</td>
<td>Glass</td>
<td>1164,1280,1286</td>
</tr>
<tr>
<td>Cartilage</td>
<td>364,521</td>
<td>0.1-1.6 × 10⁷</td>
<td>0.2-25 × 10⁵</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdominal skin</td>
<td>521</td>
<td>0.6-2.2 × 10⁷</td>
<td>0.4-4.4 × 10⁷</td>
<td>Copper</td>
<td>460,1164,1280,1288</td>
</tr>
<tr>
<td>Cancellous bone</td>
<td>364,585</td>
<td>3-3.3 × 10⁴</td>
<td>0.4-5.0 × 10⁷</td>
<td>Flint</td>
<td>364</td>
</tr>
<tr>
<td>Arterial plaque</td>
<td>--</td>
<td>--</td>
<td>0.1-1 × 10⁷</td>
<td>Cast iron</td>
<td>460,585,1280,1288</td>
</tr>
<tr>
<td>Tendon</td>
<td>362,364,521</td>
<td>1-9 × 10⁸</td>
<td>0.1-2.1 × 10⁸</td>
<td>(&amp; wrought iron)</td>
<td></td>
</tr>
<tr>
<td>Collagen</td>
<td>362,364</td>
<td>1-2 × 10⁴</td>
<td>0.2-1.4 × 10⁹</td>
<td>Mild steel (std. structural steel)</td>
<td>362,460,536,585,1164</td>
</tr>
<tr>
<td>Cornea</td>
<td>364</td>
<td>--</td>
<td>1 × 10⁶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intermediate filaments</td>
<td>364</td>
<td>4 × 10⁹</td>
<td>2.5 × 10⁸</td>
<td>Aluminum</td>
<td>362,364,460,1164,1288</td>
</tr>
<tr>
<td>Keratin (w &amp; wool)</td>
<td>364,3209</td>
<td>4 × 10¹⁰</td>
<td>2.5 × 10⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wet compact bone</td>
<td>362,364,585</td>
<td>0.5-2.7 × 10¹⁰</td>
<td>0.4-3.0 × 10⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dental enamel</td>
<td>364</td>
<td>7 × 10¹⁰</td>
<td>8 × 10⁷</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apatite</td>
<td>364</td>
<td>1.3 × 10¹¹</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone hydroxyapatite</td>
<td>364</td>
<td>1.4 × 10¹¹</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone fluorapatite</td>
<td>362</td>
<td>1.7 × 10¹¹</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Organic Materials:</td>
<td></td>
<td></td>
<td></td>
<td>Silicon</td>
<td>10,460,763,1281</td>
</tr>
<tr>
<td>Resilin</td>
<td>362</td>
<td>1.8 × 10⁸</td>
<td>3 × 10⁷</td>
<td>Boron</td>
<td>1281</td>
</tr>
<tr>
<td>Rubber</td>
<td>362,1280,1283</td>
<td>1.4-2.5 × 10⁸</td>
<td>1 × 10⁷</td>
<td>Fused quartz</td>
<td>460,1282</td>
</tr>
<tr>
<td>Abductin</td>
<td>362,364</td>
<td>1-4 × 10⁸</td>
<td>--</td>
<td>Silicon nitride</td>
<td>763,1281</td>
</tr>
<tr>
<td>Peptidoglycan, wet***</td>
<td>3149</td>
<td>1 × 10⁷</td>
<td>3 × 10⁷</td>
<td>Aluminum oxide</td>
<td>460,1283</td>
</tr>
<tr>
<td>Mature rat skin</td>
<td>521</td>
<td>4.6 × 10⁸</td>
<td>1.34 × 10¹⁰</td>
<td>Saphire</td>
<td>536,1602,1605</td>
</tr>
<tr>
<td>Celluloid</td>
<td>1280</td>
<td>2.5 × 10⁵</td>
<td>--</td>
<td>Graphite whiskers</td>
<td>10,763,1281</td>
</tr>
<tr>
<td>Cuticle (locust)</td>
<td>364</td>
<td>9 × 10⁴</td>
<td>9.5 × 10⁵</td>
<td>Tungsten carbide</td>
<td>460,536</td>
</tr>
<tr>
<td>Wood</td>
<td>362,364,585,1280</td>
<td>0.5-1.6 × 10¹⁰</td>
<td>0.02-2.2 × 10⁸</td>
<td>Boron carbide</td>
<td>460</td>
</tr>
<tr>
<td>Chitin</td>
<td>364</td>
<td>7 × 10¹⁰</td>
<td>0.2-7 × 10⁸</td>
<td>Silicon carbide</td>
<td>460,1283</td>
</tr>
<tr>
<td>Calcite shells</td>
<td>364</td>
<td>--</td>
<td>0.5-1.0 × 10⁸</td>
<td>Cubic boron nitride</td>
<td>460</td>
</tr>
<tr>
<td>Nautilus shell</td>
<td>364</td>
<td>4.4 × 10¹⁰</td>
<td>0.6-1.4 × 10⁹</td>
<td>Cubic carbon nitride</td>
<td>1284,1285</td>
</tr>
<tr>
<td>Peptidoglycan, dry***</td>
<td>3149</td>
<td>2 × 10¹¹</td>
<td>3 × 10¹⁰</td>
<td>Silicon carbide</td>
<td>460,1283</td>
</tr>
<tr>
<td>Cellophane</td>
<td>--</td>
<td>--</td>
<td>0.6-3.8 × 10⁸</td>
<td>Graphite sheet (in plane)</td>
<td>1281,2280</td>
</tr>
<tr>
<td>Cellulose fibers</td>
<td>364,3209</td>
<td>0.01-1.1 × 10¹¹</td>
<td>2.0-9.2 × 10⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silk, Bombyx mori</td>
<td>364,3209</td>
<td>0.5-1.6 × 10¹⁰</td>
<td>3.5-7.5 × 10⁸</td>
<td>Diamond</td>
<td>460,536,1321,1284</td>
</tr>
<tr>
<td>Silk, spider</td>
<td>362,364,1283</td>
<td>0.06-1.0 × 10¹⁰</td>
<td>3-15 × 10⁸</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nylon</td>
<td>364</td>
<td>4.5 x10¹⁰</td>
<td>7.5 × 10⁸</td>
<td>Fullerenes**</td>
<td>10,1278,1279,1308</td>
</tr>
</tbody>
</table>

* Failure strength may differ significantly in tension, compression, bending, or torsion; value given is representative.

** Mechanical properties are a function of nanotube chirality. 

*** Bacillus subtilis, bacterial cell wall.
9.3.5.3.2 Cleavage

Jewellers and diamond cutters have known for centuries that diamond has a “perfect cleavage” or “grain” in four directions, parallel to its octahedral crystal faces. After cutting a notch or “kerf” along the grain, a rather dull iron or steel edge is laid athwart the groove and a sharp ~20 microsec blow is struck, forcing apart the groove walls and causing the diamond sample to split along the cleavage plane of structural weakness. Large diamond crystals may be progressively fractured down to near-nanometer scale by this technique. A blade velocity of 500 m/sec should suffice for complete crack penetration through the crystal; a single cleavage of a 1 micron³ diamond cube requires 10.6 pJ; a striker blade 1 micron long and 10 nm in width gives higher fracture velocities which may produce rougher surfaces and even multiple fragmentation. Maximum crack propagation velocity is 1580 m/sec in glass, 4500 m/sec in sapphire, and 7200 m/sec in diamond. A large impedance mismatch between the gripper mechanism and the diamond crystal target optimizes pulse reflection at the holding surfaces and maximizes fragmentation. A gripper mechanism made of nylon (failure strength ~10¹¹ N/m²; Table 9.3) has an acoustic impedance of 2.9 x 10⁶ kg/m²-sec,⁵³⁶ compared to 6.3 x 10¹⁰ kg/m²-sec for diamond,⁵³⁶ a huge 22:1 impedance mismatch. A water bath also ensures acoustic reflection because of a large impedance mismatch with diamond (~1.5 x 10⁶ kg/m²-sec for water).

The weakest diamond cleavage plane with the lowest areal bond density is the [111] crystal plane which has a cleavage energy of 10.6 J/m².⁵³⁶ Thus a single cleavage of a 1 micron³ diamond cube requires 10.6 pJ; a striker blade 1 micron long and 10 nm in width must apply a pressure of ~10¹⁰ N/m² (~10,000 atm) along the contact edge, well below the failure strength of the hardest steel (Table 9.3). As a crude estimate of total cleavage decomposition energy, the progressive fracture of a 1 micron³ diamond cube into one million (10 nm³) cubes implies a thousandfold increase in surface area and the cleavage of 5.94 x 10¹⁰ nm² of diamond plane, requiring a minimum energy expenditure of ~6300 pJ to perform the decomposition. The cleavage technique is most applicable to brittle solids and may be less useful in the fragmentation of composite diamondoid materials (Chapter 11).

9.3.5.3.3 Sonication

Multiple acoustic pulses applied simultaneously to the surfaces of a crystal may elevate the interior mechanical stress above the failure strength of the material, causing it to fracture. Consider six diamond hammers of facial area L² and failure strength ~1 x 10¹¹ N/m² that simultaneously impact each of the six faces of a cubic diamond target of edge L, applying a non-self-fracturing peak pressure of 0.3 x 10¹¹ N/m² at each face. The peak pressure rises to 1.8 x 10¹¹ N/m² at the center of the cube where the plane waves converge, well exceeding the strength of diamond and shattering the target cube throughout its interior. Sonic pulverization is an analogous macroscale manufacturing process.

Single-walled fullerene nanotubes are cut by water-bath ultrasonication at 55 KHz,¹⁵²⁵ which is inferred to produce microscopic domains of high temperature¹⁵²³ following the collapse of cavitation bubbles, leading to localized sonochemistry that opens a hole in the tube side. Very high sonications of multi-walled fullerene nanotubes produces similar damage in a CH₂Cl₂ bath.¹⁵⁹⁰

9.3.5.3.4 Thermal Decomposition

Since carbon burns readily in oxygen, simple incineration in a high-pressure oxygen-rich atmosphere seems an obvious approach to diamond decomposition. A sapphire inner wall coating provides a flameproof combustion chamber with a ~200,000 atm failure strength (Table 9.3), and an externally insulated electrodynamic vacuum suspension (Section 6.3.4.4) prevents rapid thermal conduction and allows maintenance of high operating temperatures. The combustion temperature of diamond in air at 1 atm is usually given as 870-1070 K,⁶⁹¹ and higher oxygen pressures should lower this range considerably. Evans¹²⁷⁷ has studied the noncombustive oxidation of diamond in pure oxygen atmospheres as a function of temperature and pressure. Using data from Evans’ oxidation rate/pressure and Arrhenius oxidation/temperature plots, the etch rate (in meters/sec) for the [111] diamond crystal face (which oxidizes most readily) may be crudely approximated as:

\[ v_{etch} = k_1 p_{oxy} e^{-k_2/T} \quad \text{for} \quad T \leq 1050 \, K \]  
{Eqn. 9.55}

\[ v_{etch} = k_3 p_{oxy} \quad \text{for} \quad 1050 \, K \leq T \leq 1600 \, K \]  
{Eqn. 9.56}

where \( p_{oxy} \) is oxygen pressure (atm), \( T \) is diamond temperature (K), \( k_1 = 3.6 \times 10^9 \) m/sec-atm, \( k_2 = 6.5 \times 10^8 \) K, \( k_3 = 3 \times 10^{15} \) m/sec-atm, and the time required to completely oxidize a diamond cube of edge L is \( t_{decomp} = L / v_{etch} \). For \( p_{oxy} = 100 \) atm, a 1 nm diamond cube is consumed in 1 second at T ~ 750 K; a 1 nm cube heated to T ~ 660 K decomposes to oxide in 1 second, or in ~1 hour at 530 K. Although not modeled using high-pressure data, Eqn. 9.55 hints that a pure oxygen atmosphere at \( p_{oxy} ~ 1000 \) atm might support active combustion at an ignition temperature as low as ~700 K. Oxygen ions at 100-1000 eV can machine the diamond [100] face at a rate of ~0.1 nm/sec / (ion/sec-nm²) that saturates at a gas pressure >10⁻⁷ atm.²⁷⁰⁸

Extremely high temperatures are required to initiate and to sustain pure graphitization—the transformation of hard diamond into relatively soft (and more easily disposable) graphite by simple heating. In vacuo, unstressed polished hydrogenated diamond surfaces remain chemically and mechanically stable up to ~1275 K, at which temperature the passivating hydrogens are removed and crystallographic surface reconstruction begins.¹²⁹¹ Drexler¹⁰ estimates the characteristic thermal cleavage time for unstressed 556-zJ C-C bonds in diamond as ~10⁸⁵ sec at 300 K, ~10¹² sec at 700 K, and ~10¹⁶ sec at 1000 K. At 1300 K, a shear stress ≥0.18 mN/bond initiates plastic flow,¹²⁹² but diamond remains chemically stable up to ~1800 K in an inert atmosphere. Heating above 1800 K can result in extensive graphitization,¹²⁹⁰ an autocatalytic process that spreads outward from nucleation centers. The experimental graphitization rate of the [111] diamond crystal face at zero pressure is approximated by Evans¹²⁷⁷ as:

\[ v_{graphite} = k_4 e^{-E_a/kT} \quad (\text{m/sec}) \]  
{Eqn. 9.57}

where the activation energy \( E_a = 1760 \) zJ/atom, \( k = 0.01381 \) zJ/K (Boltzmann constant), \( T \) = temperature (K), and \( k_4 = 5.4 \times 10^{16} \) m/sec. At 1800 K, graphitization is very slow, about 1 nm/day; \( v_{graphite} \sim 1 \) nm/sec at T = 2150 K, ~1 micron/sec at T = 2440 K, well above the 2070 K softening point and the 2310 K melting point of sapphire.¹⁶⁰² Catalytic graphitization by Ni and Fe has been observed at ~1070 K.¹⁵⁹⁶ But because of the high temperatures involved, thermal decomposition of diamond by graphitization is impractical in vivo.
9.3.5.3.5 Molecular Mechanodecomposition

Molecular mechanosynthesis can be used to construct nanoscale objects, using either precise moiety-by-moiety placement or mechanical part-by-part placement of individual nanoscale parts (Chapters 2 and 19). Following the same process in reverse, diamondoid objects may be decomposed atom by atom into their constituent elemental components. This molecular mechanodecomposition may proceed rapidly using many tools working simultaneously, if no attempt is made to preserve or record the encoded structural pattern information.

Many of the radical-based reaction mechanisms for diamond mechanosynthesis described by Drexler may be reversible with the application of mechanical energy, since at elevated temperatures surface groups and radical vacancies apparently can migrate after placement on growing diamond surfaces. M. Krummenacker [personal communication, 1997] suggests a speculative mechanodecomposition process in which a few surface hydrogen atoms are first abstracted from an edge or corner of the diamond surface to minimize steric congestion at the disassembly site. Then a radical-based tool, possibly with an oxygen radical at the tip, bonds to an exposed carbon atom and pulls it upward far enough to insert another reactive moiety such as a nitrene (an analog to carbene) into the gap, replacing another C-C lattice bond and allowing the first carbon atom to be pulled further out of the matrix. This process is repeated on an adjoining carbon atom, resulting in a two-carbon unit with four bonds to the manipulator tool, which can then be pulled the rest of the way out by mechanically disrupting the two remaining C-C bonds to the diamond lattice without exceeding the tensile strength of the tool tip. (A great deal of sensing and chemical inference remains to be described in this scenario.) Krummenacker’s process might also be used to record structural information, if required (Chapter 19).

Evans estimates that 1760 zJ/atom of energy are required to break three C-C bonds and detach a single carbon atom from the (111) diamond crystal surface into the gas phase. With 176 carbon atoms/nm² in diamond, mechanodecomposition of bulk diamond would require ~300 nJ/micron³. Generously assuming only ~5 manipulator arm motions to remove each carbon atom and a manipulator operating frequency of ~1 MHz, a single manipulator can mechanically disassemble diamond at a volumetric rate of ~1000 nm³/sec-manipulator. The volume of each mechanosynthetic manipulator described in Section 9.3.1.4 is ~10⁻⁸ nm³, so -1000 manipulators disassembling in parallel occupy ~10⁻⁶ of a micron³ work space and can collectively mechanically decompose ~0.001 micron³/sec of diamond crystal while drawing a total of ~300 pW of continuous power. Actual power usage may be far less, since in any reasonable decomposition process some energy will be recovered when the liberated carbon atom forms new bonds; high energies will only be encountered in the transition state, as in any chemical reaction.

9.3.5.3.6 Chemical and Microbial Decomposition

Four issues are considered briefly here. First, can diamond or sapphire be chemically solvated? Second, does any microbe likely exist which is capable of attacking a diamond or sapphire surface? Third, could such a microbe evolve naturally? Fourth, could such a microbe be artificially engineered? (The author thanks R. Bradbury, M. Krummenacker, R. Merkle, and J. Soreff for helpful discussions and important contributions to this Section.)

1. Can diamond or sapphire be chemically solvated? Although carbon is soluble in molten Fe (e.g., >1808 K at 1 atm), Co, Mn, Ni, and Cr, there is no known room temperature solvent that dissolves pure crystalline diamond. Intact diamond and fullerenic surfaces are extremely inert. For example, after facet-cutting, gem diamonds are boiled in concentrated sulfuric acid for cleaning, leaving the gem surface unaffected. The outer faces of natural hydrophobic diamond may be terminated partly by hydrogen and partly by bridging oxygen, with a significant proportion of carbonyl groups and a small number of -OH and carboxyl (-COOH) groups as well.

Diamond is almost completely resistant to attack by room-temperature ground-state molecular oxygen (see Eqn. 9.55), although oxidative erosion by atomic oxygen has been shown at rates of ~0.04 nm/min and ozone or various radicals might also be effective. Molecular fluoride only fluorinates the surface, producing a “Teflon” coat (stable up to ~1120 K) while not disturbing the underlying C-C bonds because F makes only single bonds; chlorine is also taken up by diamond surfaces. Molten sodium nitrate attacks diamond at ~700 K. At high temperature or pressure, carbon from diamond can migrate and form a metal carbide phase in carbide-forming metals including W, Ta, Ti, and Zr, and metal oxides of Cu, Fe, Co, and Ni are reduced to the metal (a redox reaction with the carbon escaping as oxide) upon heating in vacuo.

Non-intact diamondoid surfaces may be more susceptible to chemical attack. For instance, open-ended single-walled fulleren nanotubes are consumed by a 3:1 mixture of sulfuric (98%) and nitric (70%) acids at 343 K at a rate of ~130 nm/hour even in the absence of sonication, the same mixture intercalates and exfoliates graphite. A 4:1 mixture of sulfuric acid (98%) and hydrogen peroxide (30%) at 343 K also etches the exposed ends of open nanotubes at ~200 nm/hour, “much like the burning of a fuse.” Etch rate may vary with the chiral indices (n, m) of the nanotubes. COCl-derivatized single-walled carbon nanotubes will solvate in organic solvents and high-strain sites such as the outer fold of a kinked nanotube are subject to attack by nitric acid.

Sapphire is primarily corundum or α-Al₂O₃, the oxide ions forming a hexagonally close-packed array and the Al ions distributed symmetrically among the octahedral interstices. The solubility of α-Al₂O₃ in neutral water at room temperature is given variously as 10⁻⁷ to 10⁻⁹ M (~60–6000 atoms/micron³) at equilibrium, but solubility rises sharply below pH 4 and above pH 9 in an almost U-shaped curve. (Human blood pH normally ranges from 7.35-7.45.) Aluminum oxide is amphoteric, forming hydrated Al⁺⁺⁺ ions in acidic solutions (pH < 4) and Al(OH)₄⁻ ions in alkaline solutions (pH > 9). Corundum dissolves slowly in boiling nitric acid and in orthophosphoric acid to 570 K, and dissolves well in potassium bisulfate at 670–870 K, or in borax or sodium fluoaluminate (cryolite) at 1070-1270 K.

Air-exposed Al metal is protected by a thin (~5 nm) adherent oxide layer having a defect rock-salt structure, easily compromised by amalgamation or halogens, or by alkali hydroxides (e.g., NaOH) with the generation of hydroxide in seconds at room temperature. Complex-forming reagents such as Cu⁺⁺ + Cl⁻ also react with oxide-coated aluminum metal in less than 1 minute.

II. Does any existing microbe possess the natural ability to attack a diamond or sapphire surface? Extensive tests have not yet been performed, and negatives cannot be proven, but the prospects are poor. Most organisms develop the means for attacking materials in their environment because these materials are abundant and provide essential molecules or elements that are needed for energy production or important biochemical functions.
For diamondophagy to evolve naturally, a microbe probably must occupy a niche in which diamond is more abundant than competing carbon sources. This is unlikely except many kilometers below the natural oil/coal deposits in the Earth's crust, though the possibility cannot be definitely ruled out because deep-crustal-dwelling bacteria and other archaic biota have been found in rocks at depths at least 2.7 kilometers, a field of study now known as geomicrobiology. However, far more abundant nonpolymeric carbon sources are usually available. Highly polymeric molecules including long-chain hydrocarbons, cellulose, and starches tend to be fairly resistant to enzymatic degradation—nature has found it relatively difficult to devise an attack strategy for cellulose (e.g., wood) and chitin (e.g., crab and insect shells). With their crystalline arrays, diamond and sapphire would seem to fall into this category as well.

A microbe naturally evolved to attack sapphire would probably be seeking to extract the aluminum (sapphirophagy), since oxygen is plentiful elsewhere. However, Al is generally considered toxic, and reports of microbes that feed on Al-rich minerals are extremely infrequent and remain unconfirmed, though microbial arduous. Acids such as HF, HCl and H2SO4 will not harm organic vessels (e.g., wax or plastic). As noted earlier, sapphire is toxic, and reports of microbes that feed on Al-rich minerals be seeking to extract the aluminum (sapphirophagy), since oxygen as well.

Arrays, diamond and sapphire would seem to fall into this category. While most of these microbes (>99%) are harmless naturally? Could a microbe capable of attacking diamond or sapphire evolve naturally? There are >10^13 bacteria living inside Earth's human population alone. While most of these microbes (>99%) are harmless or beneficial, up to ~10^14 may be undesirable pathogens possibly subject to attack by medical nanorobots. Could this large "natural laboratory" population of bacteria evolve diamondophagy or sapphirophagy as a defense against artificial nanomachines?

In the case of sapphire, most biological metal absorption is initiated by secretion of acids strong enough to dissolve the metal-containing material. Evolving an acid production capacity isn't difficult because most bacteria use H+ ion gradients as power sources. Common bacteria normally create a mildly acidic environment. Typical energy sources (e.g., glucose) are oxidized to increase the external H+ ion concentration. The flow of the H+ ions back into the cell through the ATPase enzyme generates ATP (as in mitochondria). Blocking or minimizing ATP production and limiting H+ diffusion could produce fairly high local H+ concentrations. Bacteria such as Thiobacillus ferrooxidans, found in environments where better energy sources are lacking, oxidize metal sulfides to generate energy. These result in the generation of sulfuric acid and environmental pH values of 2-3. Extremozymes that can work at a pH below 1.0 have been isolated from the cell wall and underlying cell membrane of some acidophilic bacteria, some acids (e.g., HF) that dissolve oxides (e.g., SiO2) may be safely stored in purely organic vessels (e.g., wax or plastic). As noted earlier, sapphire is also attacked by strong alkali (but see below).

Natural evolution of a diamondophagic capability is even more arduous. Acids such as HF, HCl and H2SO4 will not harm either H-terminated diamondoid surfaces or cellular lipid membranes which generally also have hydrogenated surfaces. While high H+ concentrations should remove oxygens from SiO2 or, with more difficulty, from Al2O3, a more efficient attack strategy for H-terminated diamondoid might be an extremely alkaline environment, although the most alkaliphilic bacteria (such as the "Natronobacterium gregoryi" normally found in soda lakes and high-carbonate soils) only reach a pH of 10-11. High OH- environments can destroy lipid bilayers, RNA, and standard proteases and lipases. Also, the biochemistries of acidophiles and alkaliphiles are so tuned to their natural environments that the predominantly neutral pH found in the human body might prove toxic for these microbes.

Most human-pathogenic bacteria likely to be attacked by medical nanodevices are not extremophiles. Consequently, as R. Bradbury observes, Nature's design challenge is for non-extremophiles to evolve systems of creating very high local H+ concentrations, having the energy resources available to drive this pathway and to minimize ATP production which utilizes the H+ ions. Such evolution seems unlikely. Only if medical nanodevices were constructed of an essential material that is in short supply, such as iron, or could provide a better energy source than the surrounding plasma or cytosol, could bacteria be selected which have the potential of attacking nanomachinery. Whether extremophiles might subsequently repopulate the vacated microecological niches currently occupied by natural pathogenic bacterial species is an open issue worthy of future study.

IV. Could a microbe capable of attacking diamond or sapphire be artificially engineered? Almost certainly the answer is yes. In the simplest case using conventional biotechnology, a bacterium would be designed to produce and secrete an appropriate acid or alkali, creating a highly corrosive environment local to the target material by binding to the surface and secreting the acid (or ions) into sequestered adherent pockets.

J. Soreff suggests other approaches that may employ locally far-from-equilibrium states which are harder to get (at least with fast kinetics) in normal chemistry. One strategy is to engineer artificial enzymelike structures that can bind to the target surface at one end while catalyzing the local production of molecules in excited states (e.g., by crude analogy to luciferase in bioluminescence) at the other end. Enzymes like catalase and superoxide dismutase already handle locally powerful oxidants, and in the presence of light and oxygen, C60 can pass its superfluous excitation energy onto nearby O2 molecules, creating singlet oxygen. An enzyme might be designed to catalyze HOCI → H2O2 + H + O2 (singlet state oxygen) or other singlets, triplets, quartets, or other electronically-excited oxidizers. However, survival of such an enzyme over multiple cycles is problematical. Another strategy is to design novel mecanochemic tools into the bacterium. Proteins can convert ATP hydrolysis into mechanical motion. Such motion, perhaps combined with a "snap-action" elastic energy storage mechanism to allow tool release speeds comparable to vibrational times, could be used to mechanically hammer an oxidant into a diamond surface, thus reducing the chemical energy required to be released at the point of action.

9.4 In Vivo Locomotion

One of the most important basic capabilities a medical nanorobot may possess is the ability to move about inside the human body. At its most simple, this movement may be purely statistical, with nanodevices carried along with the natural ebb and flow of bodily fluids. At the other extreme, nanorobot locomotion may be highly deterministic, including powered drive mechanisms, mapping and active navigation, and traverses of diverse histological territories.
having markedly different mechanical and chemical characteristics. The subject matter is huge and quite impossible to cover fully in a single Chapter. As a result, the discussion here is merely a preliminary survey of the most important issues and challenges of in vivo locomotion, suggesting promising new areas for future research.

Section 9.4.1 opens with an overview of fluid viscosity generally and the rheology (flow characteristics) of nanorobot-rich biofluids that might be associated with passive nanorobot locomotion. Aspects and techniques of active swimming through the bloodstream, or sanguination, are described in Section 9.4.2. This is followed by discussions of cytoabulation (cell surface walking and anchoring) in Section 9.4.3, histonation (tissue diving including diapedesis, ECM transit, and intercellular passage) in Section 9.4.4, cytopenetration (entering individual cells) in Section 9.4.5, locomotion inside the cell (Section 9.4.6), and finally cytocarriage (nanorobotic piloting of natural motile cells) in Section 9.4.7. The biocompatibility of motive mechanisms is discussed in Chapter 15.

9.4.1 Rheology of Nanorobot-Rich Biofluids

The study of biofluid flow, or biorheology, is useful in nanomedicine because it is necessary to understand the flow characteristics of fluids through which nanorobots must navigate—whether a nanorobot is actively swimming or is simply drifting with the flow, or whether the nanorobot is traveling in an ordinary Newtonian fluid such as water or in more complex viscoelastic biofluids such as mucus or saliva. Additionally, the presence of large numbers of nanorobots may dramatically alter bloodstream viscosity (depending on nanorobot number density and shape), with important medical consequences. The following discussion examines biofluid and whole-blood viscosities (Sections 9.4.1.1 and 9.4.1.2), the radial distribution of blood elements in blood vessels (Section 9.4.1.3), viscosity and bloodstream velocity profiles of nanorobot-rich blood (Sections 9.4.1.4 and 9.4.1.5), and hematocrit reduction in narrow blood vessels (Section 9.4.1.6).

9.4.1.1 Biofluid Viscosity

Viscosity is a measure of the resistance of a fluid to shearing when the fluid is in motion. Consider a plate of surface area A moving parallel to a fixed plane surface with constant velocity \( v \), being pushed laterally by a constant force \( F \). A fluid of viscosity \( \eta \) fills the volume between the two surfaces, which are separated by a distance \( d \). The fluid layer nearest the moving plate also moves at velocity \( v \) and the layer nearest the stationary plane remains stationary. In between the two surfaces, the velocity increases linearly with distance, establishing a constant gradient. This gradient is usually called the shear rate (essentially a size-normalized velocity), or \( \dot{\gamma} = v/d \) (m/sec-m, or sec\(^{-1}\)). The absolute viscosity may then be defined by:

\[
F/A = \eta \dot{\gamma} \quad \text{(N/m}^2\text{)}
\]

where \( \eta \) has MKS units of N-sec/m\(^2\), Pascal-sec, or, more simply, kg/m-sec. In a Newtonian fluid, the shear stress \( F/A \) (N/m\(^2\)) increases linearly with shear rate \( \dot{\gamma} \), so that viscosity \( \eta \) is constant over a wide range of shear rates. Many common fluids such as air, water, saline and blood serum closely approximate the ideal Newtonian fluid. The viscosity of solutions of molecules is related to the diffusion coefficient; see Eqn. 3.5.

The viscosities of some common materials are given in Table 9.4. The viscosity of an ideal gas is independent of density and independent of pressure between 0.01-10 atm; at higher pressures, intermolecular interactions lead to higher viscosities. Theory predicts that gas viscosity \( \sim T^{1/2} \) (e.g., rises with the square-root of temperature \( T \)), but a somewhat larger temperature exponent is obtained experimentally for real gases. In contrast, the viscosity of liquids increases with rising pressure (typically by a factor of 2-3 after moving from 1 atm up to 1000 atm for organic liquids) and decreases with rising temperature. In normal liquids the temperature dependence is approximated by Andrade’s formula which gives:

\[
\eta = k_v e^{E_v/kT} \quad \text{[Eqn. 9.59]}
\]

where for example the activation energy for viscosity \( E_v = 25 \text{ zJ} \) and the constant \( k_v = 2.1 \times 10^{6} \text{ kg/m-sec} \) for pure water at 1 atm. Non-electrolytes dissolved in water generally cause viscosity to rise, while solvation of electrolytes may increase or decrease viscosity (the effect is typically -10% or less for a 1M solution). Large asymmetric solute molecules increase viscosity more than an equal mass of small spherical molecules. The threshold between solid and liquid is generally taken as \( \eta = 10^{-14} \text{ kg/m-sec} \) [364].

Most biofluids are viscoelastic and non-Newtonian, with apparent viscosity \( \eta_a \) varying with shear rate and displaying other nonlinear characteristics such as hysteresis, relaxation, and creep. [362] Saliva behaves more like an elastic body than like water. Because of its high molecular weight, DNA solution is viscoelastic even at low concentrations. Sex glands produce viscoelastic fluids, including semen and uterine cervical mucus [190] (\( \eta_a \) varies ~20% during the menstrual cycle). [3575-3577] Human synovial fluid becomes increasingly incompressible at higher pressures, with \( \eta_a \sim \text{10 kg/m-sec at } 0.1 \text{ sec}^{-1} \text{ (e.g., knee flexion during very slow walking) declining to } \eta_a \sim 0.1 \text{ kg/m-sec at } 10 \text{ sec}^{-1} \text{ (very fast running), and to } \eta_a \sim 0.001 \text{ kg/m-sec at } 10,000 \text{ sec}^{-1} \text{ (experimental).} \) [362] Viscoelasticity is also an important property of respiratory tract mucus, which typically shows an \( \eta_a \sim 1 \text{ kg/m-sec at low shear rates near } 0.1-1 \text{ sec}^{-1} \), falling to \( \eta_a \sim 0.01 \text{ kg/m-sec at high shear rates near } 100-1000 \text{ sec}^{-1}. \) [362]

Even ice is a viscoelastic material, with \( \eta_a \sim 10^{10}-10^{13} \text{ kg/m-sec} \) (estimated as maximum shear stress divided by shear rate from data in Sinha et al. [1609] for \( \eta_a \sim 10^{-3} \text{-10}^{-4} \text{ sec}^{-1} \) at 262 K; ice viscosity varies with temperature (Table 9.4) as described by Andrade’s formula, with \( E_v \sim 110 \text{ zJ} \) as determined experimentally for pure ice [1069] and taking \( k_v = 6.3 \times 10^{3} \text{ kg/m-sec} \) at high shear rate.

Cytoplasm is a complex viscoelastic material having a continuous liquid phase (the cytosol) plus various suspended particles, granules, and membranous structures. More precisely, the cytomatrix is a mixed-phase body composed of a fibrillar network penetrated by a solution. [1408] As a result, viscosity is different in the various phases. From flow behavior, the viscosity of E. coli protoplasm was estimated as \( \text{1000 kg/m-sec; from measured diffusion rates of sucrose, dextran, and } \beta\text{-galactosidase, the apparent viscosity was } 3-4 \times 10^{3} \text{ kg/m-sec.} \) [1407] Cytoplasm is inhomogeneous and anisotropic at many levels of organization.

9.4.1.2 Viscosity of Whole Blood

Human blood is a suspension of cells, in plasma (an aqueous solution of electrolytes and nonelectrolytes; Appendix B). The plasma is ~90% water by weight, 7% plasma proteins, 1% inorganic materials, and 1% other organic substances. The cellular component (Appendix B) is essentially all erythrocytes, or red blood cells (RBCs; Figure 8.43), with white cells of various categories comprising less than 1/600th of the total volume of cells and platelets less than 1/800th.

Pure blood plasma is a Newtonian viscous fluid, with \( \eta_{\text{plasma}} \sim 1.1 \times 10^{3} \text{ kg/m-sec at } 310 \text{ K. Blood plasma is the environment
Table 9.4. Absolute Viscosity of Some Common Materials

<table>
<thead>
<tr>
<th>Material, (Gas, Liquid, Solid)</th>
<th>Viscosity (kg/m-sec)</th>
<th>Material, (Gas, Liquid, Solid)</th>
<th>Viscosity (kg/m-sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water @ 0°C (273 K)</td>
<td>1.787 x 10^-3</td>
<td>Earth’s mantle (geological)</td>
<td>~10^-21</td>
</tr>
<tr>
<td>@ 20°C (293 K)</td>
<td>1.002 x 10^-3</td>
<td>Pitch (room temperature)</td>
<td>~10^-14</td>
</tr>
<tr>
<td>@ 37°C (310 K)</td>
<td>0.6915 x 10^-3</td>
<td>Glass @ 720-920 K (anneal)</td>
<td>2.5 x 10^-12</td>
</tr>
<tr>
<td>@ 50°C (323 K)</td>
<td>0.5468 x 10^-3</td>
<td>Glass @ ~1300 K (blowing)</td>
<td>~1 x 10^-6</td>
</tr>
<tr>
<td>@ 100°C (373 K)</td>
<td>0.2818 x 10^-3</td>
<td>Glass @ 1500-1700 K (furnace)</td>
<td>~1 x 10^-6</td>
</tr>
<tr>
<td>Ice @ -11°C (262 K)</td>
<td>~10^-10^-11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>@ -42°C (191 K)</td>
<td>~10^-11 est.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>@ -109°C (164 K)</td>
<td>~10^-11 est.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>@ -129°C (144 K)</td>
<td>~10^-12 est.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squid axoplasm</td>
<td>1-10 x 10^4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pulmonary macrophage cells</td>
<td>0.12-0.27 x 10^6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chick fibroblast cells</td>
<td>~0.01 x 10^1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human feces, range (est.)</td>
<td>~10^-7^-10^-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. coli cytoplasm</td>
<td>&gt;1,000,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mucus and sputum (310 K)</td>
<td>10-750,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythrocyte membrane (310 K)</td>
<td>~100,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plasma membrane (~310 K)</td>
<td>10,000-100,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T-cell cytoplasm (298 K)</td>
<td>27,300 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukocyte, adhered cell (310 K)</td>
<td>66,800 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human fibroblast cytoplasm</td>
<td>5000-8000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leukocyte, free cell (310 K)</td>
<td>4500 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba cytoplasm</td>
<td>1000-30,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human chyme, normal (est.)</td>
<td>100-30,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human semen</td>
<td>9.4 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saliva (310 K)</td>
<td>4-10,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Synovial (knee) fluid (298 K)</td>
<td>1-10,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole blood, low shear rate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hct = 45%</td>
<td>~100 x 10^-1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hct = 90%</td>
<td>~1000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole blood, high shear rate:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hct = 45%</td>
<td>~10 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hct = 90%</td>
<td>~100 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RBC contents (Hb solution)</td>
<td>6-13 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human blood plasma (310 K)</td>
<td>1.1-1.2 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum/Interstitial fluid (310 K)</td>
<td>1.0-1.1 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bile fluid (probably liver)</td>
<td>0.84-2.3 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human tears (310 K)</td>
<td>0.73-0.97 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golden syrup (285 K)</td>
<td>140,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycerin (273 K)</td>
<td>12,110 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corn syrup (294 K)</td>
<td>~5,000 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycerin (303 K)</td>
<td>629 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Castor oil (310 K)</td>
<td>297 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machine oil, heavy (310 K)</td>
<td>3.5 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oleic acid (303 K)</td>
<td>25.6 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfuric acid, liquid (310 K)</td>
<td>25.4 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethylene glycol (310 K)</td>
<td>10.7 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lead, liquid (623 K)</td>
<td>2.58 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iron + 2.5% C, liquid (1673 K)</td>
<td>2.25 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propanol (310 K)</td>
<td>1.5 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanol (310 K)</td>
<td>0.472 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetone (310 K)</td>
<td>0.285 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diethylether (310 K)</td>
<td>0.202 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air, liquid (81 K)</td>
<td>0.172 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbon dioxide, liquid (303 K)</td>
<td>0.053 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neon, gas (310 K)</td>
<td>0.032 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen, gas (310 K)</td>
<td>0.021 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Air, gas (510 K)</td>
<td>0.01896 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrogen, gas (310 K)</td>
<td>0.018 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbon dioxide, gas (310 K)</td>
<td>0.0157 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrogen, liquid (~20 K)</td>
<td>0.011 x 10^-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hydrogen, gas (310 K)</td>
<td>0.0091 x 10^-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Normally encountered by micron-sized bloodstream-traversing nanorobots as they move between blood cells.

On the other hand, “whole” blood is the complete natural mixture of blood plasma and blood cells. In blood vessels whose diameters are much larger than the size of blood cells, or in the case of mechanical systems with dimensions much larger than the characteristic sizes of blood cells, whole blood must be treated as an homogeneous non-Newtonian fluid. The bulk viscosity of whole blood decreases with rising shear rate and increases with rising hematocrit (Hct), the percentage of blood volume occupied by red cells. (The human male hematocrit has a normal range of Hct = 40%-52%, average Hct = 46%,743 which optimizes oxygen transport.1313 As an example, for whole blood at a low shear rate of g = 0.1 sec^-1, then \( \eta_a = 0.001 \text{kg/m-sec at Hct} = 0\% \), 0.1 kg/m-sec at Hct = 45%, and ~1 kg/m-sec at Hct = 90%. At a high shear rate of g = 100 sec^-1, then \( \eta_a = 0.001 \text{kg/m-sec at Hct} = 0\% \), 0.01 kg/m-sec at Hct = 45%, and ~0.1 kg/m-sec at Hct = 90%.362 At still higher shear rates g > 100 sec^-1, whole blood behaves almost like a Newtonian fluid with a nearly constant \( \eta_a \) (Fig. 9.12). Viscosity also varies with disease state (if any) and slightly with temperature. For whole blood at normal Hct and g = 0.1 sec^-1, then \( \eta_a = 0.10 \text{kg/m-sec at Hct} = 45\% \text{ and Hct} = 90\% \) at 310 K and ~0.15 kg/m-sec at 283 K; for g = 100 sec^-1, then \( \eta_a = 0.010 \text{kg/m-sec at Hct} = 310 \text{ K and -0.015 kg/m-sec at 283 K}.362 

For large blood vessels, the effect of hematocrit on viscosity for Hct ≤ 45% has been modeled experimentally by Cokelet et al.318 as:

\[
\frac{\eta_a}{\eta_{\text{plasma}}} = (1 - Hct)^{-2.5} \tag{Eqn. 9.60}
\]

with Hct expressed as a fraction. This formula, which also gives good results for RBC-sized oil/water emulsions in the same volume fraction range,1316 is only useful up to volume fraction of ~10% in the case of rigid-particle suspensions, whose viscosity behaves markedly differently from that of RBC suspensions (see below).

It has long been known1313 that human red blood cells can form aggregates known as rouleaux, in which the discoid RBCs adhere loosely in a “stack of coins” configuration at shear rates <100 sec^-1.1314 Formation of linear and branched chain aggregates known as rouleaux, in which the discoid RBCs adhere loosely in a “stack of coins” configuration at shear rates <100 sec^-1.1314
(e.g., the slower the blood velocity), the more prevalent (larger size and number density) are the aggregates. As the shear rate goes to zero, it is speculated that human blood becomes one big aggregate, which then may behave as a viscoelastic or viscoplastic solid.362

As shear rate increases, the rouleaux tend to break up. Individual red cells also deform slightly, elongating and lining up with the streamlines. These two effects combine to reduce blood viscosity with rising shear, as illustrated by Figure 9.12 which shows the relative viscosity \( \frac{\eta_a}{\eta_{\text{plasma}}} \) for human blood at Hct = 45% as a function of shear rate. Note that for rigid particles (including hardened red cells, normal leukocytes, and nondeforming nanorobots), bulk viscosity is essentially independent of shear rate.

Normal blood vessel wall (peripheral zone) shear rates in physiological bloodflow range from 50-700 sec\(^{-1}\) in the larger arteries to 250-2000 sec\(^{-1}\) in the smallest arteries and capillaries, and 20-200 sec\(^{-1}\) in large and small veins.386 At shear rates >100 sec\(^{-1}\) where RBC aggregation ceases to be important, pure whole blood remains fluid even up to 98% RBC concentration by volume (e.g., Hct = 98% volume fraction).1319

9.4.1.3 Radial Distribution of Blood Elements

At low shear rates, red cells aggregate into rouleaux and migrate inward, forming a network of aggregates in the core of the tube. Individual rouleaux may incorporate 10-20 red cells, or more, creating by far the largest cellular elements normally present in the blood. At the highest shear rates, the rouleaux break up entirely into single red cells, and the red cells then distribute themselves more uniformly in the radial direction.

The radial distribution of white cells is also a function of flow conditions.1325,1333-1336 At low shear rates, under conditions allowing red cell aggregation, the white cells are displaced to the periphery of the flow by the much larger red cell rouleaux. At the highest shear rates, white cell concentration is highest along the tube axis, displacing some red cells, since most white cells are larger than individual red cells.

The radial distribution of the local platelet concentration during blood flow in tubes has also been investigated.1337,1338 In plasma containing only platelets, the platelet distribution is radially uniform. However, in whole blood flow under all shear conditions, the platelet concentration is highest near the vessel wall. In arterioles, the platelet number density is about two times higher near the wall than in the center of the vessel.1342 Platelets are much smaller than either red cells or white cells, thus tend to be crowded out of the center whenever red cells or white cells are present.

These experimental observations are consistent with the general principle that during blood flow in a vessel, the largest particles, or “flow units,” move toward the axial region, leaving the smallest particles more concentrated at the periphery.1352 In most cases, bloodborne medical nanorobots will be ~2 microns in diameter or smaller. As such, they will normally constitute the smallest particles in the bloodstream. More than any other blood element, free-floating nanorobots should tend to migrate nearest the blood vessel walls, although at high shear rates the radial diffusivity is significantly increased due to local fluid motions generated by red cell rotations (Section 3.2.2). Small molecules and complexes such as lipoproteins are less subject to erythrocyte-induced diffusivity enhancement, hence may migrate toward the walls.

9.4.1.4 Viscosity of Nanorobot-Rich Blood

The presence of large numbers of relatively rigid nanorobots dramatically alters bloodstream viscosity. Figure 9.13 shows the relative viscosity \( \frac{\eta_a}{\eta_{\text{plasma}}} \) for human blood at 298 K with shear rate >100 sec\(^{-1}\) as a function of particle volume fraction, compared to the relative viscosity of suspensions of latex rigid spheres, rigid disks, emulsion droplets, and sickled erythrocytes (which are virtually nondeformable), as determined experimentally.1312 Figure 9.13 reveals that a 50% suspension of micron-sized rigid nanorobots will increase blood viscosity by a factor of ~350, seriously impeding flow especially in the smaller vessels. However, a plasma suspension of microspheres at a ~10% particle volume fraction has a relative viscosity indistinguishable from Hct = 10% whole blood. This suggests a conservative 10% volume-fraction limit for the maximum bloodstream concentration of medical nanorobots (e.g., a maximum “nanocrit” or Nct = 10%), a limit that may also ensure free flow of the fluid (see also Sections 9.4.1.5 and 9.4.2.6).

Fig. 9.12. Viscosity of human blood as a function of shear rate, at Hct = 45% (relative apparent viscosity at 310 K, redrawn from Chien1314).

Fig. 9.13. Viscosity of nanorobot-rich human blood at high shear rate (relative apparent viscosity at 298 K, shear rate > 100 sec\(^{-1}\); redrawn from Goldsmith and Mason,1312 Cokelet and Lichtman1402).
Relative viscosity also depends on particle size, though the effect due to the presence of medical nanorobots is usually minor. As a conservative upper limit in the smallest vessels:1315

\[ \frac{\eta_a}{\eta_{\text{plasma}}} = \left[ 1 - \left( \frac{D_{\text{nano}}}{d_{\text{tube}}} \right)^4 \right]^{-1} \]  

\text{(Eqn. 9.61)}

where \( D_{\text{nano}} = 2 \times r_{\text{nano}} \) is the maximum nanorobot diameter (radius) and \( d_{\text{tube}} = 2 \times r_{\text{tube}} \) is the blood vessel diameter (radius). Taking \( d_{\text{tube}} = 8 \) microns (capillaries), then \( \eta_a / \eta_{\text{plasma}} = 1.0002 \) for \( D_{\text{nano}} = 1 \) micron, or 1.07 for \( D_{\text{nano}} = 4 \) microns (the largest bloodborne nanorobot; see Sections 5.2.1 and 8.2.1.2). In larger blood vessels, this effect is even smaller for micron-sized medical nanorobots.

Relative blood viscosity also depends on nanorobot shape. Chien’s measurements\(^{314} \) of effective viscosity as a function of particle shape in dilute suspensions found that minimum viscosity is achieved by hard spheres or by 1:1 hard cylinders. Thin disks or long cylinders have higher viscosity. For example, 10:1 rods (10 times longer than wide) produce a suspension with -10 times higher viscosity than a suspension containing an equal volume of spheres; for 100:1 rods, viscosity increases -2500-fold. The implication for nanodevice design is that a large population of bloodborne medical nanorobots will have the minimum impact on blood viscosity if each nanorobot is closest to spherical in shape. Long rod or flat disk shapes will greatly increase blood viscosity, in comparison to spheres, although at the lowest nanodevice number densities the total impact on blood viscosity may be negligible.

Chien’s results\(^{314} \) also suggest that metamorphic nanorobots capable of continuous surface deformations in response to flow conditions (like RBCs) may further reduce their contribution to blood viscosity by at least a factor of 2-6, depending on shear rate (Fig. 9.12). Goldsmith and Turitto\(^{386} \) show that at shear rates over 200 sec\(^{-1} \), typical in physiological blood, the optimum shape for red cells is ellipsoidal, positioned at an angle to the flow, with the surface rotating in the direction of flow in a tank-tread-like motion. In experiments with flowing emulsions, the deformation of a liquid droplet results in its migration across the streamlines away from the tube wall. Thus in physiological blood over the whole range of normal hematocrits and typical flow rates, there is a plasma-rich (blood-cell-rare) or “plasmatic” zone \( \delta_{\text{plasma}} \approx 2-4 \) microns deep at the walls of vessels whose diameters exceed 100 microns.\(^{362,1319} \)

Such lateral migration is not observed with small rigid particles of any shape at high concentrations and at low Reynolds numbers (Section 9.4.2.1) \( N_R \approx 10^{-3} \) (e.g., arterioles and smaller vessels; Table 8.2).\(^{319} \) However, for vessels with \( N_R \approx 1 \) (e.g., arteries and veins; Table 8.2), inertial effects do come into play and rigid free-floating nanorobots will be pushed away from the wall to produce a particle-free zone. The thickness of this “plasmatic” zone \( \delta_{\text{plasma}} \) decreases sharply with increasing nanorobot concentration (Nct). For example, at Nct = 2%, \( \delta_{\text{plasma}} \approx 0.3 \times r_{\text{tube}} \); at Nct = 10%, \( \delta_{\text{plasma}} \approx 0.1 \times r_{\text{tube}} \); at Nct = 30%, \( \delta_{\text{plasma}} \approx 0.01 \times r_{\text{tube}} \).\(^{1320} \)

In terms of individual nanorobot motion, a rigid sphere initially placed near the tube wall migrates inward, while a rigid sphere placed near the tube axis migrates outward. Known as the “tubular pinch effect,” rigid spheres started in either position converge to an intermediate equilibrium radius position (as measured from the tube axis) of \( r_{\text{eq}} = (0.6-0.7) \times r_{\text{tube}} \) for \( R_{\text{nano}} / r_{\text{tube}} \ll 1 \), or \( r_{\text{eq}} = 0.5 \times r_{\text{tube}} \) (farther from the wall) for \( R_{\text{nano}} / r_{\text{tube}} \approx 0.25 \).\(^{1320,1321} \) By analogy with Brownian translational diffusion and Eqn. 3.1, a radial dispersion coefficient \( D_r \) may be defined as \( \Delta = 2 (\pi D_r) \) (metres), where \( \Delta \) is the RMS radial displacement of a bloodborne object in an observation time \( \tau \). The analogy is imperfect because these radial motions are not random, but are due to multibody collisions determined by the local velocity gradient, particle concentration, and surface deformations of the objects. At any given concentration, displacements are greatest at radial distances between 0.5-0.8 \( r_{\text{tube}} \). For local shear rates of 5-20 sec\(^{-1} \) and volume concentrations from 20%-70%, \( D_r = 1-20 \times 10^{-12} \) m\(^2\)/sec both for red cells and for rigid 2-micron diameter microspheres,\(^{386} \) and 3-86 x 10\(^{-11} \) m\(^2\)/sec for platelets in whole blood,\(^{1358} \) as determined experimentally.\(^{1358} \) Thus the mean time for a nanorobot (\( R_{\text{nano}} = 1 \) micron) to migrate a radial distance \( \Delta = 1 \) micron is \( \tau \approx 25-5000 \) milliseconds—about an order of magnitude faster than simple Brownian diffusion (Section 3.2.1).

### 9.4.1.5 Bloodstream Velocity Profiles

Consider a Newtonian fluid of viscosity \( \eta \) flowing through a cylindrical tube of length \( l_{\text{tube}} \) and radius \( r_{\text{tube}} \) with pressure differential \( \Delta p \) between the ends. The average fluid velocity (\( v_{\text{ave}} \)) for laminar or Poiseuille flow* is given by Eqn. 9.26, but imposing a no-slip condition at the vessel wall produces a radius-dependent parabolic velocity profile:62

\[ v_{\text{ave}}(r) = k_p (r_{\text{tube}}^2 - r^2) \text{ (m/sec)} \]  

\text{(Eqn. 9.62)}

where \( r \) is radial distance from the tube axis and \( k_p = \Delta p / 4 \eta l_{\text{tube}} = 2 v_{\text{flow}} / r_{\text{tube}} \). Maximum flow velocity \( v_{\text{max}} = 2 v_{\text{flow}} \), and occurs at the centerline (\( r = 0 \)).

Of course, as fluid enters a tube from a large reservoir, there is an entrance region called the inlet length (\( l_{\text{inlet}} \)) which lies between the entrance and a point downstream, where the parabolic profile is in the process of being established asymptotically (Fig. 9.14). For \( N_R \approx 1 \), \( l_{\text{inlet}} \approx 1.3 \times r_{\text{tube}} \) while for \( N_R \approx 30 \), \( l_{\text{inlet}} \approx 0.16 N_R r_{\text{tube}} \) producing, post-inlet, a \( <1\% \) deviation from an ideal Poiseuille (parabolic) profile.\(^{361,1331} \) Inlet conditions prevail throughout the entire length of the aorta and most of the major arteries, but entrance effects are minimal in the smaller vessels.

Now suppose that red blood cells are added to the plasma. Due to the inward migration of red cells and other factors, the velocity profile of whole blood is affected by flow rate and hematocrit,

---

* For well-developed turbulent flows, the velocity profile may be expressed as \( v_{\text{ave}} = v_{\text{ave}} (1 - (r / r_{\text{tube}}))^2/3 \) away from the laminar sublayer near the wall, where \( m \approx 7 \) for a wide range of Reynolds numbers.1380

---

Fig. 9.14. Establishment of parabolic velocity profile in Poiseuille tube flow (redrawn from Goldsmith and Turitto386).
particularly in vessels <500 microns in diameter, becoming blunted (Fig. 9.15). The degree of blunting decreases with increasing flow rate and increases with rising hematocrit. But the ability of RBCs to deform under the influences of cell crowding and fluid stresses in shear flow allows whole blood to continue to flow up to Hct ~ 98%. It is theoretically possible that metamorphic nanorobots could approach this level of performance, though this issue has not yet been studied extensively.

What if rigid spherical nanorobots are added to the plasma instead of red cells? At small R_{nano} and low Nct, the flow profile remains parabolic. Onset of velocity profile blunting generally occurs when Nct ≥ 20% or when R_{nano} > 0.05 \, r_{tube}. Such blunted flow is sometimes called partial plug flow. Partial plug flow was not observed experimentally when (Nct \, R_{nano}/r_{tube}) ≤ 0.6%. Thus, non-plug, purely Poiseuille flow can probably be maintained for rigid spheres and disks. As Nct and R_{nano}/r_{tube} continue to rise, complete plug flow eventually ensues (Fig. 9.16(B)), wherein the entire mass of fluid moves stiffly at constant velocity, like toothpaste squeezed from a tube. Plug flow requires much higher pumping power (and pumping pressure) than laminar flow. Complete plug flow was observed at Nct = 38% and R_{nano} = 0.112 \, r_{tube}, that is, when (Nct \, R_{nano}/r_{tube}) ≥ 3.8%. If complete plug flow can be avoided by holding R_{nano}/r_{tube} < 0.38 at our assumed maximum Nct = 10% (Section 9.4.2.6), then for the smallest d_{tube} = 4 micron human capillary, R_{nano} < 0.76 microns, allowing up to D_{nano} = 1.5 micron diameter for bloodborne nanodevices at the maximum nanocrit.

Figure 9.16 shows the experimentally-determined effects of Nct, R_{nano}/r_{tube}, and fluid flow rate on the velocity profile of suspensions of rigid spheres and disks. As Nct and R_{nano}/r_{tube} continue to rise, complete plug flow eventually ensues (Fig. 9.16(B)), wherein the entire mass of fluid moves stiffly at constant velocity, like toothpaste squeezed from a tube. Plug flow requires much higher pumping power (and pumping pressure) than laminar flow. Complete plug flow was observed at Nct = 38% and R_{nano} = 0.112 \, r_{tube}, that is, when (Nct \, R_{nano}/r_{tube}) ≥ 3.8%. If complete plug flow can be avoided by holding R_{nano}/r_{tube} < 0.38 at our assumed maximum Nct = 10% (Section 9.4.2.6), then for the smallest d_{tube} = 4 micron human capillary, R_{nano} < 0.76 microns, allowing up to D_{nano} = 1.5 micron diameter for bloodborne nanodevices at the maximum nanocrit.

Further analysis that might consider the effects (on bloodborne nanorobot velocity profiles) of vessel geometry (including bifurcations, nozzling, and curved paths), vessel wall elasticity (including collapsible tubes), pulsatile flow, and adding nanorobots of various mixtures of sizes and shapes to whole blood rather than plasma, would be useful but is beyond the scope of this book.

9.4.1.6 Hematocrit Reduction in Narrow Vessels

Fahraeus found that when blood of a constant hematocrit Hct is allowed to flow from a large feed reservoir into a small tube, hematocrit in the tube (Hct_{tube}) decreases as the tube diameter decreases; Fahraeus and Lindqvist found a decrease in apparent viscosity when tube diameter is reduced to below 300 microns (Fig. 9.17). Barbee and Cokelet crudely approximated the experimental data using the following (slightly modified) empirical formula:

\[
\text{Hct}_{\text{tube}} = \left( \exp \left( \frac{a}{\ln(d_{\text{tube}})} + b \right) \right) \text{Hct} \quad (\%)
\]

where \(a = 0.196\), \(b = -0.117\), and \(d_{\text{tube}}\) is expressed in microns. Given a normal human male hematocrit of Hct = 46%, in smaller vessels this falls to Hct_{tube} = 36% at d_{tube} = 100 microns, 25% at 30 microns, and 13% at 8 microns. Surveying the literature, Gaethgens concludes that minimum hematocrit Hct_{tube} generally occurs at d_{tube} ~ 15-20 microns, which Cokelet suggests marks the transition from multi-file flow to single-file flow among the red cells. Gaethgens also showed that the relative viscosity of human RBC suspensions reaches a minimum at about 7-8 microns.

Hematocrit decreases in small blood vessels for several reasons. First, a cell-free layer approximately equal to RBC radius exists near the wall, so the smaller the vessel, the larger the fraction of volume occupied by this layer, hence the lower the hematocrit. Second, a vessel side branch that draws mainly from the cell-free layer produces a lower hematocrit in that side branch (an effect called plasma skimming). Third, red cells are elongated and oriented along the direction of shear flow, making it less likely that they will enter a side branch aligned perpendicular to the direction of flow (an effect at d_{tube} ≲ 29 microns, sometimes called screening or steric hindrance). All three factors should be less important for near-spherical rigid nanorobots that are smaller than RBCs, hence any reduction in nanocrit during passage through narrow blood vessels should be quite modest.

The velocity of a cell or nanorobot located on the axis of a small vessel is greater than the mean velocity of the suspending fluid, but is always slightly less than the fluid in the immediate vicinity on the axis (Section 9.4.1.5). The simplest model is the stacked-coins model discussed by Whitmore, wherein the nanorobot velocity \(v_{nano}\) is given by:

\[
\frac{v_{nano}}{v_{flow}} = 2 \left[ 1 + \left( \frac{D_{nano}}{d_{tube}} \right)^2 \right]^{-1}
\]

Taking \(d_{tube} = 8\) microns and \(v_{flow} = 1\) mm/sec for a typical capillary (Table 8.2), a nanorobot with diameter \(D_{nano} = 2\) microns has \(v_{nano} = 1.88\) mm/sec, which is faster than \(v_{flow}\) but is slower than \(v_{max} = 2\) mm/sec.

9.4.2 Sanguinatation

Medical nanorobots will often be called upon to travel from place to place within the human body by actively swimming through the bloodstream, a process called sanguinatation. This Section describes the general nature of the process in terms of the Reynolds number (Section 9.4.2.1), rotations and collisions with vessel walls and cellular blood components likely to be experienced by free-floating or powered nanorobots (Section 9.4.2.2), nanorobot hydrodynamics (Section 9.4.2.3), general force and power requirements for submersive swimming (Section 9.4.2.4), various specific natrixation mechanisms (Section 9.4.2.5), and some additional considerations regarding sanguinatation (Section 9.4.2.6).
9.4.2.1 Reynolds Number

Consider an object of characteristic dimension $L$ moving at velocity $v$ through a fluid of density $\rho$ and viscosity $\eta$. The object's movement is resisted by two forces—inertia and viscous drag. The inertial force on the object is of order $F_{\text{inertial}} = \rho v^2 L^2$; the viscous drag force is of order $F_{\text{viscous}} = \eta v L$. Thus a slow human underwater swimmer with $L \sim 1$ meter, $v \sim 0.1$ m/sec, $\rho \sim 1000$ kg/m$^3$, and $\eta \sim 10^{-3}$ kg/m-sec must apply $F_{\text{inertial}} = 10$ N plus a minor additional $F_{\text{viscous}} = 10^{-4}$ N of motive force in order to keep moving forward. Clearly, the human swimmer lives in a world of predominantly inertial forces.

A bacterial swimmer faces entirely different challenges. A bacterium of size $L \sim 1$ micron and velocity $v \sim 10$ micron/sec must apply $F_{\text{inertial}} = 10^{-4}$ fN (femtonewtons; 1 fN = $10^{-15}$ N) but also a much larger $F_{\text{viscous}} = 10$ fN of motive force in order to keep moving forward. The ratio of the two forces is still $10^5:1$, but the roles have reversed. The bacterium (or any micron-scale medical nanorobot) lives in a world dominated by viscosity, where, as an example, the phenomenon of “coasting” essentially ceases to exist. For instance, if motive power to a swimming nanorobot with radius $R_{\text{nano}} = 1$ micron and velocity $v_{\text{nano}} = 1$ cm/sec is suddenly stopped, then the nanorobot will “coast” to a halt in a time $t_{\text{coast}} = \rho R_{\text{nano}}^2 / 15 \eta = 0.1$ microsec and in a distance $x_{\text{coast}} \sim v_{\text{nano}} t_{\text{coast}} = 1$ nm. If the nanorobot is rotating at a frequency $\nu_{\text{nano}} = 100$ Hz when its rotational power source is suddenly turned off, $v_{\text{nano}}$ decays exponentially to zero in a time $t_{\text{decay}} = \nu_{\text{nano}}^{-1} / \rho R_{\text{nano}}^2 / 15 \eta = 0.1$ microsec and stops after turning $\theta_{\text{rot}} = 2 \pi v_{\text{nano}} \rho R_{\text{nano}}^2 / 15 \eta \sim 40$ microradians.

The ratio of inertial to viscous forces is called the Reynolds number $N_R$, or:

$$N_R = \frac{F_{\text{inertial}}}{F_{\text{viscous}}} = \frac{\rho v L}{\eta}$$  \hspace{1cm} \text{[Eqn. 9.65]}

---

Fig. 9.16. Dimensionless velocity profiles in flowing aqueous nanorobot suspensions (rigid spheres and discs in rigid tubes; modified from Karnis, Goldsmith, and Mason\textsuperscript{1322}). A) Effect of concentration ($N_c = \%$ nanocrit). B) Effect of particle size ($R_{\text{nano}} / \ell_{\text{tube}}$). C) Effect of fluid flow rate ($\dot{V}_{\text{HP}}$). D) Effect of nanorobot shape.
which is a dimensionless number. In the examples given above, \( N_R = 10^4 \) for the human swimmer and \( N_R = 10^5 \) for the bacterium. Purcell\(^{389}\) notes that for a man to be swimming at the same Reynolds number as his own sperm, he would have to be placed in a swimming pool full of molasses and then be forbidden to move any part of his body faster than 1 cm/min, roughly the speed of the minute-hand of a large wall clock.

The Reynolds number has already been introduced in connection with laminar tube flow (Section 9.2.5), and elsewhere it has been noted that \( N_R = 100-6100 \) in the arteries, 200-900 in the veins, 0.0004-0.003 in the blood capillaries (Table 8.2), and \( 10^{-6}-1 \) for lymph vessels (Table 8.5). However, these figures are relevant only when considering flow phenomenon on a scale large enough such that the cellular graininess of human blood may be ignored. In the case of microscopic motile cells and medical nanorobots, this assumption is not valid.

On the contrary, nanorobotic sanguinatators will find themselves negotiating a viscous Newtonian-fluid plasmatic environment, punctuated by numerous closely-spaced free-floating cellular obstacles. Powered sanguinatation thus may involve traversing opportunistic clear volumes of blood plasma between red cells, then altering course to take advantage of the next available open space further along, following a zig-zag path generally in the desired direction. While the sizes and shapes of these clear volumes are strongly time- and position-dependent, their characteristic size is \( x_{\text{clear}} \sim (MCV \times 100\%-\text{Hct}) / \text{Hct} \).\(^{113}\) Taking the Mean Cell Volume \( MCV = 94 \) micron\(^3\) for red cells (Section 8.2.1.2), then \( x_{\text{clear}} \sim 5 \) microns in the arteries where Hct = 46\% and \( x_{\text{clear}} \sim 10 \) microns in the capillaries where Hct = Hct\text{tube} = 10\% (Section 9.4.1.6). If paths taken through clear volumes at a mean velocity \( v_{\text{nano}} \) average an angle \( \theta_{\text{path}} \) relative to the desired direction of travel, then the net forward velocity in the desired direction is \( v_{\text{net}} = v_{\text{nano}} \cos(\theta_{\text{path}}) \). Taking \( v_{\text{nano}} = 1 \) cm/sec giving \( N_R = 10^2 \), then \( v_{\text{net}} \sim 0.5 \) cm/sec through heavy traffic in the vessel lumen assuming \( \theta_{\text{path}} \sim 60^\circ \).

Brownian displacements of RBCs are negligible by comparison—\( v_{\text{brownian}} \sim \Delta X / \tau \sim 0.1 \) micron/sec for \( \Delta X = 1 \) micron (Eqn. 3.1)—while shear velocities in small arteries are typically \(-1 \) mm/sec (Section 9.4.2.2). For powered trajectories exclusively confined to the cell-free “plasmatic” zone near the blood vessel wall (the vascular “express lane” for the fastest-moving nanorobot traffic; Section 9.4.2.6), \( x_{\text{clear}} \) is largely independent of Hct, \( \theta_{\text{path}} \sim 0^\circ \), and \( v_{\text{net}} \sim v_{\text{nano}} \).

### 9.4.2.2 Rotations and Collisions in sanguo

From Eqn. 9.62 describing the axial fluid velocity \( v_{\text{fric}} \) as a function of radial distance \( r \) in laminar tube flow, the velocity gradient, or shear rate \( \gamma \), is seen to increase linearly with \( r \), as:

\[
\gamma = \frac{\Delta p}{2 \eta l_{\text{tube}}} = \frac{4 \pi \eta v_{\text{flow}}}{r_{\text{tube}}^2} \quad \text{r}^{-1} \quad \text{[Eqn. 9.66]}
\]

When a suspending liquid is subjected to laminar flow, the fluid stresses on the surface of suspended rigid bodies cause those bodies to rotate as they travel down the tube.\(^{1339}\) Rigid free-floating spherical nanorobots will rotate with uniform angular velocity, or, more specifically, with a rotational frequency of:

\[
v_{\text{nano}} = \frac{\gamma}{4\pi} \left( \frac{\Delta p}{8\pi \eta l_{\text{tube}}} \right) r = \left( \frac{\eta v_{\text{flow}}}{\eta^2 r_{\text{tube}}} \right) r \quad \text{Hz} \quad \text{[Eqn. 9.67]}
\]

For example, a rigid free-floating spherical nanorobot traveling in an artery with \( r_{\text{tube}} = 500 \) micron and \( v_{\text{flow}} = 100 \) mm/sec (Table 8.2) rotates at \( v_{\text{nano}} = 63 \) Hz at \( r = 495 \) microns (very near the vessel wall), which is much faster than the random \(-0.1 \) Hz Brownian rotation typical for micron-scale nanorobots (Section 3.2.1). Non-spherical rigid particles\(^{1343}\) and red cells\(^{1344}\) display variable angular velocity but spend more time in each orbit aligned with the direction of flow. Platelets can be seen tumbling in arterioles, especially near the vessel wall,\(^{1340}\) but they tend to align with the flow, a tendency which is strongest nearest the wall.\(^{1341}\)

Gyroscopic effects are present in every rotating body, with pitch, roll, yaw, and turning producing torques which cause precessions, giving periodic increases in bearing pressures and stresses. Fortunately, these stresses remain modest in most nanomechanical designs. For example, consider a rapidly spinning component inside the tumbling spherical nanorobot described in the previous paragraph. This internal component is a diamondoid disk of radius \( r_{\text{disk}} \), thickness \( h_{\text{disk}} \), and density \( \rho_{\text{disk}} \); spinning at an angular velocity \( \omega_{\text{disk}} \) around an axis oriented in the plane of the nanorobot tumble (tumbling is at frequency \( v_{\text{nano}} \), and supported by two coaxial bearings located a distance \( xbearing \) apart. The gyroscopic reaction force on the bearings is given by:

\[
F_{\text{bearing}} = \pi^2 \frac{r_{\text{disk}}^4 h_{\text{disk}} \rho_{\text{disk}} \omega_{\text{disk}} v_{\text{nano}}}{xbearing} \quad \text{[Eqn. 9.68]}
\]
Taking $r_{disk} = 100 \text{ nm}$, $h_{disk} = 20 \text{ nm}$, $\rho_{disk} = 3510 \text{ kg/m}^3$, $\omega_{disk} = \omega_{max} = 10^{10} \text{ rad/sec}$ (Eqn. 4.17), $v_{nano} = 63 \text{ Hz}$, and $v_{tumbling} = h_{tubing}$, then $F_{tumbling} = 2 \text{ pN}$.

In general, if two micron-scale objects are translating at the same mean speed, more energy is dissipated by the one that produces the larger amount of vorticity—that is, an asymmetric object that rotates as it swims is less efficient than the nonrotating swimmer.\(^\text{1389}\) The tangential force required to neutralize the hydrodynamic torque acting on a rigid sphere of radius $R_{nano}$ and tumble frequency $v_{nano}$ is torque divided by radius, or:

$$F_{tumble} = 16 \pi^2 \eta R_{nano}^2 v_{nano}$$

For $R_{nano} = 1 \text{ micron}$, $v_{nano} = 63 \text{ Hz}$, and $\eta = 1.1 \times 10^{-3} \text{ kg/m-sec}$ for plasma at 310 K, then $F_{tumble} = 11 \text{ pN}$. Application of such small forces may be controlled using measurements of absolute orientation taken from an onboard nanogyroscope (Section 4.3.4.1). Given that $F_{tumbling}$ and $F_{tumble}$ are of similar magnitude, direct anti-tumble gyrostabilization and gyroscopic rotational locomotion (gyrorotation) may be possible in some nanorobot applications.

The velocity gradient often brings suspended particles into close proximity. Kinetic theories of flowing suspensions have been developed from experiments on model systems involving two-body collisions between rigid and deformable spheres\(^\text{1345}\) and between rigid cylinders.\(^\text{1343}\) Goldsmith and Mason\(^\text{1312}\) estimate that the per-object collision frequency $K$ for equal-sized rigid spheres of volume concentration $c_i = (4/3)\pi R_i^3$ expressed as a fraction. Each object occupies a cubic fluid volume of dimension $L = n_i^{-1/3} = (4/3)\pi c_i^{1/3}$, and in laminar flow, two radially-adjacent boxes move past each other (a “collision”) at a mean relative velocity of $v_{ij} = v_i - v_j$ when $R_i, R_j < r_{hydro}$ and interact for a time $t_{ij} = L_i / v_{ij}$ (j = 1, 2). From the viewpoint of object $i$, the probability that after $N_{coll,ij}$ collisions at least one object $j$ has been encountered is $P_{ij} = 1 - (1 - c_i)N_{coll,ij}$ Taking $p_i < 0.9$ (the final result is relatively insensitive to the exact threshold selected), and multiplying by 2 because there is a concentric layer on either side of object $i$, then the mean collision rate experienced by each object $i$ with an object $j$ also present in the suspension is $K_{ij} = 2(N_{coll,ij} v_{tij})^{-1}$, which, at a radial distance $r$ in the tube, may be written as:

$$K_{ij} = 2 \left( \frac{48}{\pi} \left( \frac{v_{flow} c_i^{1/3}}{r_{tube}^{2/3}} \right) \left( \frac{R_i + R_j}{R_i} \right) r \log_{10} \left( \frac{1 - c_i}{1 - c_j} \right) \right)$$

and the mean free path ($l_{ij}$) between an object $i$ and an object $j$ in the suspension (neglecting all velocity profile and margination effects) is given approximately by:

$$l_{ij} = \left( \frac{4}{9} \right)^{1/3} \left( \frac{v_{flow} c_i^{1/3}}{r_{tube}^{2/3}} \right)^{1/3} \left( \frac{R_i + R_j}{R_i} \right)^{1/3} r \log_{10} \left( \frac{1 - c_i}{1 - c_j} \right)^{-1}$$

For example, in a suspension consisting of $R_i = 1 \text{ micron}$ nanorobots with volume concentration $c_i = 0.10$ (Nct = 10%), and $R_j = 1.5 \text{ micron}$ platelets with $c_j = 0.0035$ (mean physiological plateletocrit), flowing through a small artery with $r_{tube} = 500 \text{ microns}$ at mean velocity $v_{flow} = 100 \text{ mm/sec}$, and at a radial distance $r = 495 \text{ microns}$ (near the tube wall), then for nanorobot/platelet interactions, $K_{ij} = 2 \text{ collisions/sec}$, mean collision velocity $v_{ij} = 2 \text{ mm/sec}$, and $l_{ij} = 9.7 \text{ microns}$. For nanorobot/nanorobot interactions, $K_{ij} = 40 \text{ collisions/sec}$ (cf. $K = 200 \text{ collisions/sec}$ using Eqn. 9.70), $v_{ij} = 1.6 \text{ mm/sec}$, and $l_{ij} = 3.5 \text{ microns}$. These figures are crude estimates at best, since they ignore many possible complicating factors such as aggregation and margination, collision inelasticities and n-body collisions ($n > 2$), nonspherical object shapes, the presence or absence of specific macromolecules, receptor interactions, pulsatile flow, ionic strength of the medium, and cell surface electrostatics.\(^\text{1325,1354}\)

Free-floating nanorobots that collide with blood vessel walls produce negligible shear forces, given the no-slip condition at the wall. Powered nanorobots of radius $R_{nano}$ that impact a vessel wall at velocity $v_{nano}$ may apply a maximum shear stress of $p_{shear} = \rho_{nano} v_{nano}^2$; taking $\rho_{nano} = 1000 \text{ kg/m}^3$ and $v_{nano} = 1 \text{ cm/sec}$, then $p_{shear} \approx 0.1 \text{ N/m}^2$. By comparison, the time-averaged shear stress for blood circulation in normal vessels\(^\text{362,1346,1347,1352}\) is $1 - 2 \text{ N/m}^2$ (range 0.5-5.6 N/m²\(^\text{386}\), reaching up to 10-40 N/m² when small arteries and arterioles are partially occluded as by atherosclerosis or vascular spasm.\(^\text{1348,1349}\) This may also be compared to the threshold limit for shear stress-induced platelet aggregation of 6-9 N/m²\(^\text{362}\),1346,1349,1351 the shearing stress of 5-100 N/m² acting at the interface between a leukocyte and an endothelium when the leukocyte is adhering to or rolling on the endothelium of a venule,\(^\text{366}\) and the critical shear stress of 42 N/m² known to initiate major changes in the endothelial cells in the arteries.\(^\text{365}\)

9.4.2.3 Disturbed Flows, Hydrodynamic Interactions, and Entropic Packing

Two kinds of fluid flow have already been described—laminar or Poiseuille streamline flow, and turbulent or random flow. However, in branching or nonuniform-diameter blood vessels, nanorobots may also be required to negotiate an additional flow regime that is not observed in straight tubes of uniform diameter. This third regime, called “disturbed flow,” involves secondary fluid motions in directions away from that of the primary flow, often with separation of the streamlines from the vessel walls to form a vortex or a recirculation zone between the forward flowing mainstream and the wall.\(^\text{1358}\) Disturbed flow is most common in the larger blood vessels.

Studies have been conducted to observe the flow patterns in various vessels of simple and complex geometries, using flow visualization and cinemicrographic techniques.\(^\text{1358-1366}\) In one extensive series of experiments,\(^\text{1359}\) tracer polystyrene microspheres were photographed at various flow rates as they traveled through glass tube models or through chemically-transparentized natural blood vessels. The developed movie films were then projected on a drafting table and analyzed frame by frame. Figure 9.18 illustrates some results for progressively higher-angle blood vessel bifurcations. The cross-streamline and vortex patterns are quasi-stable; in pulsatile flow, vortices vary periodically in size and intensity, with the axial location of the vortex center and reattachment point oscillating in phase with the upstream fluid velocity between maximum and
minimum positions about a mean. A model study using red cells found that over time periods long in comparison with the orbital period, single cells and small aggregates <20 microns in diameter migrate outward across the closed streamlines and exit the vortex after describing a series of spiral orbits of continually increasing diameter until rejoining the mainstream. Erythrocyte rouleaux >30 microns in size remain trapped within the vortex, either assuming equilibrium orbits or staying at the center. The vortex also provides favorable conditions for the spontaneous aggregation of normal human platelets through shear-induced collisions of particles circulating in their orbits. An intimate knowledge of disturbed flow patterns in all human blood vessels will be essential in any nanomedical mission design requiring sanguinatating nanorobots, but further detailed discussion is beyond the scope of this book.

A variety of early studies attempted to model the hydrodynamics of micron-scale swimmers, including descriptions of the hydrodynamics of an individual swimming cell, the hydrodynamic interaction of two magnetotactic bacteria, and the interaction of two parallel swimming cells. Wall effects are well-known. For instance, the velocity of a flagellar swimmer decreases as it moves closer to a solid boundary; the boundary effect is strong at low N<sub>R</sub>, with speed reduced 5% even at 10 object radii from the wall. Flagellates swim in curved paths near a solid boundary. Another effort to model the hydrodynamic interaction of pairs of flagellar-driven 1-micron diameter bacterial swimmers found that cells are attracted toward each other when they are swimming side by side and are repelled when swimming one behind the other. The researchers also modeled two microswimmers approaching each other from opposite directions along parallel but non-coaxial paths. As the two cells pass, they move closer together, causing mutual cell rotations which lead to a significant change in swimming orientation. Eventually the cells move away from each other along a straight, now-coaxial path whose deflection angle to the original path is determined by the initial off-axis separation of the two cells.

Fig. 9.18. Disturbed flow streamlines in progressively higher-angle blood vessel bifurcations. A) Axisymmetric-constricted tube with no bifurcation (redrawn from Karino et al). B) 45° bifurcation (redrawn from Motomiya and Karino). C) 90° bifurcation (redrawn from Karino, Motomiya, and Goldsmith). D) 150° bifurcation (redrawn from Karino et al).

Fig. 9.19. Hydrodynamic interaction of approaching flagellates (redrawn from Dillon, Fauci, and Gaver).
also known as Stokes law. This result applies only in the case of a single sphere in an infinite expanse of homogeneous viscous fluid, and neglects all inertial terms. If the fluid container is finite in size, or if there are other spheres in the neighborhood, or if inertial forces cannot be entirely ignored because $N_R \gg 1$ (e.g., $R_{nano} = 1$ micron, $v_{nano} > 1$ m/sec in water), then the equation must be corrected. $^{374}$ Dynamic effects due to sphere oscillation, sudden release from rest, or variable speed also require corrections, $^{374}$ and there are other corrections for neighboring spheres. $^{373}$ Most of these corrections are of order near-unity for micron-scale objects.

Eqn. 9.73 provides a useful approximation of the motive force required to drive a spherical nanorobot through blood plasma. Assuming a perfectly efficient drive mechanism, the power requirement $^{372}$ is at least:

$$P_{nano} = F_{nano} v_{nano}$$

$$= 6 \pi \eta R_{nano} v_{nano}^2 \text{ (watts)} \quad \text{[Eqn. 9.74]}$$

For example, a force of $F_{nano} = 200$ pN and a power of at least $P_{nano} = 2$ pW are required to drive a $R_{nano} = 1$ micron spherical nanorobot at $v_{nano} = 1$ cm/sec through blood plasma at 310 K with $\eta = 1.1 \times 10^{-3}$ kg/m/sec. These formulas may also be used to estimate nanodevice velocity, given $F_{nano}$ or $P_{nano}$.

Of course, drive mechanisms are not perfectly efficient. Propulsion efficiency is often poor for objects that swim at low Reynolds numbers (e.g., typically $N_R < 10^{-3}$ for flagellates). $^{3578}$ A sphere driven by a helical propeller (e.g., flagellar propulsion; Section 9.4.2.5.2) of arbitrary length may have a propulsion efficiency as low as $e^6 - 0.01\%$. $^{389}$ Propulsive efficiency for organisms with spherical heads 10-40 times larger than their flagellar radius (optimum shape), using helical flagellar beats, ranges from $e^4 = 0.10-0.28$ (10%-28%), $^{3577}$ requiring a propulsive input power of $-P_{nano}/e^6 \approx 7-20$ pW in the previous example. Other drive mechanisms may prove more efficient.

Force requirements for nonspherical nanorobots are similar in magnitude. For example, consider a cylinder of radius $R_{nano}$ and length $L_{nano}$ translating uniformly in a direction normal to its axis. Lighthill $^{367,1376}$ shows that the force (applied at the midpoint) required to move this object at velocity $v_{nano}$ is given by:

$$F_{nano} = \frac{8 \pi \eta L_{nano} v_{nano}}{1 + \ln \left( \frac{L_{nano}^2}{R_{nano}^2} \right)} \quad \text{[Eqn. 9.75]}$$

for $L_{nano} >> R_{nano}$. $^{363}$ (If the perpendicular force is applied at a distance $c$ from one end of the cylinder, then the $L_{nano}^2$ term in Eqn. 9.75 is replaced by $4c$ ($L_{nano} - c$) when $c >> R_{nano}$.) Eqn. 9.75 does not hold for a perpendicular force applied near the ends of the cylinder. Also, a constant $F_{nano}$ actually produces a slowly varying velocity field along the cylinder length; for a midpoint-applied force, $v_{nano}$ is at maximum at the midpoint, but the percentage variation over most of the cylinder is not large. $^{363}$ Experiments show that needle-shaped bodies fall in viscous media about half as fast sideways as they do end-on; $^{1378}$ that is, the force on a long cylinder translating parallel to its axis is $F_{nano} \approx - (1/2) F_{nano}$.

The power requirement is then calculated as in Eqn. 9.74.

Interestingly, there seems to be a sharp minimum size limit of ~0.6 microns for free-swimming foraging microbes, below which size locomotion has no apparent benefit. $^{3581}$ This theoretical conclusion is supported by the observation that the smallest 97
genera of motile bacteria have a mean length of 0.8 microns, whereas 18 of 94 nonmotile genera are smaller.

9.4.2.5 Nanomechanisms for Natation
Swimming motions consisting purely of reciprocal deformations (e.g., shape A deforms into shape B, then retraces the same motion back to shape A) cannot produce forward progress in a viscosity-dominated environment. For example, in the macroscale world, a scallop may open its shell slowly, then close it rapidly to squirt out water, relying upon inertia and “coasting” to carry it forward a little bit each cycle. At the low Reynolds numbers of most medical nanorobots, such single-hinge deformations yield only back-and-forth motion. Indeed, Fukuda et al estimate that a vibrating-fin-driven water-swimming robot shorter than ~6 mm can no longer overcome viscous drag. Purcell notes that the simplest possible mechanical swimmer may require at least two hinges and a cyclical deformation loop traced out in a two-dimensional configuration space.

At least four distinct classes of mechanisms for natation are readily distinguished, as described below.

9.4.2.5.1 Surface Deformation
The use of flexible or metamorphic surfaces (Section 5.3) allows natative structures to deform asymmetrically. For example, it is impossible to “row a boat” that is fully submerged in a viscous fluid, because the stiff oars are simply reciprocating. But if the structures are flexible, then the oar can bend one way during the first half of the stroke and the other way during the second half, providing an asymmetry that permits the object to advance (Fig. 9.20). If the oar is metamorphic and can also vary its total surface area throughout the cycle, then this asymmetry can be enhanced.

Another example is the metamorphic doughnut-shaped nanorobot (Fig. 9.21A). An invaginating torus (e.g., a “smoke ring” motion) can swim along a vector normal to the torus plane because its outer surface is larger and moving faster than its inner surface (in the doughnut’s “hole”), giving rise to a differential viscous force in the same direction as the inner surface is moving—the opposite travel direction from what might be expected in an inertia-dominated environment. Differential rotation of circumaxial surface segments allows steering.

Yet another example proposed by Purcell is a single device constructed as two reversibly-adhered continuously-rotating spheroids (Fig. 9.21B). The spheroids stick together but are free to roll all over each other’s surfaces, by selectively binding and releasing mateable surface elements. A differential viscous force arises because the outermost surface of the pair is always fully bathed in viscous fluid but the inner contact surface is partially shielded from the fluid. By altering speeds and rotational axes, the pair can establish a velocity vector in any direction in three-dimensional space. A similar locomotive method is provided by Solem’s “viscous-lift helicopter” design (Fig. 9.22). Viscous drag on each of the four rotating wheels is

\[ F_{\text{wheel}} = \left(\frac{16 \pi}{3}\right) \eta \nu_{\text{wheel}} R_{\text{wheel}}^2 \]

for wheels of radius \( R_{\text{wheel}} \) turning at frequency \( \nu_{\text{wheel}} \) (Hz) in a medium of viscosity \( \eta \) and device mass is

\[ m_{\text{nano}} = 2 \pi \rho_{\text{nano}} R_{\text{wheel}}^3 \]

for a device of density \( \rho_{\text{nano}} \). Taking \( \nu_{\text{wheel}} = 10 \text{ KHz} \), \( R_{\text{wheel}} = 1 \text{ micron} \), \( \rho_{\text{nano}} = 2000 \text{ kg/m}^3 \), and \( \eta = 10^{-3} \) for water, then \( F_{\text{wheel}} \approx 170 \text{ pN per wheel} \), or ~0.7 nN for the whole device. There is a vast range of possibilities, according to Purcell: “Turn anything—if it isn’t perfectly symmetrical, you’ll swim.”

While incompressible tangential surface deformations alone are not propulsive, even single-sphere swimmers can translate or rotate using specific cyclic, nonreciprocal, compressible surface distortions (i.e., traveling waves). Such waves have been suggested as the mode of locomotion employed by appendageless spheroidal cyanobacteria, and surface undulations passing backward from the advancing edge of the cell have been observed in mammalian fibroblasts. Swimming speed for an object of radius \( R_{\text{nano}} \) is ~10 \( R_{\text{nano}} \) per second using a radial deformation of \( \varepsilon = 0.05 \) (5%); with \( n = 10 \) surface ripples, energy efficiency is approximated by \( e\% \sim (3 \times 10^{-3} \nu_{\text{wheel}} + 0.1)^{-1} \).

* The smallest known swimming fish is the dwarf pygmy goby (Pandaka pygmaea), measuring 7.1-9.7 mm long in the adult form.
A more familiar example of a deformable natation structure is the simple cilium (Section 9.3.1.1). In ciliate protozoa, cilia are typically much shorter than body length and are arranged in large numbers in rows on the body surface. Individual cilia have a regular beat pattern, typically including a rapid forward power stroke at full extension, followed by a slower return or recovery stroke with the cilium bent close to the body surface. Ciliary arrays (Section 9.3.4) may also display metachronism, or time-synchronized wave patterns, that can be used for precision steering and speed control, as, for example, by a 60 micron x 220 micron (~600,000 micron3) Paramecium caudatum, which rotates as it progresses forward in its path (Fig. 9.23) using the ~2500 cilia on its outer surface. The swimming paramecium may adjust its wave patterns while negotiating a liquid medium of varying viscosity. These and similar results should be generalized and extended to the case of nanorobotic ciliary propulsion.

For a medical nanorobot coated with propulsive cilia with wave velocity $v_{cilium}$ then $v_{nano}/v_{cilium} = 0.125-0.25$ and $e\% = 0.125-0.30$ (12.5%-30%). If $v_{nano} = 1$ cm/sec, then $v_{cilium} = 4-8$ cm/sec giving $p_{shear} \sim 2-6$ N/m² (Section 9.4.2.2), well within the normal range for human blood and probably nonthrombogenic to platelets. Ciliary propulsion even by large protozoa costs relatively little energy. For example, a 220-micron long paramecium swimming at ~1 mm/sec in water experiences ~1 nN of drag force and consumes $P_{drag} \sim (1 \text{ pW} / e\%)$ of power to locomote, but the animal can be very inefficient because ~10,000 pW are available to a paramecium with power density ~$10^4$ watts/m³ (est. from Table 6.8). Observed propulsion velocities of paramecia range from 0.2-2.5 mm/sec. These and similar results should be generalized and extended to the case of nanorobotic ciliary propulsion.

Another class of mechanisms for natation makes use of the inclined plane, a basic mechanical device that can convert viscous forces into forward motion. The simplest example is a threaded screw. In standard propeller theory at high Reynolds number, forward thrust is proportional to the rate at which a mass of fluid can be ejected out the rear (e.g., inertial forces). However, at low Reynolds number, the fluid that is pushed backwards by the rotating tilted planes does not provide thrust primarily by its inertial movement, but rather serves as a resistive medium against which the device can push itself forward. In the world of the nanorobot, the environment is very thick and viscous. The motive effect is not unlike the forward motion achieved by a threaded screw as it is screwed into a piece of wood using a screwdriver.

The motive force and power consumption of a microscale screw drive (Fig. 9.24) with pitch angle $\varphi$ and mean radius $R_{screw}$ may be very crudely approximated as follows. Consider a helical ribbon of width $w_{thread}$ and total length $l_{thread}$ that is wrapped around an axially-translating cylindrical body, making a pitch angle $\varphi$ as measured from normal to the direction of travel of the screw body. From Stokes law (Eqn. 9.73), a square element of that ribbon with area $w_{thread}^2$ experiences a maximum drag force of $F_{max} \sim 6 \pi \eta w_{thread} v_{thread}$. There are $n_{elements} = l_{thread}/w_{thread}$ square elements in the entire ribbon; neglecting flow field interactions of the elements and of the solid center for this approximation, the maximum laminar drag force on the entire ribbon is $F_{max} \sim 6 \pi \eta l_{thread} v_{thread}$. Viscous drag is lowest at $\varphi = 0^\circ$ (edge on) and highest at $\varphi = 90^\circ$ (face on); a factor of $(3-\cos(2\varphi))/4$ captures the experimental behavior of needle-shaped bodies which fall in viscous media about half as fast sideways as they do end-on (Section 9.4.2.4), with periodicity of $\pi$. The number of threads around the screw is $N_{thread} = l_{thread} \cos(\varphi) / \pi$. 
(2 \pi R_{\text{screw}}) and the screw rotates at a frequency \( \nu_{\text{screw}} = \nu_{\text{thread}} / (2 \pi R_{\text{screw}}) \), hence the total force required to turn the screw is:

\[
F_{\text{screw}} = 6 \pi^3 \eta \nu_{\text{screw}} N_{\text{thread}} R_{\text{screw}}^2 (3 - \cos(2\eta)) / \cos(\eta) \]

{Eqn. 9.76}

with total drag power \( P_{\text{screw}} = F_{\text{screw}} \nu_{\text{thread}} = 12 \pi^4 \eta \nu_{\text{screw}}^2 N_{\text{thread}}^2 R_{\text{screw}}^3 (3 - \cos(2\eta)) / \cos(\eta) \). To further simplify the calculation, we assume a “no slip” condition such that each complete revolution of the screw carries the nanorobot forward by a distance \( \approx 2 \pi R_{\text{screw}} \tan(\eta) \), although more slip may occur as the thread becomes looser (e.g., at high \( \eta \) and low \( N_{\text{thread}} \)). To avoid turbulence in fluid passing through the threads, from Eqs. 9.29 and 9.65, we must require \( \nu_{\text{thread}} / R_{\text{screw}} < 2000 \eta / \rho L \), a condition easily met for \( L < 1 \) micron devices. Under “no slip” conditions, the velocity of forward translation is approximated by \( v_{\text{nano}} = 2 \pi R_{\text{screw}} \nu_{\text{screw}} \tan(\eta) \), giving from Stokes law a net forward towing force of \( F_{\text{nano}} = 6 \pi \eta R_{\text{screw}} v_{\text{nano}} \), and a net mechanical efficiency \( e% = 2 \cos(\eta) \tan(\eta) / [\pi N_{\text{thread}} (3 - \cos(2\eta))] \).

Taking \( R_{\text{screw}} = 1 \) micron, \( \nu_{\text{thread}} = 0.1 \) micron, \( N_{\text{thread}} = 1 \) turn, \( \eta = 60^\circ \), \( 1.1 \times 10^{-3} \) kg/m-sec for plasma at 310 K, and \( \nu_{\text{screw}} = 920 \) Hz, then \( v_{\text{nano}} = 1 \) cm/sec, \( F_{\text{nano}} = 0.6 \) pN, \( \nu_{\text{thread}} = 0.1 \) cm/sec, total power requirement is \( P_{\text{screw}} = 7.6 \) pW, efficiency is \( e% - 0.27 \) (27%), and the pressure at the screw thread surface is \( \rho_{\text{thread}} = F_{\text{screw}} / (\nu_{\text{thread}} \nu_{\text{thread}}) = 10^3 \) N/m² \( \ll \) \( -5 \times 10^5 \) N/m² (required to induce transient cavitation in water at this frequency; Section 6.4.1). The outside edge of the screw thread is blunted to minimize energy transfer to impacted biological blood elements. A second counterrotating reverse-threaded screw mounted coaxially doubles the motive force while reducing net viscous torque on the natactor to zero. If each screw of the screw-pair is mounted on gimbals, the nanorobot can achieve controlled translation and rotation in any direction in three-dimensional space; changing leading screw rotation from clockwise (CW) to counterclockwise (CCW) enables the nanodevice to reverse direction or to undertake more complex motions.

Another well-known instance of the inclined plane in locomotion is the corkscrew drive (Fig. 9.25), of which the bacterial flagellum is the most familiar biological example.\(^{216,581,1395,1397}\) The flagellum is a closely-packed rigid helix \(~20 \text{ nm in diameter (with a ~3 nm flagellin protein core)}, and its length is almost always more than 100 times its thickness,\(^{1386}\) up to 10 microns long. The bacterial flagellum is turned by a \(~0.0001 \text{ pW} \) motor that rotates up to 300 Hz at 310 K (\~15 Hz under load) and can reverse its direction of rotation in \~1 millisecond (Section 6.3.4.2). The bacterium typically uses about 0.1% of its available metabolic energy (under growth conditions) to run the flagellum.\(^{581}\) Forward motion may be achieved using either planar waves or (more efficient) spiral/helical waves. The highest swimming speed attainable by a flagellate (body length up to 50 microns with flagella \>100 microns long in eukaryotes) is \(~50\)% of the front-to-back wave speed of its flagellum,\(^{1380}\) although \(~20\)% is more typical.\(^{1401}\) Measured swimming speeds are up to 100-200 microns/sec in various species of bacteria.\(^{1429}\) The flexural rigidity of the bull sperm flagellar tail is \( 30 \times 10^{-21} \text{ N-m}^2 \) in the rigor state, and \( 2 \times 10^{-21} \text{ N-m}^2 \) in the more flexible state in the presence of ATP.\(^{1451}\) See Eqn. 8.6. Energy efficiency may vary widely (Section 9.4.2.4).

9.4.2.5.3 Volume Displacement

Locomotion may also be achieved in a high-viscosity medium by physically removing fluid from the path ahead, then occupying the void before it closes up, and filling the volume behind with the vacated fluid—a process not unlike a tunneling mole. Forward velocity is largely determined by the speed at which blocks of fluid can be transferred from one end of the nanorobot to the other, whether along internal or external pathways.

Fluid pumping (Section 9.2.7) provides the simplest means of volume displacement. Consider a tube of radius \( r_{\text{tube}} = 100 \) nm and length \( L_{\text{tube}} = 1 \) micron installed across the diameter of a 1-micron wide spherical nanorobot \( (R_{\text{nano}} = 0.5 \) micron), along with a single positive displacement piston pump of single-cycle displacement \( V_{\text{cycle}} = 0.008 \) micron\(^3\) (Fig. 9.1). From the relations given in Section 9.2.7.2, operating the pump at \( V_{\text{pump}} = 1 \) MHz produces a time-averaged fluid flow velocity of \( \nu_{\text{flow}} = 26 \) cm/sec, a volume flow rate of \( V_{\text{pump}} = 8000 \) micron\(^3\)/sec, a nanorobot velocity of \( v_{\text{nano}} = V_{\text{pump}} / (\pi R_{\text{nano}}^2) = 1 \) cm/sec, and a power draw of \( P_{\text{flow}} = 1800 \) pW. From Eqn. 9.25, \( \Delta P = 2 \) atm across the nanorobot diameter, though this might be reduced to \~0.02 atm by flaring the entry and exit portals. Power usage, operating pressure, and the risks of clogging or fouling are relatively high, but this system has the virtue of simplicity, occupies less than 8% of total device internal volume, and can be expanded to three-dimensional motility by adding orthogonal entry portals. Jet propulsion is energy inefficient at small sizes and high speeds.\(^{2022}\)

9.4.2.5.4 Viscous Anchoring

By extending ahead of it a structure whose surface area is capable of being expanded, a nanorobot can set an anchor forward along its path, then winch itself towards the anchor. In a typical cycle, the deflated anchor is thrust forward. Then the anchor is expanded tenfold in area (thus increasing its viscous resistance to backwards movement by approximately tenfold). The nanorobot pulls itself forward by attempting to reel in the anchor. The anchor is then deflated, and the cycle is repeated. The anchor mechanism may resemble an umbrella that is opened and closed, or a balloon that is inflated and deflated, at the end of a telescoping rod. A second anchor placed aft and countercycled with the first permits continuous motion by alternating pushing with pulling; noncoaxial anchor settings allow arbitrary rotation and access to all headings in three-dimensional space.
Energy efficiency can be good because less fluid may be sheared during nanorobot movement, although in 1998 this method had not been widely studied. If the forward anchor is pushed many diameters ahead before inflation, then even the expansion of an anchor much larger than the nanorobot might prove less energy-expensive than anything except a device with skin than could flow backwards to match laminar streamlines (Section 5.1).

9.4.2.6 Additional Considerations

The need for safe shear stresses (Section 9.4.2.2), and nanorobot power requirements for swimming (Section 9.4.2.4), both suggest that the maximum sanguinationation velocity to be employed by nanorobots in the human body in normal circumstances should be \( \leq 1 \text{ cm/sec} \). Also, from Eqn. 9.65, a 1-micron nanorobot in water has a Reynolds number \( \mathsf{N}_R \sim v \); hence, to remain in the purely viscous regime, \( v < < 1 \text{ m/sec} \). The following analysis provides additional support for this speed limit.

Consider a spherical nanorobot of radius \( R_{\text{nano}} \) swimming at velocity \( v_{\text{nano}} \) in the bloodstream, that impacts a passing blood cell. Virtually all cells encountered in this way will be erythrocytes. The theory could be accommodated by administering compensatory velocity \( v_{\text{nano}} \) in the bloodstream, that impacts a passing blood cell.

\* But only “relatively”—for example, the multi-body collision frequency is appreciable among red cells even at hematocrits as low as 5%.32

Note that a 1 cm\(^3\) therapeutic dose of 1 micron\(^3\) nanorobots includes \(-10^{12}\) devices, each of cross-sectional area \(-1 \text{ micron}^2\). Such a fleet would occupy \(-1 \text{ m}^2\) if spread out uniformly on a surface, adjacent and one layer thick. The total surface of the entire human vascular system is \(-313 \text{ m}^2\), but the summed area of all large veins and main arterial branches alone totals \(-1 \text{ m}^2\) (Table 8.1). Thus, small therapeutic doses of medical nanorobots may safely travel the cell-free plasma “freeway” at somewhat higher speeds than previously estimated, though possibly at the expense of significantly greater power consumption (Section 9.4.2.4) and possibly, at the highest number densities, interfering with the lubrication effect normally provided by the plasmatic layer, especially in the capillaries.

The need to maintain moderate viscosity of nanorobot-rich blood (Section 9.4.1.4) and to avoid complete plug flow in the narrowest human capillaries (Section 9.4.1.5) implies that the maximum nanocrit to be employed by nanorobots in the human bloodstream in normal circumstances should be \( N_{\text{ct}} \leq 10\% \). The following simple analysis provides another constraint that is consistent with this estimate, and is valid for rigid and metamorphic nanorobots alike.

The frequency of close encounters and collisions among nanorobots may be taken as a conservative metric of the likelihood of physical jamming, mission interference, and other pathological effects of crowding. Consider a spherical nanorobot of radius \( R_{\text{nano}} \) and volume \( V_{\text{nano}} = (4/3) \pi R_{\text{nano}}^3 \), and a second spherical nanorobot of identical size that approaches the first until the two are in contact, defining a “collision” event. If the number density \( n_{\text{nano}} = 3 \text{ Nct} / [4 \pi R_{\text{nano}}^3 (1-\text{Hct})] \) of a uniformly distributed population of such nanorobots exceeds 2 devices within a spherical volume of radius \( 2R_{\text{nano}} \), then, on average, all devices are in collision. This defines a maximum upper bound for nanocrit of \( N_{\text{ct}} \leq 15\% \) for Hct = 46% in the arteries, 22.5% for Hct = 10% in the capillaries. Similarly, if the nanorobot number density is less than 1 device within a spherical volume of radius \( 2R_{\text{nano}} \) and volume \( V_{\text{nano}} = (4/3) \pi R_{\text{nano}}^3 \), then, on average, all devices are in collision. This defines a maximum upper bound for nanocrit of \( N_{\text{ct}} \leq 30\% \) for Hct = 46% in the arteries, 40% for Hct = 10% in the capillaries. Thus, in the arteries, an \( N_{\text{ct}} \leq 6.8\% \) is relatively noncollisional, \* an \( N_{\text{ct}} > 13.5\% \) is highly collisional, and a \( 6.8\% < N_{\text{ct}} < 13.5\% \) (midpoint \( N_{\text{ct}} = 10\% \)) is transitional.

\* But only “relatively”—for example, the multi-body collision frequency is appreciable among red cells even at hematocrits as low as 5%.32
9.4.3 Cytoambulation

Many applications of medical nanorobots will require the ability to “walk” across biomaterial-embedded cellular tissues, packed-cell blocks, or vascular surfaces, a process called cytoambulation. Cytoambulation may involve treading the surfaces of blood or lymph vessel walls which are lined with endothelial cells; negotiating flexible surfaces such as the walls of the urinary or gall bladders; traversing interior synovial or bursal surfaces such as the articular cartilage inside skeletal joints; and walking the walls of various excretory ducts, bulbs, tubes, and glands. Cytoambulating nanorobots may be required to cross various classes of tissue and cellular membranes. Tissue membranes are typically composed of epithelial and connective tissues and include:

a. serous membranes lining the cavities of the body and surrounding the various organs (e.g., the pericardium, pleura, peritoneum, and meninges);

b. mucous membranes lining the alimentary, respiratory, and genitourinary tracts; and

c. fibrous membranes composed entirely of connective tissues (e.g., perichondrium, periosteum, and synovial membranes).

Traversing such tissue membranes will often involve crossing the individual plasma membranes of tissue-embedded cells of various classes. Nanorobots also may cytoambulate across the surfaces of motile or free-floating cells, including individual erythrocytes and red-cell rouleaux, leukocytes, fibroblasts, platelets, large protozoa, and even some bacteria.

A comprehensive survey of all possible cytoambulatory modes and requirements is beyond the scope of this book. This Section is restricted to a discussion of basic cell-surface walking, including the single footpad contact event (Section 9.4.3.1), cell plasma membrane elasticity (Section 9.4.3.2), anchoring and dislodgement forces (Section 9.4.3.3), the physical limits of contact event cycling (Section 9.4.3.4), and examples of a few specific physical nanomechanisms for cytoambulation (Sections 9.4.3.5-9.4.2.8). The reader is encouraged to review Section 8.5.3 on cell structure before proceeding further.

9.4.3.1 Ambulatory Contact Event

Stripped to its fundamentals, the act of ambulation requires that a mobile surface (e.g., a footpad) is brought into adhesive contact with a fixed surface to be traversed. After the contact event has occurred, the point of adhesion becomes a fulcrum allowing mechanical leverage to be applied against the fixed surface, thus permitting forward motion of a mass attached to a lever. The mobile surface may then be detached from the fixed surface and the cycle repeated.

Consider a footpad approaching the extracellular surface of a cell. The footpad first encounters 5-8 nm thick strands of the circumcellular glycocalyx at a number density of \(10^5\) strands/micron\(^2\), typically at a distance of 10-100 nm from the plasma membrane surface (see Section 8.5.3.2). A reversible binding site embedded in the footpad contact area (Section 4.2.8) can recognize glycocalyx chains unique to specific cell types or tissues, or more generally any glycoprotein chain, and then securely bind to such chain until mechanically released. If cell-type specificity is not an important design requirement, then a simple reversible mechanical grappling system may be employed. Each carbohydrate chain may be up to 100-200 nm in length, so a footpad adhering to a cell surface in this fashion may experience up to \(-100\) nm of horizontal or vertical free play in its anchorage. This can seriously degrade the efficiency of motive levers (e.g., nanorobot legs) \(\leq 1\) micron in length, and can make ambulatory locomotion nearly impossible for motive levers \(\leq 100\) nm in length. (If the footpad adheres to many chains, free play is substantially reduced.)

Alternatively, footpads may employ spikelike structures and glyophobic coatings to penetrate the glycocalyx without injuring it. At 10-20 nm above the plasma membrane surface, these footpad structures will encounter the extracellular polar regions of integral proteins that are embedded in the cellular plasma membrane. Integral proteins, typically \(-100,000\) daltons, are present in RBCs at a number density of \(-10^7/micron^2\); however, \(-100\) different integral protein types are present, reducing the number density of any one common type (e.g., red cell glycoporphins) to \(-10^5-10^4/micron^2\). Free-floating integral proteins (constituting 20%-70% of the total number\(^{1435}\) exhibit considerable lateral diffusion but can provide adequate anchorage if the dwell time is brief enough. Transmembrane proteins are often immobilized by links to protein networks located near the cytoplasmic side of the plasma membrane, possibly reducing free play to \(\sim 20-30\) nm, although many transmembrane proteins are required to be unattached to the internal cytoskeleton, permitting them to circulate so they can mediate multi-receptor signal transduction.

Penetrating still further to the plasma membrane surface, reversibly-hydrophilic footpads may bind with the polar phospholipid heads of the \(-650\) dalton molecules comprising the lipid bilayer, with number density \(-2.5 \times 10^5/micron^2\) in each of the two layers. (The erythrocyte plasma membrane contains well over 100 different lipid species, \(^{1430}\) in much lesser concentrations.) Adjacent phospholipid molecules generally are not bound covalently, so individual molecules may be extracted from the lipid bilayer with a modest vertical force (taking \(-10\) zJ/nm\(^2\) interlipid hydrophobic binding energy (Section 3.5.1) and lipid-lipid contact area \(~4\) nm\(^2/\)lipid divided by single-layer extraction distance \(~4\) nm implies a single-lipid, single-layer extraction force \(F_{\text{lipid}} \sim 10\) pN, comparable to the estimate by Evans\(^{1415}\) and hence provide relatively weak anchorage in the vertical direction. Although such anchorage may suffice in many applications, artificial amphipathic transmembrane anchored structures (Section 9.4.3.3) can provide an even more secure bond for locomotion and parking.

After the contact event has occurred, the adhesion structure is used as a fulcrum to apply mechanical leverage against the cell plasma membrane surface, giving rise to tensile or shear forces at the point of attachment. How large are these forces? As a minimal estimate, we assume that each transmembrane protein is anchored to the internal cytoskeleton by only a single actin microfilament. An actin microfilament has a failure strength of \(\sim 2.2 \times 10^6\) N/m\(^2\) (Table 9.3) and a cross-sectional area of \(~30\) nm\(^2\) (Section 8.5.3.1), thus requiring a tearing force of \(-108\) pN for detachment. \(^{362}\) A glycoprotein molecule consisting primarily of a chain of C-C covalent bonds may have a rupture strength on the order of \(-10,000\) pN (Section 3.5.1).

* J. Hoh emphasizes that a large thermal contribution exists for forces in the piconewton range: the smaller the force, the larger the time-dependent contribution. Normal biomolecular interactions, such as antibody-antigen interactions at \(\sim 100\) pN, have finite lifetimes and will unbind on a timescale of hours or days in the absence of any force. Even a very small loading force may shift the unbinding timescale significantly downward. As a result, the binding forces for nanorobot footpads will be critically linked to the speed at which the nanorobot is moving. Hoh believes these variations in force may span orders of magnitude; additional research on nanorobot footpad mechanics is clearly needed.
When mechanical leverage is applied against the cell surface, plasma membrane components feel a force in the direction opposite to the direction of travel of the ambulating nanorobot. Since the plasma membrane is a fluid, components within the membrane may be free to slip backwards, reducing traction. For simplicity of analysis, consider a spherical nanorobot of radius $R_{nano}$ that inserts a spherical footpad of radius $R_{foot}$ into a lipid bilayer membrane, at the end of each of $N_{legs}$ legs each of a width that is small in comparison with $R_{foot}$ (so that the contribution from the legs may be ignored). The viscosity of the plasma membrane that resists the backward movement of the footpad is approximately $\eta_{membrane} = 10$ kg/m·sec and the viscosity of the extracellular fluid is $\eta_{extra} = 10^{-3}$ kg/m·sec (Table 9.4). The nanorobot ambulates across the cell surface through the extracellular fluid at a velocity $v_{nano}$ against a "headwind" of $v_{headwind}$ (the fluid speed relative to plasma membrane surface), during which movement the footpad slips backwards at a velocity $v_{foot}$. Equating the forces on the nanorobot and the footpad from Eqn. 9.73 gives:

$$v_{nano} = (\kappa_{traction} N_{leg} v_{foot}) - v_{headwind} \quad \text{(m/sec)} \quad \text{[Eqn. 9.79]}$$

where the coefficient of traction $\kappa_{traction} = \eta_{membrane} R_{foot} / \eta_{extra} R_{nano}$. Taking $N_{leg} = 2$, $R_{nano} = 1$ micron and $R_{foot} = 10$ nm, then $\kappa_{traction} = 200$. Thus with $v_{headwind} = 0$, the footpads of a bipedal nanorobot traveling forward at $v_{nano} = 1$ cm/sec slip backward within the lipid bilayer membrane at only $v_{foot} = 50$ microns/sec. This process is functionally similar to the footpad system modeled by kinesin microtubule transport (Section 9.4.6) wherein productive steps are directed by protein-protein interaction affinities.

Repeated application of physical forces against a cell surface may activate mechanical signal transduction pathways mediated by specific transmembrane proteins such as the integrins and cadherins (Sections 8.5.2.2 and 8.5.3.11). These signals can trigger major changes in cell behavior and in cellular protein expression. Microfilaments readily transduce 0.01-2 Hz mechanical signals throughout the cell. In 1998, it was not yet known whether much higher frequencies, such as might be employed in cytoambulation, also induce a significant cellular response, except, of course, in specialized mechanoreceptor cells such as the cochlear stereocilia. Footpad target proteins, motions, and operating frequencies should be selected with the objective of minimizing mechanical signal transduction effects.

**9.4.3.2 Cell Plasma Membrane Elasticity**

Does the temporary attachment and passage of a cytoambulating nanorobot across a cell surface have a significant mechanical effect on that surface? Cell and plasma membrane deformation can be quite complex. However, taking the plasma membrane to be a two-dimensional, incompressible elastic solid, a deformation will always consist of one or more of three fundamental independent deformations: expansion without bending or shear (characterized by the area expansion modulus $K$, in N/m), elongation (shear) without expanding or bending (shear modulus $\mu_{shear}$ in N/m), and bending without shear or expansion (bending modulus $B$, in N-m). Each mode is discussed briefly, below.

9.4.3.2.1 Plasma Membrane Areal Expansion Elasticity

A nanorobot with a footpad of width $L_{foot} = 10$ nm that applies a pressure of $P_{nano} = 10$ pN / $L_{foot}^2$ across a span $L_{foot}$ of plasma membrane surface (and normal to it) creates an isotropic tension of $T = P_{nano} L_{foot} = 1 \times 10^3$ N/m within the membrane, causing the plasma membrane under the nanorobot to expand uniformly without shearing or bending. For comparison, the tension required to smooth out the thermal undulations or "Brownian motion" of the outer membrane of artificial phospholipid vesicles 10-20 microns in diameter (typical cell size) was determined experimentally to be 0.01-0.1 x $10^3$ N/m.668 At the other extreme, osmotically swollen red cells will instantaneously lyse* at a plasma membrane tension of $T_{lyse} = 10-20 \times 10^3$ N/m$^2$ or at an elastic modulus with respect to area dilation of $K_{lyse} / h_{cell} \approx 3 \times 10^6$ N/m²,1422 where red cell wall thickness $h_{cell} = 8$ nm.

The relative area expansion of the plasma membrane under the nanorobot footpad is $\Delta A / A = T / K$,1415 where the experimentally-determined area compressibility modulus is $K = 0.45$ N/m for red cell plasma membrane at 298 K,1414-1416,3171,3172 (change in area compressibility modulus with temperature is $-6 \times 10^{-3}$ N/m-K).1411 This gives $\Delta A / A = 0.2$% or 0.06% for the nanorobot footpad of size $L_{foot}$ described above (2%-4% areal expansion produces lysis1415), or a mean linear strain of $\varepsilon = (\Delta A / A)^{1/2} = 0.04$ (4%) or 0.02 (2%). For $\varepsilon << 1$, the footpad depresses the surface by $\Delta x = L_{foot} (\varepsilon / 2)^{1/2} = 1.5$ nm into the red cell's interior. This deflection represents only -15% of lipid bilayer membrane thickness and must be considered modest in most circumstances. For leukocyte plasma membranes, measured $K = 0.636$ N/m.846

The design of cytoambulation mechanisms should attempt to minimize the activation of sodium-ion (and many other) stretch-activated channels—gated transmembrane channels that are activated by simply stretching the plasma membrane, or by tension or stress development in cytoskeletal elements associated with the cell membrane.1417-1419 Such channels have already been implicated in the maintenance of cell volume.939,1416 Biochemical transduction of mechanical strain has also been investigated in bone cells during normal loading. Linear strains of $\varepsilon < 0.05$% were nonstimulative; those between 0.05%-0.15% maintained normal bone mass, and strains of $\varepsilon > 0.15$% stimulated osteoblasts to increase bone mass.1417-1419 Linear strains >1% induced osteoblasts to alter morphology, becoming fibroblast-like,1420 and red cells lyse at $\varepsilon \geq 20\%$.362 A limit of $\Delta A / A - 0.05$% for a 10 nm footpad on a red cell surface would give an estimate of ~2 pN for the activation threshold. However, a force of 1 pN across a protein stretch sensor of cross-sectional area ~1 nm² represents an energy density of $10^6$ joules/m³, or 10 J/(-2 kT) for a ~10 nm² sensor volume, which is probably close to the limit for biological force detection consistent with earlier estimates of nanosensor detection limits (Section 4.4.1).

From these comparisons, it appears that a biological response, stimulated by plasma membrane areal expansion due to the passage of nanorobot footpads across the cell surface employing forces typical in cytoambulation (e.g., ~20 pN; Section 9.4.3.5), cannot be ruled out. However, in 1998 it was not yet known how efficiently mechanical pressures cycled at 10-100 KHz are transduced by plasma membrane stretch sensors, since most mechanical cell stimulation

---

* $T \approx 6 \times 10^{-7}$ N/m produces red cell plasma membrane lysis after ~10 sec; time of lysis is a stochastic function that increases to ~120 days (maximum red cell lifetime) at $T \approx 3-4 \times 10^{-6}$ N/m.1421

**Specifically: 0.05 Hz, 0.1 Hz, 0.25 Hz, 0.33 Hz, 0.5 Hz, 1 Hz, 1.67 Hz, 2 Hz, 3 Hz, and 5 Hz.**
experiments have been conducted at frequencies between 0.05-5 Hz. Energy coupling may be poor due to a large mechanical impedance mismatch. Looser glycocalyx anchorages also could minimize stretch activation effects.

9.4.3.2.2 Plasma Membrane Shear Elasticity
A plasma membrane surface can deform by “shear”—elongation in one dimension while narrowing in another dimension—without bending or increasing surface area. If a deforming force normal to a plasma membrane surface is suddenly removed, the membrane surface will recover its normal unstressed shape in a characteristic time $t_{\text{shear}} = \eta_{\text{surface}} / \mu_{\text{shear}} \sim 0.1$ sec, where elastic shear modulus $\mu_{\text{shear}} = 6.6 \times 10^{-6}$ N/m for the red cell plasma membrane at 298 K. As determined experimentally, and the measured coefficient of surface viscosity $\eta_{\text{surface}} \sim 10^{-6}$ N-sec/m, is the product of plasma membrane thickness (~8 nm for RBC) and plasma membrane viscosity (~80 nm-sec for RBC), with measured $\eta_{\text{surface}} = 0.6-1.2 \times 10^{-6}$ N-sec/m for the red cell plasma membrane. (The change in elastic shear modulus with temperature is $-6 \times 10^{-8}$ N/m-K.) Shape recovery is dominated by plasma membrane viscosity, not cytoplasm viscosity, in some systems, particularly RBCs, however. J. Hoh [personal communication, 1999] notes that this does not hold for all cells, and that for most cell types it is the cytoplasm that dominates the rheological properties, as experimentally measured, for example, by indentors or piezo controlled microtapes. The time required for a leukocyte to enter a capillary is ~1000-2000 times longer than the time required by red cells, suggesting $t_{\text{snapback}} \sim 100$ sec for white cells.

With nanorobot legwring frequency $\nu_{\text{leg}} >> t_{\text{snapback}}^{-1}$, a section of plasma membrane distorted by the passage of one nanorobot footpad may not be able to resume its fully relaxed state before the arrival of another footpad in the same spot as the first. If a traverse through a narrow passage of cell-rich tissue by large numbers of ambulating nanorobots is contemplated for a particular application, then footpad placements should be randomized as much as possible to avoid repeat-deformation grooving, rutting, or other persistent indentation features in the plasma membrane surface.

9.4.3.2.3 Plasma Membrane Bending Elasticity
The bending stiffness (bending modulus) of the red cell plasma membrane at room temperature is $B = 1.8 \times 10^{-19}$ N-m. The energy required to bend the plasma membrane into a hemispherical cap is $\sim 4 \pi \frac{B}{2} \sim 2300$ J. A nanorobot footpad of size $l_{\text{foot}}$, that exerts a backward force $F = 10$ pN against the plasma membrane may cause the membrane surface behind the footpad to fold into a ripple with radius of curvature $r_{\text{curve}} \sim B / F \sim 20$ nm. The dynamic resistance to RBC plasma membrane folding appears to be limited by viscous dissipation in the cytoplasmic and external fluid phases. The plasma membrane has only a very weak resistance to bending.

9.4.3.3 Anchoring and Dislodgement Forces
A footpad which is noncovalently bound only to a single lipid or unanchored protein molecule embedded in the plasma membrane will slowly move as the membrane molecule(s) to which it is attached experiences translation diffusion according to the well-known Einstein-Smoluchowski equation:

$$\Delta X = (2D\tau)^{1/2} \quad \text{[Eqn. 9.80]}$$

which is readily obtained by combining Eqs. 3.1 and 3.5. For lipid probes, translational diffusion coefficients in artificial fluid bilayer systems generally range from $D = 1-10 \times 10^{-12}$ m$^2$/sec; for plasma membrane lipids in human fibroblasts, $D = 1.6 \times 10^{-12}$ m$^2$/sec at 310 K. Proteins reconstituted in phospholipid bilayers show diffusion coefficients from $D = 0.7-8 \times 10^{-12}$ m$^2$/sec. LDL receptors in human fibroblasts have $D = 0.2 \times 10^{-12}$ m$^2$/sec at 310 K; MHC Class I (HLA) antigens in human neutrophils, lymphocytes and fibroblasts have $D = 0.05-0.07 \times 10^{-12}$ m$^2$/sec, or 0.15-0.30 x $10^{-12}$ m$^2$/sec in human endothelial cells. There is a weak dependence on protein molecular size according to Monte Carlo simulations, a plasma membrane lattice covered up to 82% with impermeable (protein) domains diminishes $D$ by only a factor of 20 in comparison to the zero concentration limit. In erythrocyte plasma membranes, spectrin-anchored proteins such as the anion transport protein band 3 have $D = 0.0045 \times 10^{-12}$ m$^2$/sec, while spectrin-depleted cells where the same proteins are no longer anchored, have $D = 0.25 \times 10^{-12}$ m$^2$/sec. Additionally, for a cylindrical protein of radius $R = 3$ nm traversing a plasma membrane of thickness $h = 10$ nm, the rotational diffusion coefficient $D_{\text{rot}}$ is given by:

$$D_{\text{rot}} = \frac{4 \pi \eta_{\text{membrane}} R^2 h}{kT} = 380 \text{ sec}^{-1} \quad \text{[Eqn. 9.81]}$$

taking $\eta_{\text{membrane}} = 10$ kg/m-sec at 310 K, which gives a diffusional rotation of $\Delta \alpha = (2D_{\text{rot}} \tau)^{1/2} = 0.1-0.3$ radian taking $\tau \sim 10-100$ microsec.

For much longer parking times (e.g., $\tau \sim 1$ sec), in the classical fluid mosaic model of the cell plasma membrane (Section 8.5.3.2), $\Delta X \sim 1$ micron for footpads attached to lipids or unanchored proteins. However, recent work by Kusumi and Sako has demonstrated that a substantial fraction of unanchored proteins are transiently confined to domains somewhat smaller than this. According to their membrane-skeleton fence model, a spectrin-like meshwork (Fig. 8.43) closely apposed to the cytoplasmic face of the plasma membrane sterically confines transmembrane proteins to regions on the order of the cytoskeletal mesh size. The fences appear elastic, because, for example, transferrin receptors rebound after they strike barriers, and a small fraction of these receptors seem to be fixed to the underlying cytoskeleton by spring-like tethers. For cadherins, transferrin receptors, and epidermal growth factor receptors, the domains (e.g., the barrier-free path or BFP) are 300-600 nm in diameter.
and confinement lasts 3–30 sec. BFPs for the lipid-linked and membrane-spanning isoforms of the MHC (Section 8.5.2.1) antigens are ~1700 nm and ~600 nm, respectively, at a temperature of 296 K. Confinement was also found for a lipid-linked isoform of neural cell adhesion molecules (NCAMs) in muscle cells (which cannot be directly trapped by the cytoskeletal network), with BFP domains ~280 nm in diameter and a mean trapping time of ~8 sec. Still, footholds to well-anchored proteins may be preferred for the longest parking times.

An annular footpad of radius $R_{foot} = 10$ nm firmly attached around its circumference to a ring of plasma membrane-surface lipid molecules, taking each molecule of area $A_{lipid} \sim 0.4$ nm$^2$/molecule and assuming a single-lipid extraction force $F_{lipid} = 1$ pN (Section 9.4.2.2), gives a typical dislodgement force of $F_{dis} \sim 2$ pN on a 1-micron$^2$ nanorobot. (The net attractive force of ~70 pN between red cells in a rouleaux is also ~2 N/m$^2$; locomotion has been measured as ~165 nN, or ~1000 pN/micron$^2$ giving $t_{swing} = 8$ microsec and $t_{arc} = 80$ nm giving $t_{swing} = X_{arc}/X_{arc}/N_{leg} = 8$ nm in a typical footpad environment. (There are also ~100 are phospholipids, ~50 are cholesterols, 10-75 are glycolipids, and the remaining 25-90 are other lipids.) Alternatively, mechanical grippers that close and open on a 5-nm wide molecular target using a 1 cm/sec jaw speed require ~1 microsec per gripping cycle.

Finally, a mechanical manipulator appendage traveling at a conservative ~1 cm/sec (Section 9.3.1) can transit the 10-100 nm glyocalyx, reaching the plasma membrane surface, if necessary, in 1-10 microseconds. A similar amount of time must be allowed for the limb to be retracted from the cell surface. These figures are all consistent with a maximum legswing frequency $v_{leg} \sim 1$ KHz.

9.4.3.3 Legged Ambulation

Consider a 1 micron$^3$ nanorobot cytoambulating using legs tipped with appropriate footpads to traverse a vascular wall. Each leg is assumed to be similar in size and function to the 100-nm long, 30-nm wide cylindrical telescoping nanomanipulator described in Section 9.3.1.4. Each nanorobot has a total of 100 legs, occupying ~7% of the 10$^6$ nm$^2$ underside area of the device. To allow tenfold redundancy, at any one time only $N_{leg} = 10$ legs are deployed and in use. The remainder are stowed as spares.

Ignoring ~1% traction losses (Section 9.4.3.1), the total force that must be supplied by all $N_{leg}$ working legs is:

$$F_{total} = \sum_{leg} F_{dis} = 100 \text{pN}$$

The maximum dislodgement or “headwind” force normally encountered along blood vessel walls is $F_{dis} \sim 40$ pN (Section 9.4.3.3). The viscous force on each leg is approximated by Eqn. 9.75 as $F_{visc} = \eta v_{leg} L_{leg} / R_{leg}$, giving $F_{visc} \sim 100$ pN, $R_{leg} = 15$ nm, and $v_{leg} = 1$ cm/sec. From Eqn. 9.73, $F_{nano} = 100$ pN, taking $\eta = 1.1 \times 10^5$ kg/m-sec for plasma at 310 K, $R_{nano} = 0.5$ micron, and $v_{nano} = 1$ cm/sec, giving $F_{total} = 200$ pN, easily within the capacity of a single leg (Section 9.3.1.4), giving an allocation of $F_{total}/N_{leg} = 20$ pN per leg. Maximum safe towing force is $\approx 300$ pN/leg, assuming 100 nm$^2$ footpads and a $3 \times 10^6$ N/m$^2$ membranolytic limit for plasma membrane.

Many N-podal gait possibilities are possible. In the most conservative gait, only 1 leg is moved at a time while the remaining ($N_{leg} - 1$) legs stay anchored at their footpads. Given a full center-to-center working arc of $X_{arc} = 80$ nm, each leg must travel $X_{swing} = X_{arc}/N_{leg} = 8$ nm in a time $t_{swing} = X_{swing} / v_{leg} = 0.8$ microsec at a velocity $v_{leg} = 1$ cm/sec, with a per-leg duty cycle of $d_{leg} = N_{leg}^{-1} = 10\%$ and an operating frequency of $f_{leg} = 100$ Hz. From Eqn. 9.74, nanorobot motive power is $P_{nano} = F_{total} v_{nano} / e\% \sim 10$ pW, taking $e\% = 0.20$ (20%). Conservatively taking each footpad binding event as costing $E_{bind} = 100$ zJ (Section 4.2.1), then footpad binding power requirement is $P_{bind} = N_{leg} V_{leg} E_{bind} \sim 0.1$ pW, a negligible contribution.

For the least conservative gait, only 1 leg stays anchored while the remaining ($N_{leg} - 1$) legs are in motion. In this case, $X_{swing} = X_{arc} = 80$ nm giving $t_{swing} = 8$ microsec and $v_{leg} \sim 10$ KHz, but motive power, nanorobot velocity, and leg duty cycle are unchanged. Doubling leg length to $L_{leg} = 200$ nm while holding $R_{leg}, V_{leg}$ and $V_{nano}$ unchanged decreases operating frequency to $v_{leg} \sim 5$ KHz while increasing $F_{total}$. 

9.4.3.4 Contact Event Cycling

How fast can ambulatory contact events be cycled? Molecular receptors ~10 nm$^2$ in size can bind large (5-10 nm), “common” (~10$^{-2}$/nm$^3$) molecules in t$_{EQ} \sim 0.2$ microsec (Section 4.2.1). With -10$^4$/micron$^2$ glyocalyx strands and a similar number of integral proteins near the cell surface (Section 9.4.3.1), there are ~10 strands and ~10 proteins under each 100 nm$^2$ footpad. For either glyocalyx strands or integral proteins, mean separation is ~3 nm giving a concentration of ~4 x 10$^{-7}$/nm$^3 > 10^2$/nm$^3$, thus qualifying as “common” molecules in the footpad environment. (There are also ~250 bilayer lipid heads directly beneath a 100 nm$^2$ footpad, of which ~100 are phospholipids, ~50 are cholesterols, 10-75 are glycolipids, and the remaining 25-90 are other lipids.) Alternatively, mechanical grippers that close and open on a 5-nm wide molecular target using a 1 cm/sec jaw speed require ~1 microsec per gripping cycle.

For the least conservative gait, only 1 leg stays anchored while the remaining ($N_{leg} - 1$) legs are in motion. In this case, $X_{swing} = X_{arc} = 80$ nm giving $t_{swing} = 8$ microsec and $v_{leg} \sim 10$ KHz, but motive power, nanorobot velocity, and leg duty cycle are unchanged. Doubling leg length to $L_{leg} = 200$ nm while holding $R_{leg}, V_{leg}$ and $V_{nano}$ unchanged decreases operating frequency to $v_{leg} \sim 5$ KHz while increasing $F_{total}$ to
230 pN and $P_{\text{nano}}$ to 12 pW. Doubling the velocities doubles force and operating frequency, and quadruples the power demand. Perhaps counterintuitively from common macroscale experience, for small ambulators traveling in viscous-dominated media (e.g., low Reynolds number ambulators), shorter legs may produce the highest motive velocity for the lowest power requirement and applied force.

As a biological analog, tiny extensible hydraulic tube feet specialized for burrowing and stepping locomotion, often with terminal suckers, have been extensively described in echinoderms.1472-1475

### 9.4.3.6 Tank-Tread Rolling

When an injured cell “calls for help” from leukocytes, the cell secretes cytokines such as IL-1 and TNF, causing nearby vascular endothelial cells to express P-selectin and E-selectin on their luminal surfaces.415 Passing white cells adhere to these protruding molecules because their carbohydrate coat contains complementary structures. When a leukocyte touches a venule wall, its rate of movement slows because of the stickiness between the cell and the wall, but the force of the circulating blood keeps the cell moving with a tank-tread motion in what is called a rolling interaction (Fig. 9.28). The average white cell rolling velocity is only <4% of the mean blood flow velocity in the venule, typically ~10-40 microns/sec.1027,1484,1507

This form of locomotion is easily implemented by rigid medical nanorobots. A primitive example would be a spherical device with numerous flexible knobs projecting radially from its surface, located in the tip of each knob is switched on or off, depending on the direction because only footpads must be operated. In the limit of a flexible knob-dense surface, a nanorobot with ~1 micron$^2$ of continuous contact surface could, in theory, tow up to 3,000,000 pN of load before reaching the $\sim 3 \times 10^6$ N/m$^2$ membranolytic limit.1422

![Fig. 9.26. Protoplasmic streaming in a monopodal amoeba with velocity profiles at three loci (modified from Holberton).](image)

### 9.4.3.7 Amoeboid Locomotion

The most common cells exhibiting amoeboid locomotion in the human body are white cells, having moved out of the blood into the tissues in the form of tissue microphages or macrophages. Fibroblasts and normally sessile germinial skin cells can move into damaged areas to assist in wound repair. Embryonic cells in the fetus such as neurons often migrate long distances to their final location by amoeboid movement, after which they become fixed in tissue as sessile cells. The 200-600 micron carnivorous Amoeba proteus also displays this form of locomotion and was the subject of the earliest studies, hence the name.

Amoeboid movement is associated with two properties—cytoplasmic streaming, and the extension and retraction of pseudopods—the motive effects of which could be simulated by medical nanorobots using metamorphic exterior surfaces (Section 5.3). As shown in Figure 9.26, the monopodal amoeboid cell progresses by establishing a series of attachment points or focal contacts with the surface it is traversing. Viewed from the side, the amoeba steps forward on new pseudopods that make adhesive contact with the surface, with the tail region and retracting pseudopods lifted clear of the surface.1466 The plasma membrane with its mucus coat is a relatively permanent structure that plays a passive role during locomotion,1465 rotating forward during locomotion something like a water balloon rolling across a tilted tabletop.

Inside the plasma membrane are two regions of cytoplasm with varying viscosity—the actin microfilament-rich outer portion, called the ectoplasm or cell cortex, and the inner portion called the endoplasm. To move forward, the ectoplasm at the front end becomes thin, causing a pseudopodium to bulge forward; the ectoplasm at the tail end contracts, pushing endoplasm into the pseudopodium and extending it further.1466 Actin monomers in the pseudopods are gelated to actin polymers in the pseudopods, then solated back to monomers in the tail. (Gels depolymerize when Ca$^{++}$ concentration is reduced.) The surface for new pseudopod formation is unfolded from the sink of convoluted surface in retracted pseudopods and in the tail. In each pseudopod, the surface rolls forward over the stationary ectoplasmic tube on a lubricating layer of hyaloplasm dispersed from the hyaline cap.1380

Isolated amoebic cytoplasm, when injected with ATP, streams spontaneously at up to 160 microns/sec.1394

The formation of a focal contact between fibroblast and surface is a biphasic process in which the fast (~1 sec) initial phase of establishing small (~0.25 micron$^2$) immature contacts at large separation distances is followed by a slower (~15 sec) phase of widening mature contacts with narrowing separation distances.1457 Leukocyte pseudopodia are ~1 micron in length with a typical extension velocity of ~0.08 microns/sec.840 so white cells and fibroblasts readily cytoambulate at ~0.05-0.1 microns/sec;159,1513 chemotactically-stimulated leukocytes can locomote at ~0.7 microns/sec.1516 An individual fibroblast has a towing strength of ~165 nN,1461 giving an implied low-load motive power demand of only ~0.02 pW (force x velocity). Measured motive force of a whole amoeba (~towing force) is 50-290 nN.1462,1469 or ~100 nN per

![Fig. 9.27. Schematic of neutrophil cytoambulation (redrawn from Horwitz).](image)
50-micron pseudopod; internal hydrostatic pressures have been estimated from 10-100 N/m$^2$. Ambulatory velocity ranges from 1-50 microns/sec, but averages ~10 microns/sec. A nanorobot using metamorphic pseudopods spaced ~1 micron apart and cycling its adhesive contacts at 10 KHz could achieve $v_{nano} \sim 1$ cm/sec within a ~10 pW motive power budget (Section 5.3.1.4).

Neutrophils walk from one site to another by forming and breaking integrin-mediated attachments to a matrix (Fig. 9.27). Integrin receptor surface density is ~2000/micron$^2$ on neutrophils, with each fibringen attachment having a detachment strength of ~2 pN.

### 9.4.3.8 Inchworm Locomotion

Another ambulatory strategy that has been investigated for medical robots is inchworm locomotion. In this mode, the tail end of a flexible tubular nanodevice is anchored. The front end extends in the desired direction across the cell surface by pushing against the anchor. Once fully extended, the front end sets anchor. The rear end unlocks, slides itself forward toward the front anchor across the cell surface, then resets in preparation for the next cycle. Inchworm mobility may be particularly applicable to environments with extremely rough, uneven terrain. A related strategy is "Slinky" mobility may be particularly applicable to environments with uneven, uneven terrain.

### 9.4.4 Histonatation

Histonatation, or swimming through biological tissues, will be necessary for medical nanorobots that need to reach a specific histological or cellular target to begin diagnostic or repair work. Each tissue that must be traversed has its own unique set of biochemical, immunochemical, mechanical, electrophysiological, and other characteristics that will affect the precise mode and manner of locomotion. A complete survey of all possible tissue types is beyond the scope of this book. This Section briefly considers a few important issues in the following general circumstances—exiting endothelial-cell-lined blood vessels to enter the tissues (Section 9.4.4.1), passage through acellular tissue (Section 9.4.4.2), passage through cell-dense tissue (Section 9.4.4.3), and tissue-resident nanorobot conjugation and partition (Section 9.4.4.4).

### 9.4.4.1 Nanorobot Diapedesis

The passage of formed blood elements, especially white cells, through the intact walls of blood vessels, is called transendothelial migration or diapedesis (Fig. 9.28). Diapedesis is a stereotyped behavior of certain motile cells that may occur everywhere in the human body except (usually) through the brain-blood barrier. Normally, ~75% of circulating neutrophils are adherent to the endothelium at any time. Following its rolling interaction (Section 9.4.3.6), a leukocyte may abruptly attach to the wall of a venule and leave the bloodstream by squeezing between adjacent endothelial cells. Prior to attachment, the leukocyte does not stop rolling until certain binding events occur between the vessel wall and the white cell plasma membrane. Rapid triggering (e.g., in a few seconds) of integrin-mediated adhesion is required for the arrest of bloodborne lymphocytes and neutrophils at sites of leukocyte recruitment from the blood. This adhesion is also mediated by pertussis toxin-sensitive G$\alpha_i$ protein-linked receptors of the rhodopsin-related seven-transmembrane or serpentine chemoattractant family.

### 9.4.4.2 Nanorobot Diapedesis

Nanorobots engaging in diapedesis may release a minute quantity of interleukin-8 (IL-8) chemoattractant receptors in vivo. Particular chemokines (e.g., MIP-2p, SDF-1t, and 6-C-kine) induce flowing lymphocytes to decelerate and adhere to an integrin target protein within ~500 milliseconds. Adhesion is then followed by diapedesis, the final step in extravasation (exiting the blood vessel), given the presence of appropriate haptotactic (adhesion-gradient) or chemoattractant (0.0001-1 nanomolar$^{1512,1515}$ concentration-gradient) signals. The same $\beta_2$ and $\alpha_4$ integrins involved in lymphocyte arrest, in conjunction with other adhesion receptors, trigger the final steps in transendothelial migration. The junctions between adjacent endothelial cells in capillaries and venules are normally very narrow, with parallel plasma membranes forming an intercellular cleft of ~10-20 nm. The discontinuous-type endothelium found in the liver, spleen, and bone marrow (whose functions include the addition or extraction of whole cells from the blood) have resting gaps ~150 nm wide. In the final step of diapedesis, biochemical mediators acting on capillary endothelial cells cause those cells to loosen their attachments to their neighbors. The wall gaps in the capillaries open more widely, allowing diapedesis of white cells into the tissues. A wide range of pharmacologically active compounds (including prostaglandins, angiotensin, serotonin, histamine, epinephrine, and nicotine) have been observed to encourage the splitting of endothelial junctions along their symmetry plane, followed by cell separation and the appearance of gaps large enough to allow passage of whole cells—especially in the postcapillary venules. It also appears that the emigration of neutrophils and monocytes from the vasculature is regularly regulated by at least three distinct molecular signals—selectin-carbohydrate, chemoattractant-receptor, and integrin-Ig interactions—acting in sequence, not in parallel. This establishes a three-digit "area code" for cell localization in the body, "as if leukocytes carried cellular phones."

The sequence in which these signals act on neutrophils and lymphocytes may differ. (A neutrophil surface typically has $5 \times 10^5$ Mac-1 integrin receptors and $2-4 \times 10^4$ selectin receptors.) Properly configured medical nanorobots can undoubtedly read these "area codes" as well (Section 8.4.3).

Diapedesis typically requires ~3-10 minutes for the white cell to complete. Leukocyte emigration occurs by insertion of a pseudopod into the enlarged gap between endothelial cells, followed by amoeboid-like migration of the white cell through the blood vessel wall (Fig. 9.28). Electron microscopic studies have demonstrated that an efficient protein-tight seal is maintained between endothelial cells and a migrating leukocyte during all stages of its escape. Endothelial cells are typically 0.5 microns thick, although in postcapillary venules they may be cuboidal and much thicker.

Nanorobots engaging in diapedesis may release a minute quantity (e.g., ~0.0005 micron$^{-3}$/micron$^3$ of the RGDS tetrapeptide (a known adhesion inhibitor), $\beta_2$ integrin-mediated arrest of neutrophils can be triggered through stimulation of the formyl peptide, leukotriene B4, eotaxin, platelet activating factor, or cytoambulate through the widened
endothelial gap possibly using a metamorphic surface to maintain a protein-tight seal if necessary, then release endothelial movement inhibitors such as cytochalasin B or endothelial growth inhibitor protein (EGIP)\textsuperscript{1496} to encourage endothelial reattachment, returning the gap to its normal size. (Active chemokine concentrations are typically ~10 molecules/micron\textsuperscript{3}; Section 7.4.5.2.) The actual transit of the gap should require only milliseconds assuming ~mm/sec ambulation velocities, but a time on the order of seconds could be required for biochemically-mediated endothelial cell gap width management. Physical force may be used to spread or to close a temporary gap much faster (Section 9.4.4.3) but may be inconsistent with endothelial cell homeostasis. As already noted in Section 8.2.1.3, the lymphatic endothelium can probably open gaps at least ~22.5 microns wide,\textsuperscript{851} so micron-sized nanorobots should also find easy passage into and out of the lymphatic system.

9.4.4.2 ECM Brachiation

Taking the standard reference male human body volume of 60,000 cm\textsuperscript{3} and subtracting ~16,000 cm\textsuperscript{3} (27%) for fluids, digestive contents, and expandables (Section 8.2.5), then subtracting another ~27,000 cm\textsuperscript{3} (45%) for the volume of all tissue cells (Section 8.5.1), leaves ~17,000 cm\textsuperscript{3} (28%) which comprises the intercellular tissue volume. This volume is thoroughly penetrated by the extracellular matrix or ECM, a fibrous scaffolding that helps organize cells into tissues. The ECM contains both protein and carbohydrate components. (Matrical water is not in the form of a dilute aqueous solution but is strongly influenced by the macromolecules that are present; diffusion of small molecules is much slower in matrix water than in bulk water.)

Nanorobots with at least two appendages can alternately grasp and release a succession of adjacent ECM elements, brachiating,\textsuperscript{421} “hand over hand” through tissue in a manner crudely analogous to a scuba diver pulling himself along an underwater rope mesh ladder. Similarly, fibroblasts migrating through tissue toward the site of a wound move toward a chemotactic distress gradient by extending lamellipodia toward the stimulus while their opposite poles remain firmly bound until released (haptotaxis). ECM fibrils strongly influence the direction of migration since cultured fibroblasts tend to align and migrate along discontinuities in substrate to which they are attached (contact guidance), and only along, but not across, fibronectin fibrils.\textsuperscript{1537}

ECMs are composed of different collagen types, elastin, large glycoproteins (e.g., fibronectin, laminin, entactin, osteopontin), and proteoglycans that contain large glycosaminoglycan side chains (e.g., heparan sulfate, chondroitin sulfate, dermatan sulfate, keratan sulfate, hyaluronic acid).\textsuperscript{938,985,1511} While all ECMS share these components, the organization, form, and mechanical properties of ECMS vary widely in different tissues. For example, interstitial collagen (e.g., Types I & III) self-assemble into a three-dimensional lattice which binds fibronectin and proteoglycans. This type of ECM hydrogel forms the backbone of loose connective tissues such as dermis.\textsuperscript{985} In contrast, basement tissue membrane collagens (Types IV & V) assemble into planar arrays. When these collagenous sheets interact with fibronectin, laminin, and heparan sulfate proteoglycan, the result is a planar ECM scaffolding. The ability of tendons to resist tension, and of cartilage and bone to resist compression, similarly results from local differences in the composition and organization of the ECM.\textsuperscript{985}

Tissues are dynamic structures that exhibit continual turnover of all molecular and cellular components. The ECM helps to maintain tissue pattern integrity, allowing cells that are lost due to injury or aging to be replaced in an organized fashion. For instance, when tissue cells are killed by freezing or poisoning, all of the cellular components may die and be removed, but the basement membrane often remains intact. These residual scaffoldings ensure correct repositioning of cells (e.g., cell polarity) and proper restoration of different cell types to their correct locations (e.g., muscle cells in muscle basement tissue membrane, nerve cells in nerve sheaths, endothelium within vessels). Conversely, loss of ECM integrity during wound healing may cause permanent disorganization of tissue patterns, producing dermal scars.\textsuperscript{985} The ECM also plays a role in controlling cell function and morphology. Hepatocytes cultured upon fibronectin or laminin at low fiber surface density (1-50 ng/cm\textsuperscript{2}) exhibit differentiation, but switch to growth at high fiber surface density (1000 ng/cm\textsuperscript{2}).\textsuperscript{1502} Capillary-like tube formation by endothelial cells can be induced by purely mechanical means (simulating mechanical stresses in the natural ECM).\textsuperscript{1502} Cells cultured on substrates containing adhesive islands (mimicking spatial patterns of ECM adhesivity) change shape to match the shape of the islands.\textsuperscript{718}

The mesh size of the fibrous components of the ECM varies widely according to tissue type, but a few generalizations and specific examples can be given. For instance, the matrical function of collagen and elastin is mostly structural, while the function of fibronectin and laminin is primarily adhesive. Yet collagens are also adhesive for cells and other macromolecules, and fibronectin provides the major matrical support in clot and early granulation tissue formation (the first stages of wound repair).\textsuperscript{985}

Collagens are the major proteins of the ECM and comprise ~25% of total mammalian protein mass.\textsuperscript{936} The collagen family currently contains 19 members.\textsuperscript{971,1497,1498,3614} The collagen molecule is a long (~300 nm), thin (~1.5 nm) fiber, the center of which is a characteristic triple helix running ~95% of the length.\textsuperscript{938} Micrographs of endothelial cells cultured on a disorganized collagen-rich gel show that the cells quickly remodel the substrate into a network of collagen pads, cords, cables and bridges with grid sizes ranging from 100-1000 microns.\textsuperscript{1499} Photomicrographs of excised tissue samples show a gap of 4-12 microns between adjacent collagen fibrils in skin tissue and in tendons.\textsuperscript{935} Cartilage (which contains much collagen) has large aggregates of proteoglycans, with as many as 100 molecules attached to a single hyaluronic acid, giving a length of ~10 microns and a diameter of 500-600 nm; a tissue micrograph of fibrous cartilage in the disks between the vertebral bodies shows gaps of 10-30 microns between adjacent fibrils.\textsuperscript{938}

Fibronectin is usually formed as a disulfide-linked dimer consisting of two 220-250 kilodalton peptide subunits, each containing binding sites for collagen, for heparin, for the clotting protein fibrin, and for several other ECM molecules, along with cell surface receptors.\textsuperscript{938}

The main function of laminin (~50 nm x 70 nm cross-shaped molecule)\textsuperscript{938} is to mediate the binding of cells to Type IV collagen. Fibronectins or laminins may be present in ECM at low density (<1 ng/mm\textsuperscript{3}) or high density (>5 ng/mm\textsuperscript{3}).\textsuperscript{985} giving, from Eqn. 8.7, an estimated typical intermolecular spacing of L\textsubscript{grid} ~ 10 microns at low density and L\textsubscript{grid} ~ 4 microns at high density.

Although the matrix grid can be as narrow as 0.3-3 microns in special cases such as the ECM between the epidermis and the somites of the axolotl embryo, or between the highly flexible elastin fibers in the walls of the aorta,\textsuperscript{938} in most cases the gaps between ECM elements should be wide enough for micron-size nanorobots to slip through without much difficulty. Even quite large rigid objects should be able to negotiate the ECM. Experimentally, teflon particles up to 80 microns in diameter were injected extravascularly and
migrated to distant locations around the body including lymph nodes, lungs, and kidneys.²⁴⁶

From Eqn. 9.75, an isolated cylindrical ECM element ~1 micron long and ~0.1 micron wide could be shoved aside at a lateral velocity of ~1 mm/sec in plasma with a force of ~5 pN, easily within the strength limits of medical nanorobots having the appropriate manipulatory and motive appendages. Chemosensor pads also will allow the nanorobot to identify fiber type and tissue type (having unique chemical signatures) as the device crawls or brachiates through the ECM forest. For example, only two kinds of carbohydrate are found attached to the hydroxylsines of collagen. Elastin lacks the amino acid methionine. The family of integrins binds fibronectin, and fibronectin has lots of binding sites. Proteoglycans typically have a long core protein to which are attached up to 100 very large glycosaminoglycan (GAG) chains, with each population of core proteins having many different carbohydrates, varying in size, charge, and even composition.²³⁸ Sensoric pads can make a cell-type or tissue-type determination in as little as 2 millisecond (Section 8.5.2.2), suggesting an upper limit on nanorobot ECM brachiation of \( v_{\text{brach}} \approx 500 \text{ Hz} \) if chemopositional validation is required at every step.

What is the maximum safe nanorobot brachiation speed through the ECM? In theory, a diibrachial armswing of ~1 micron repeated at ~100 Hz allows an ECM transit speed up to 100 microns/sec through clear fluid lanes in the matrix. From Eqn. 9.73, the force required to pull a 1 micron spherical object through 310 K interstitial fluid (Table 9.4) at 100 microns/sec is ~2 pN, somewhat less than the ~10 pN likely to stimulate a mechanically-transduced cellular response (Section 9.4.3.2.1). (One study found that the force required to break individual integrin bonds to fibronectin is ~2.1 pN;¹⁰⁰⁸ another study found ~1000 cell adhesion proteoglycan molecules on each marine sponge cell surface, with pairs on adjacent surfaces having an experimentally measured adhesion of 40-400 pN per pair, depending upon \( \text{Ca}^{2+} \) concentration.¹¹⁴⁸) Assuming a pair of 1 micron long, 0.1 micron thick cylindrical nanorobot arms and applying Eqn. 9.75, an additional ~0.5 pN is required to move each arm forward, giving a total power requirement of only ~0.0003 pW for brachiation at ~100 microns/sec. For comparison, a proposal to drive a millimeter-size “seed” through human brain parenchyma using external magnetic forces, while maintaining what was believed to be an adequate margin of safety, was tested experimentally at a net tissue transit speed of ~8 microns/sec.¹₂⁵⁶,¹₂⁵⁷

Natural movement of biological entities through ECM is much slower. The ECM provides a handhold-like scaffold for fibroblast migration.³⁹⁹ Leukocyte and fibroblast amoeboid motion through extracellular tissue is typically ~0.050-0.7 microns/sec (Section 9.4.3.7), although white cell diapedesis through the blood vessel walls (~0.005 microns/sec; Section 9.4.4.1) is much slower. Chain migration of neuronal precursor cells across a restricted pathway in the brain, through a complex parenchyma between two regions separated by several millimeters, occurs at ~0.008 microns/sec.⁴⁹⁷ The migratory speed of cultured smooth muscle cells taken from coronary and peripheral arterial walls is also only ~0.006-0.015 microns/sec.¹⁵⁸ Epithelial cells migrate over a wound surface in a sheet at ~0.035 microns/sec, while wound contraction proceeds at 0.007-0.009 microns/sec.³⁵⁹ Osteoclasts tunnel through old bone, remodeling it, advancing at a rate of only ~0.0006 microns/sec.⁵³¹ However, inside cytomatrix-rich cells, membrane-bound vesicles undergo axonal transport at a typical speed ~2-4 microns/sec,³³⁸⁻³³⁹ pigment granules in chromatophores move at 2-10 microns/sec,³³⁸⁻³³⁹ and mitochondria (~micron-sized organelles) are shuttled around the cellular interior at up to ~10 microns/sec,⁴⁵³ so faster ECM transit speeds up to ~100 microns/sec are probably safe.

9.4.4.3 Intercellular Passage

With just a few notable exceptions, virtually all tissue cells lie within ~2-3 cell widths of a capillary, or ~50 microns.¹¹³,¹³¹,¹³² Thus a bloodborne nanorobot may reach any tissue cell by rapidly traveling most of the way by capillary, then exiting the capillary and crossing at most 1-2 cells to arrive at a given target cell. In cell dense tissue, it may become necessary to crawl between adhering cells.

Tissue cells are not packed so tightly that adjacent cell surfaces are in direct contact with each other. There is usually a space of at least 20 nm between the opposing plasma membranes of adjacent cells, wide enough to admit a slender nanorobot manipulator arm (Section 9.3.1). This space is filled with extracellular fluid and provides a pathway for substances to pass between cells on their way to and from the blood in nearby capillaries. Most tissues employ desmosome or gap junctions²⁹² wherein the opposing plasma membranes come within 2-4 nm of each other over a space of ~5 nm on the membrane surface, with these junctions spaced ~20 nm apart across the plasma membranes and opposing plasma membranes, retaining a fluid-filled 25-35 nm gap between them. Tight or “occluding” junctions (as in the endothelial blood-brain barrier and certain epithelial surfaces comprised of intestinal cells, bladder cells, and some exocrine cells) have <2.5 nm wide intermembrane gaps.³⁶¹ None of these spaces is wide enough to allow passage of whole (even metamorphic) medical nanorobots. To pass between cells in cell-rich tissue, it is necessary for an advancing nanorobot to disrupt some minimum number of cell-to-cell adhesive contacts that lie ahead in its path. After that, and with the objective of minimizing biointrusiveness, the nanorobot must reseal those adhesive contacts in its wake, after passage, crudely analogous to a burrowing mole.

A full treatment of all types of cell-to-cell contacts and anchoring mechanisms is beyond the scope of this book, but a few specific examples can be given. For instance, spot desmosome (Section 5.4) are ~30-nm-long “spot weld” molecules linking neighboring cells, spaced ~8 nm apart across the apposed cellular surfaces. A force of 6-10 pN is required to separate each connexion-32 hepatic cell gap junction unit (Section 5.4.2). Integrins may average ~20 nm separation, and require ~2.1 pN to separate from a fibronectin molecule in the ECM,¹⁵⁰⁸ although direct cell-ECM connections are relatively scarce inside cell-cell junctions.

As a crude estimate, assume that intercellular adhesion molecules \( h_{\text{adhes}} \sim 30 \text{ nm in length are spaced } x_{\text{adhes}} \sim 10 \text{ nm apart across the cell surfaces and require } F_{\text{adhes}} \sim 10 \text{ pN to pull apart,}¹²²³ \) and that an \( l_{\text{nano}} \sim 1 \text{ micron wide nanorobot wishes to pass through by dynamically clearing a path ahead that is } l_{\text{nano}} \text{ wide and } l_{\text{nano}} \text{ long. Along the leading edge, } l_{\text{nano}} / x_{\text{adhes}} \text{ adhesive joints must be detached one by one, requiring a detachment energy of } E_{\text{detach}} \sim l_{\text{nano}} F_{\text{adhes}} h_{\text{adhes}} / x_{\text{adhes}} \sim 30,000 \text{ zJ to advance a distance } x_{\text{adhes}}. \) This gives a power requirement of \( P_{\text{travel}} \sim E_{\text{detach}} v_{\text{travel}} / x_{\text{adhes}} \sim 0.003 \text{ pW} \) if the velocity of forward travel through the cell-cell junction is \( v_{\text{travel}} \sim 1 \text{ micron/sec, allowing transit between two adhered } 20-\text{micron cells in } 20 \text{ sec. Adding manipulator arm energy dissipation raises the total to } \sim 0.01 \text{ pW (viscous drag power is } \sim 10^{-5} \text{ pW; Eqns. 9.74 and 9.75). Total detachment rate is } v_{\text{detach}} \sim l_{\text{nano}} v_{\text{travel}} / x_{\text{adhes}} \sim 10,000 \text{ detachments/sec, a burden which can be shared by more than one manipulatory appendage, each equipped with appropriate lytic end-effectors (e.g., trypsin-like tool tips; Section 9.3.2). Rejoining the parted cell junctions astern requires, at worst, a similar energy expenditure; in the case of noncovalent bonds, unassisted
rejoining may be energetically favorable if the edges have been left in relative proximity. Reattachment may be facilitated using a metamorphic nanorobot integument, together with a surfacial hydrophilic solvation wave drive system (Section 9.4.5.3), to establish a teardrop-shaped cross-section that can slowly adduct the separated plasma membrane faces and maneuver the detached stubs of parted anchor macromolecules into close proximity. Traffic density and frequency should be locally restricted to minimize mechanical stimulation of unwanted ECM and cytomatrical responses (Sections 9.4.3.2.1 and 9.4.4.2), and especially detachment-triggered apoptosis (Section 10.4.1.1), perhaps by restricting intercellular passage to nanorobots traveling single-file through a relatively small number of channels.

9.4.4.4 Nanorobot Conjugation and Partition

Occasionally it may be necessary for two or more motile nanorobots to dock, or conjugate, during histonatation (e.g., Section 8.3.3). Assuming the existence of in vivo communication and navigation networks (Sections 7.3 and 8.3), two nanorobots that are attempting to dock can exchange positional information until the two devices are within a range of ~3 microns. Once within such close range, the devices switch to ~GHz chirps which, from Eqn. 4.52, attenuate ~0.001%/nm in 310 K plasma. Using an acoustic such close range, the devices switch to ~GHz chirps which, from the two devices are within a range of -3 microns. Once within such close range, the devices switch to ~GHz chirps which, from Eqn. 4.52, attenuate ~0.001%/nm in 310 K plasma. Using an acoustic receiver sensitive to $10^{-6}$ atm pressure changes, and 0.01-atm transmitted pulses, and assuming that no other nanorobots are present and chirping in the immediate vicinity at the same time, the two nanorobots can locomote up the attenuation gradient to within ~10 nm proximity, approximately the minimum scale of the docking mechanism (e.g., metamorphic bumpers; Section 5.4). Conjugating microbes employ chemical communications to similar effect.3628

Once physical contact is made, nanomechanical communication at ~1 GHz transfers ~1000 bits during a ~1 microsec contact event (Section 7.2.4), sufficient to mutually validate identity. A universal clamp may be used for unrestricted docking among devices of any machine species that meet in any orientation. Device shape or surface features can be used to control docking orientation (Section 5.4.1). Docking clamps can also be designed to connect in only one way if the spatial or rotational positioning of the conjugated pair is an important mission requirement; from Eqn. 3.1, Brownian motion provides ~4000 random docking trials per second for two 1-micron nanorobots maintaining a mean 10-nm separation in 310 K interstitial fluid. Docking clamps may serve a purely adhesive function, or may additionally provide communicating junctions that transfer fluids, power, or data between conjugated devices. Similar procedures permit three or more nanorobots to conjugate in vivo, and may allow a growing nanorobot aggregate to sequentially add new members to the collective at specific locations in the developing structure.

Partition, or separation of linked nanorobots, requires a defined procedure in which all transtegumental openings in nanorobots are sealed or valved shut prior to release of docking clamps and physical separation. Once separated, and depending upon their mode of locomotion and the timing of their release, partitioned nanorobots must temporarily observe a dense-traffic protocol involving slower and lower-amplitude motions of motive appendages or rotating elements in order to:

1. avoid mechanical conflict among such appendages or elements between neighbors;
2. avoid potentially chaotic hydrodynamic interactions (Section 9.4.2.3); and
3. clear a sufficient interdevice distance to preclude significant error in positional readings from the navigation grid (Section 8.3.3), due to spurious acoustic refraction and attenuation effects if there is a large number density of nanorobots in the vicinity.

9.4.5 Cytopenetration

Upon arriving at a target tissue cell, a cellular repair nanorobot may require entry into the cytosol. The following is a necessarily brief and incomplete discussion of several techniques and issues involved in penetrating the plasma membrane, and by extension, in penetrating other intracellular membrane structures including the doubled-walled nuclear envelope and the multilayer outer membranes of Gram-negative bacteria (Section 10.4.2.5). Micron-size siliceous diatoms and “nanobacteria” with calcium shells and will require additional penetrative mechanisms that are not described here. This discussion also is applicable to the insertion of sensory appendages or manipulatory devices through cell or organelle lipid membranes.

9.4.5.1 Transmembrane Brachiation

One simple approach to cytopenetration is to insert a narrow manipulator arm through the plasma membrane, perhaps assisted by a wall-breaching tool tip analogous to the T4 lysozyme enzyme that opens a hole in bacterial cell peptidoglycan walls or the T4 DNA-injection system that is specifically designed for lipid bilayer penetration. The terminus of the manipulator firmly attaches to an actin or microtubular fixed component of the interior cytoskeleton, then retracts, towing the nanorobot through the lipid bilayer surface. Metamorphic reshaping of the nanorobot aspect minimizes the total number of noncovalent lipid bilayer bonds that must be disturbed during transit. Cell surfaces tolerate forces up to ~1 nN per 100 nm$^2$ before breaking, so a ~1 nN arm tip smaller than 100 nm$^2$ in cross section should be able to push itself through the cellular plasma membrane and into the interior of the cell. As a very crude estimate of the penetration energy required, taking the membranolytic limit as 3 x $10^6$ N/m$^2$, then a 1-micron long, 1-nm...
wide tear can be opened up by applying ~3000 pN of force through a distance of ~10 nm (typical plasma membrane thickness) costing ~30,000 zJ of energy, a ~0.003 pW power demand during a ~10 millisecond transit time at a ~100 micron/sec transit speed. It may also be possible to reduce entry forces by inducing lipid bilayer membrane demixing using a tangentially-applied electric field.\textsuperscript{1613}

9.4.5.2 Metamorphic Screw Drive

A nanorobot with a hydrophobic (e.g., diamondoid) metamorphic surface may locally reconfigure its prow in the shape of a screw having two traveling raised ridges, with a thread width at least as wide as the plasma membrane plus the plasma membrane skeleton or cell cortex, perhaps 30-50 nm thick (Section 8.5.3.11). After the screw has initially penetrated the surface and the first thread course is locked in place (Section 9.4.5.1), the thread pattern may then be translated toward the rear of the nanorobot, allowing the device to screw itself into the cell's interior (Fig. 9.29). A ~100 micron/sec transit speed requires ~0.001 pW for a 1 micron\textsuperscript{2} cross-section nanorobot to overcome viscous drag (Section 9.4.2.4) and ~0.01 pW to operate a 100-nm wide section of metamorphic surface assuming ~0.35 micron\textsuperscript{2} of surface dissipating ~0.03 pW/micron\textsuperscript{2} (Sections 5.3.1.4 and 5.3.6). Cell entry is completed in ~10 milliseconds.

9.4.5.3 Solvation Wave Drive

The lipid bilayer membrane is hydrophilic on the outermost and innermost surfaces, and lipophilic in the interior (Fig. 8.37). Much like the physical screw described in the previous Section, the hydrophilicity of the nanorobot exterior surface may also be manipulated to produce a traveling helical solvation wave that establishes temporary noncovalent bonds with elements of the plasma membrane, allowing the nanorobot to pull itself through the lipid bilayer (Fig. 9.30). The plasma membrane is typically 6-10 nm thick (Section 8.5.3.2). The cholesterol-poor lipid membranes of the mitochondria and the endoplasmic reticulum have a 2.5 nm inner hydrophobic region, while cholesterol-rich lipid membranes like those of the plasma membrane and the endosomes have a 3.1 nm wide hydrophobic region.\textsuperscript{1113-1115} Each contact point may be envisioned as a semaphore-like mechanism (Section 5.3.6) by which lipophilic or hydrophilic moieties are rotated into an exposed position (facilitating noncovalent bonding with a particular lipid bilayer phase), then moved a short distance relative to the nanorobot body (transmitting force to the nanorobot), then rotated back into a nonexposed position (breaking the bond to the lipid membrane).

From Eqn. 9.73, a 1-micron nanorobot traveling at ~100 microns/sec must generate ~2 pN of motive force in order to overcome the viscous resistance of plasmalike fluid at 310 K. Given a single-lipid extraction force of $F_{\text{lipid}}$ ~ 1 pN (Section 9.4.3.1), ten simultaneous contacts producing a total of 2 pN reduces applied force per contact to a reasonably safe ~20% $F_{\text{lipid}}$. Assuming 10 simultaneous contacts spaced around the nanorobot perimeter and 10 nm of available longitudinal travel for each bonded semaphore mechanism, a 1-micron nanorobot requires ~1000 surface semaphores in staggered configuration. Allowing ~300 nm\textsuperscript{2} per semaphore presentation face (Section 5.3.6), the entire solvation wave drive system occupies just 5% of nanorobot surface area. Breaking ~1000 noncovalent semaphore-membrane bonds, assuming ~100 zJ/bond (Section 3.5.2), during a 10 millisecond transit requires ~0.01 pW of power; the same figure is obtained for 300,000 nm\textsuperscript{2} of semaphores dissipating ~0.03 pW/micron\textsuperscript{2} (Section 5.3.6). Total power required to overcome viscous resistance is ~0.0002 pW during the ~10 millisecond transit time (Section 9.4.2.4).

9.4.5.4 Vesicle Fusion and Endocytotic Entry

A nanorobot may also gain entry to a cell by releasing amphipathic lipid molecules from its surface, which will self-assemble into a thin lipid bilayer coat, enveloping the entire nanorobot. Natural viruses\textsuperscript{1721} and synthetic carriers\textsuperscript{1722} use similar methods to convey extracellular DNA across outer cell plasma membranes and nuclear membranes. The materials comprising a 10-nm thick lipid bilayer surrounding a spherical object ~1 micron in radius can be stored in a ~0.1 micron\textsuperscript{3} internal tank, occupying only ~3% of total internal device volume. The fusion of two distinct lipid bilayers is energetically unfavorable in the absence of specialized proteins.\textsuperscript{1536} Thus after approaching the target plasma membrane, the enveloped nanorobot must emit specialized plasma membrane fusion proteins, known as fusogens,\textsuperscript{1587,3658} similar to the sperm protein PH-30\textsuperscript{α−β} that allows sperm and egg to merge,\textsuperscript{1082} the HIV virus envelope protein gp120-gp41 with a typical contact area of ~8 nm\textsuperscript{2},\textsuperscript{1531,1587} the envelope glycoprotein H employed by the herpesviruses,\textsuperscript{3657} and...
the hemagglutinin protein found on the surface of influenza virus.\textsuperscript{532} (Nonprotein fusogens also are well-known, including lipids such as N-acyl phosphatidylethanolamines (NAPEs),\textsuperscript{3659} fatty acids such as arachidonic acid,\textsuperscript{3660} carbohydrates such as polyethylene glycol,\textsuperscript{3661} and simple solvents such as dimethyl sulfoxide (DMSO).\textsuperscript{3662}) The nanorobot lipid bilayer, now fusion active, can join with the plasma membrane layer and unfold, bringing the formerly fully enclosed nanorobot into direct sealed contact with the cytosol with a negligible energy expenditure (Fig. 9.31). This process will add 0.5%-1.0% new lipid, by volume, to the existing -14-24 micron\textsuperscript{3} plasma membrane volume (Table 8.17) of a 20-micron tissue cell. Added lipids that don’t exactly match the native lipids of the cell might disrupt the normal protein/lipid membrane ratio and modify membrane fluidity, possibly altering signal timing and other cell behaviors. If this is a problem, nanorobots employing this entry technique could be designed to reabsorb and recycle the foreign lipid, once inside the cell.

Other means of cell entry, such as “induced uptake,” are well known. Like bacterial invaders, a nanorobot may emit cytochemical “entry mediator” signals (e.g., the outer wall membrane protein invasin, in \textit{Yersinia} bacteria) to induce the target cell to form an endocytotic vacuole around the device and suck it into the interior of the cell. Because nanorobots need to enter cells fast, a fusion-type tool may subsequently be used to escape from the resulting endosome. Macrophages ingest up to -25% of their volume per hour,\textsuperscript{996} a volume equivalent to hundreds of micron-sized nanorobots. Many bacteria\textsuperscript{1012,1561} and viruses\textsuperscript{530,1553} gain entry to cells by this means, and phagocytosis is mechanically similar. Both are initiated by ligand-receptor interactions that activate host signaling, with the actin cytoskeleton providing the necessary force to internalize the particle into a membrane-bound vacuole.\textsuperscript{1012} Invasive bacteria utilize two major types of induced uptake:

1. a “zipper” type mechanism involving direct contact between bacterial ligands and cellular receptors that sequentially encircle the organism (e.g., \textit{Yersinia}, \textit{Listeria}); and;

2. a “trigger” mechanism in which bacterial signals to the cell induce dramatic plasma membrane ruffling and cytoskeletal rearrangements resulting in macropinocytosis with virtually passive entry of the bacterium (e.g., \textit{Salmonella}, \textit{Shigella}) into the cell.\textsuperscript{1012,1561,3191}

### 9.4.5.5 Cytosolic Leakage During Transit

If the nanorobot-membrane interface is not tight, some leakage of material into or out of the cell (depending upon relative hydraulic and osmotic pressures) may occur during nanodevice transit. For example, from Eqn. 3.1 the Brownian diffusion distance during a transit of 10 millisecond time is ~3 microns for small molecules like amino acids (MW - 100 daltons), -1 micron for large protein molecules (MW - 100,000 daltons). An annular leakage ring of width h\textsubscript{leak} - 10 nm around the circumference of a nanorobot of radius R\textsubscript{nano} - 1 micron may pass V\textsubscript{leak} ~ 2 π R\textsubscript{nano} h\textsubscript{leak} V\textsubscript{transit} ~ 0.006 micron\textsuperscript{3} of cytosolic fluid during each transit (~0.00008% of cell volume), assuming V\textsubscript{transit} - 10 microns/sec for random hydrodynamic flows induced by thermal fluctuations inside a cell (Section 8.5.3.12). Alternatively, modeling the leak as a ring of N\textsubscript{pipe} = 2 π R\textsubscript{nano} / r\textsubscript{tube} = 1257 nanopipes each of radius r\textsubscript{tube} = h\textsubscript{leak}/2 and length l\textsubscript{tube} - 10 nm, and taking Δp - 0.001 atm for the interstitial/cellulular pressure differential (Section 4.9.1.2), then from Eqn. 9.25 the total leakage volume from the ring of nanopipes is V\textsubscript{leak} ~ N\textsubscript{pipe} V\textsubscript{tip} V\textsubscript{transit} ~ 0.03 micron\textsuperscript{3} of cytosolic fluid during each transit, or ~0.0004% of cell volume. Leakage may be reduced by encouraging a tighter seal during transit using a dynamic ring pattern of lipophilic semaphores at the nanorobot surface (Section 9.4.5.3) that tracks the device’s passage through the plasma membrane. In some cases such as nerve cells, significant ionic leakage could produce depolarization and thus should be avoided.

### 9.4.5.6 Breach Sealing and Intrusiveness

Some cytopenetration techniques involve tearing the lipid bilayer. This breach must be resealed after transit. A fractured lipid membrane presents two “greasy” surfaces to water; it is thermodynamically favorable for these two cut surfaces to fuse together so as to eliminate the unfavorable water-fat interface.\textsuperscript{72,2316} Such fusion may be encouraged mechanically by the passing nanorobot (Section 9.4.4.3), although the edges of the tear will attempt to round off and self-seal, without rejoining.

Nevertheless, it is not uncommon to observe rapid natural resealing of plasma membranes with little loss of intracellular contents, a useful property given the many transient plasma membrane disruptions that commonly occur in cells that experience significant mechanical stress, such as gut, skin, endothelium, and muscle.\textsuperscript{1534,3665} In one experiment, tissue cell plasma membranes were punctured using 2-3 micron diameter micropipettes and a 300 millisecond transit (wounding) time, and the torn plasma membrane spontaneously resealed in 10-30 sec with relatively little visible loss of injected dye.\textsuperscript{1534} Exocytosis-based resealing\textsuperscript{3666-3668} of a microneedle puncture through the fibroblast plasma membrane occurs in 5-10 sec,\textsuperscript{3667} but a second puncture at the same site heals faster than the initial wound\textsuperscript{3668}—at first wounding, the cell uses existing endocytotic compartment to add membrane necessary for resealing, but Ca++ entry at the first wound stimulates vesicle formation from the Golgi apparatus, resulting in more rapid resealing of the second membrane disruption.\textsuperscript{3668}

Watertight breach sealing might even be possible for nanorobots trailing narrow, untensioned tethers with lipophilic coatings, although such tethers may cause other problems (Sections 6.4.3.6 and 7.3.3): 0.1-micron optical fiber tips have been poked through a cellular plasma membrane to measure the pH of the cytoplasm inside, in single cells and in single rat embryos, without ill effect on these large cells.\textsuperscript{577} In small cells (2-15 micron diameter), “stab” micrinjection at high pressure (0.1-0.2 atm) is problematic because the nucleus-to-cytoplasm ratio is higher so the nucleus is more likely to be damaged during the stab. In one experiment, less than 5% of neutrophils survived the high-pressure stab intact, but a low-pressure (~0.01 atm) injection through a lipid bridge produced a ~100% survival rate.\textsuperscript{2346}

How many nanorobots (or other pieces of exogenous matter) can safely squeeze into a living cell? This question is addressed further in Chapter 15, but an extremely conservative volumetric injection limit is ~50-100 micron\textsuperscript{3} per tissue cell (0.5-1% of typical cell volume) without any significant observed effect on cell viability\textsuperscript{1192}—equivalent to ~3-100 bloodborne nanorobots, depending upon object size. Individual lymphocytes (~200 micron\textsuperscript{3})\textsuperscript{677} have also been observed circulating for hours inside large living cells, with no evident ill effect (Section 8.5.3.12).

In some nanorobotic system configurations and applications, it may be useful to deploy smaller intracellular probes tethered to larger extracellular command centers. For example, axonal “cleanout” could be an important anti-aging activity, given the evidence that axonal transport slows with age presumably due to debris accumulation or a decline in energy supplies; yet full-size nanorobots might not conveniently fit inside some neural axons. Transmembrane tethers

---

**Basic Capabilities • Manipulation and Locomotion** 323
(tipped with relatively small unfurlable mechanisms) may be passed into the cell’s interior by threading them through an anchored transmembrane sleeve device. Fluid leakage between tether and inner sleeve wall, either into or out of the cytosol, can be minimized if both surfaces are hydrophobic. There are, of course, many important drawbacks to the use of tethered systems generally (Sections 6.4.3.6 and 7.3.3).

9.4.5.7 Nuclear Membrane Penetration
Most of the methods proposed for plasma membrane penetration can also be applied to the nuclear membrane, with several important differences. First, the entry procedure should proceed at a much slower velocity, to preclude any possibility of damaging the chromatin (Section 8.5.4.7). Second, chromatin is attached to the nuclear cortex away from the nuclear pores, and also in specific chromosomal territories, so special care must be taken to avoid disturbing these configurations. Third, the region to be traversed is much deeper—two lipid bilayer membranes separated by a perinuclear space, and lined on the nucleoplasmic side by a nuclear cortex of varying width (Fig. 8.46)—which may require a somewhat different design than systems used for single-membrane bilayer penetration.

9.4.6 In Cytolocomotion
Once inside a cell, a motile nanorobot must navigate a cluttered and highly viscous cytomatrix-rich environment (Section 8.5.3). (When microscopic iron particles are introduced into the cytoplasm, they move erratically (rather than smoothly) under the influence of external magnetic fields.) Any of the methods described for cytoamputation (Section 9.4.3), histonation (Section 9.4.4), or cytopenetration (Section 9.4.5) may also be employed in modified form within the cell. From Eqn. 9.65 and the data in Table 9.4, and assuming a top speed of $v \approx 10$ microns/sec (Section 9.4.4.2), then the Reynolds number of an $L \approx 1$ micron object inside a red blood cell (a nucleus-free floppy bag filled with hemoglobin solution) is $N_R \approx 10^{-6}$ For a nanorobot inside a free leukocyte, $N_R \approx 10^{-9}$; inside an E. coli bacterium, $N_R \approx 10^{-11}$, taking the higher viscosity into account.

There are many natural transport mechanisms inside cytomatrix-rich cells which may serve as analogs for in cyto nanorobot locomotion. For example, vesicles and granules -100 nm in diameter or larger are carried at a peak speed of up to ~2 microns/sec (although mean unloaded kinesin motor speed is usually 0.5-0.8 microns/sec) on the back of a 60-nm kinesin transport molecule (Fig. 9.32) that takes 8-nm ATP-powered steps along microtubule tracks running throughout the cell with a stall force of 5-7 pN; typically kinesin takes ~100 steps along a microtubule, then lets go. Nanorobots could be designed to brachiate along these tracks as well. Inside giant amoebas, mitochondria measuring 1-3 microns in length are carried along microtubules at speeds up to 10 microns/sec, and the pseudopods of fibroblasts (and amoeobas) can also extend at ~10 microns/sec. Faster speeds may be possible for well-designed nanorobots, but an upper limit of ~10 microns/sec through fluid-rich intracellular clear channels seems reasonably conservative.

It is well-known that several pathogenic bacterial species, once free in the cytoplasm of a human cell, propel themselves through the cytosol using a continuous actin polymerization process that takes place at one pole of the bacterium. Actin assembly is visible as a tail of polymerized F-actin that remains stationary in the cytosol while the bacterium moves ahead. The polymerizing tail rectifies the random thermal motions of the bacterium, preventing it from diffusing backwards while permitting forward diffusion; thus the tail doesn’t actually “push” the bacterium forward. Actin-based motility is mediated by a single bacterial protein—ActA (610 amino acids) in Listeria and IcsA/VirG (120,000 daltons) in Shigella—localized in the polar regions of the bacterium. Actin microfilaments (and tubulin) typically self-assemble at ~0.1-1 micron/sec; actin polymerization gives a stall force of ~10 pN per fiber, sufficient to drive a 1-micron object at ~1.5 micron/sec against a ~1 pN load in free fluid cytoplasm. This probably defines the top speed for actin-based bacterial mobility.

To progress through dense cytomatrix regions, motile nanorobots of similar size to bacteria must cut or detach cross-bridged cytoskeletal elements lying across the path ahead, then attempt to reattach or reconstruct those elements after the nanorobot has passed through the breach, because these elements lack sufficient elasticity (and grid sizes are too small) to be pushed completely out of the way. Such cytoskeletal elements will most commonly include intermediate filaments and actin-based microfilaments. This procedure is crudely analogous to the process employed by fibroblasts transiting the ECM during wound repair. Fibroblast movement into cross-linked fibrin blood clots or tightly woven ECM requires an active proteolytic system that cleaves a pathway for migration; known enzymes serving this purpose include plasminogen activator, interstitial collagenase (MMP-1), the 72 kilodalton gelatinase A (MMP-2), and stromelysin (MMP-3).

For an average filament grid size of $L_{grid} \approx 100$ nm for the cortical actin cytogel and a nanorobot of radius $R_{nano} \approx 1$ micron, a minimum of two long diagonal cuts each of length $2R_{nano}$ (simplistically, making four triangular flaps) to allow passage requires $N_{cut} = 4R_{nano}/L_{grid} = 40$ transected grid segments in order to advance a distance $L_{grid}$ or $v_{cut} = 4R_{nano}\nu_{nano}/L_{grid}^2 = 400$ cuts (or reattachments) per second at $v_{nano} \approx 1$ micron/sec. This compares favorably...
with the $v_{\text{brach}} \approx 500$ Hz estimated earlier for a manipulator arm used for ECM brachiation, where chemopositional validation is required at every step (Section 9.4.4.2), and is well below the $v_{\text{detach}} \approx 10$ KHz estimated for intercellular passage where such validation is not required (Section 9.4.4.3). An important additional design issue for in cyto brachiation systems is to minimize accidental mechanical signal transduction into the nucleus, which may trigger unwanted cytological responses.

Power requirements for intracellular mobility are typically modest but vary widely, depending mainly upon velocity, the type of cell, and the path chosen. In highly fluidic clear channels, viscosity should more closely resemble that of the red cell interior ($\eta \approx 10^{-2}$ kg/m-sec), so from Eqn. 9.74 and taking $R_{\text{nano}} = 1$ micron and $v_{\text{nano}} = 10$ microns/sec, then $P_{\text{nano}}/e\% = (0.00002 \text{ pW})/e\%$ for pure cytonatation. Nanorobots traversing pathways through filament-rich regions of the cell will experience much higher effective viscosities, on the order of $\eta \approx 10-1000$ kg/m-sec (Table 9.4), and will also travel more slowly (e.g., $v_{\text{nano}} \approx 1$ micron/sec), giving a viscous power requirement for the nanorobot body of $P_{\text{nano}}/e\% = (0.0002-0.02 \text{ pW})/e\%$, from Eqn. 9.74. A telescoping manipulator arm measuring 500 nm in length, performing the filament cuts and joins at $v_{\text{cut}} = v_{\text{join}} \approx 400$ Hz with ~2.5 microns of tool-tip travel per cycle, moves at a tip speed of ~1 mm/sec, producing ~0.025 pW of mechanical losses under no-load conditions (Section 9.3.1.4); there are at least two operational arms, one fore and one (or more) aft. Multiple arms may be needed in a dense filament network. Another ~0.025 pW is required to overcome the viscous resistance against each moving arm (using Eqn. 9.75), giving a total power cost of 0.1-0.3 pW for transfilamentary intracellular locomotion at $v_{\text{nano}} \approx 1$ micron/sec assuming locomotice efficiency $e\% = 0.10$ (10%).

Some of the difficulties and limitations of in nucleo locomotion have already been mentioned in Sections 8.5.4.7 and 9.4.5.7. The most important constraint is to observe a maximum speed limit that avoids mechanical chromatin damage. Although a 500 kilodalton protein would diffuse from one side of a 5-8 micron nucleus to the other in only ~5-6 sec (~1 micron/sec average), normal chromosomal movements during interphase have been estimated by observing the motions of individual centromeres, typically 0.002-0.003 microns/sec, and peak chromosome transport speeds during mitosis are ~0.1 micron/sec. A strict in nucleo speed limit of ~0.1 micron/sec on all nanorobot bodies and most exposed appendages appears prudent. A force limit of ~50 pN for intranuclear locomotion should also be observed (Section 8.5.4.7) — a 1-micron nanorobot picking its way through a filament-rich medium of net viscosity ~10 kg/m-sec requires the application of ~20 pN to overcome viscous drag at a travel speed of ~0.1 micron/sec.

### 9.4.7 Cytocarriage

The commandeering of natural motile cells by medical nanorobots, known as cytocarriage, offers an alternative mode of in vivo transport. During cytocarriage, one or more medical nanorobots may enter a motile cell, ride or steer the cell to a desired destination inside the human body, then vacate the cell upon arrival. The discussion here is a necessarily brief overview of the objectives of cytocarriage (Section 9.4.7.1), cytovehicle selection (Section 9.4.7.2), initiating cytocarriage (Section 9.4.7.3), steering and control during the journey (Section 9.4.7.4), cytocarriage navigation and sensing (Section 9.4.7.5) and cytovehicular behavioral control (Section 9.4.7.6).

#### 9.4.7.1 Objectives of Cytocarriage

One of the most important utilities of cytocarriage is the "cell herding" function. Nanorobots can marshall significant resources drawn from the body's natural immune and wound repair systems, and redirect them to particular sites that are in urgent need of assistance. In this role, medical nanorobots may act as immune system or wound repair homeostatic accelerators, regardless of the nanorobots' own additional inherent reparative functions, which may be considerable. Nanorobots may also seize control of pathogen and parasite mobility systems, then guide them to natural disposal sites within the body.

While nanorobot control of the natural motility apparatus might slightly improve its performance, as a practical matter the maximum speed of cytovehicles being driven through tissues is probably limited to near the maximum observed natural speeds of in vivo progression of uncontrolled cells (e.g., ~1-10 microns/sec). This will normally be considerably slower than the maximum transfurcellular speeds available to self-propelled medical nanodevices (e.g., 100-1000 microns/sec). Also, cytocarriage is usually energetically inefficient compared to self-propelled medical nanorobots. However, in some circumstances a strategy of cytocarriage may avoid the need to include certain propulsive mechanisms in the nanorobot design for some segments of the mission pathway, thus reducing nanodevice design complexity and freeing up scarce onboard storage volume for additional consumables or mission-critical machinery. Cytovehicles (e.g., leukocytes) engineered to display organ-specific tissue homing could be preloaded with passenger nanorobots prior to injection into the patient. Once injected, the nanorobots would be delivered to the intended destination with acceptable reliability. Of course, surface-placed antigen semaphores (Section 5.3.6) may directly confer a similar homing ability on bloodborne nanorobots, with equal or superior reliability.

Due to the small size of most readily available cytovehicles, conservatively only one or at most a few nanorobots may be able to safely enter and occupy the interior of each commandeered motile cell without triggering unwanted cellular responses (Chapter 15). However, additional nanorobot passengers may adhere to the exterior of the plasma membrane of the cytovehicle (after disabling certain natural haptotactic reactivities), or may be towed by the motile cell, allowing the motile cell to serve as a "pack animal" or "biological tractor." Interestingly, protozoa called mixotrichs that are present in the termite gut are propelled by several thousand bacterial spirochetes attached to the protozoan as obligate symbiotes, S. Vogel observes that "the protozoa have adopted bacteria as engines the way a human might use a team of horses." Even heavily burdened cytovehicles operate at very low Reynolds numbers (Section 9.4.2.1), so the total mass of passengers ($M_{\text{pass}}$ determining the inertial load) is relatively unimportant compared to the total surface area of the passengers ($A_{\text{pass}}$ determining the viscous load); $M_{\text{pass}} - N_{\text{pass}}$ (the number of passengers per cell), but $A_{\text{pass}} - N_{\text{pass}}^{1/3}$ for efficiently packed passengers. Giving a towing strength of 165 nN for the human fibroblast, a properly harnessed single fibroblast cell traversing interstitial fluid at a speed of 10 microns/sec can in theory tow behind it a 20-micron wide, 1000-micron long cylindrically-packed aggregate of $N_{\text{pass}} = 300,000$ 1-micron$^3$ nanorobot passengers with a towing force of only $F_{\text{nano}} = 0.01$ nN (Eqn. 9.75). (A fibroblast pulling 165 nN at 10 microns/sec consumes ~2 pW of power.)

#### 9.4.7.2 Cytovehicle Selection

Which motile human cells may be the most appropriate cytovehicles? Since almost all cells exhibit migratory behavior at the...
embryonic stages of development, in theory all cells can probably be mobilized, given access to the nuclear chromatin and the ability to manipulate it (Chapter 20). However, it is a safer and simpler matter to choose cells which are exclusively or at least facultatively motile in the human adult. The obvious candidates include white cells (e.g., neutrophils, monocytes, dendritic cells, eosinophils, basophils), lymphocytes (up to ~0.5 microns/sec), macrophages, and fibroblasts—the speediest and most widely distributed cytovehicles—as well as Kupffer cells, osteoblasts and osteoclasts, endothelial cells, smooth muscle cells, sperm cells, malignant tumor cells (after safing), neuronal cells, and some migratory epithelial (e.g., Langerhans’ cells) and mesenchymal cells. Motions analogous to crawling transform blood platelets from smooth disks into spiny spheres to plug vascular leaks after injury. Most bacteria are too small for whole-device entry, although their motility could be controlled by spoofing the external chemosensors, perhaps making them useful for piggybacking or towing (Section 9.4.7.1), or for steering to disposal sites. Other possibilities include parasites such as Entamoeba histolytica (found in the gut during amoebic dysentery, the liver during hepatic amebiasis, the bloodstream, and elsewhere), Giardia lamblia (an intestinal parasite), Plasmodium (a bloodstream parasite producing malaria), and so forth.

Erythrocytes are a special case. Though not motile, red blood cells are plentiful in the body, flow to within ~50 microns of virtually every tissue cell in the body, and have a tough, floppy outer plasma membrane. Experimentalists often replace the hemoglobin solution of red cells with an isotonic aqueous solution, producing “ghost cells” that mechanically behave much like the original cells minus the respiratory functions. In theory, nanorobots could enter an erythrocyte ghost cell and erect an artificial internal cytoskeleton to control cell shape, rotation, and possibly simulate amoeboid motions. However, naturally motile cells such as white cells already possess hardwired programs for creeping, diapedesis, object engulfment, etc. that can be triggered mechanically or biochemically with a minimum of effort (Section 9.4.7.6), hence these cells are much more attractive cytovehicles.

Passive cytocarriage is widely practiced in the microscopic world. Dengue-2 virus enters macrophages and B lymphocytes, hitching a ride;3082 bacteria stably associated with human cells (e.g., E. coli in the intestinal mucosa) are well known. Larger microbes such as the tuberculosis bacterium also employ opportunistic cytocarriage. These bacteria are specially adapted to recognize and lock on to the large macrophages distributed throughout pulmonary tissue, and then to induce these cells to ingest them. Once inside the macrophages, the microbes get a free ride into the blood or the lymphatic system, enabling them to reach destinations all over the human body.384

9.4.7.3 Cytocarriage Initiation

There are ~7000 leukocytes/mm³ in adult human blood (Appendix B) with a mean separation of ~50 microns, providing numerous and readily-available cytovehicles. A white cell that is cytoambulating at ~1 micron/sec moves only ~0.02 micron during the ~2 millisecond required to establish identity (Section 8.5.2.2), and another ~0.1 micron during the ~10 millisecond required for the piloting nanorobot to penetrate the plasma membrane (Section 9.4.5). Nevertheless, it may be desirable to temporarily halt the progression of the cytovehicle during the period in which nanorobot control is being established. This is readily accomplished by biochemical means. For example, when leukocytes are placed in an isotonic solution containing EDTA, active spontaneous movement stops and the cells respond passively to external loads. Excess zinc immobilizes macrophages. Integrins respond to intracellular messages, altering both selectivity and binding strength: the αβ integrin can be inactive, a receptor for collagen, or a receptor for both collagen and laminin, depending on the type of cell that produces it and the signals received from inside the cell. Once inside a 10-micron diameter leukocyte, a nanorobot traveling at ~1 micron/sec may require ~5 sec to migrate to an appropriate location near the center of the cell and establish control. Leukocyte towing harnesses may be quite simple in design but must be non-destructive. A required towing force of 0.01 nN (Section 9.4.7.1) requires only ~5 noncovalent bonds (breaking strength ~2 pN/bond; Section 9.4.4.2) to the glyocalyx or plasma membrane; a single 100 nm² footprint (Section 9.4.3.3) can provide up to 0.3 nN of harness anchoring force.

9.4.7.4 Steering and Control

A complete understanding of motile cell mobility biological control systems is not yet available. However, present knowledge is sufficient to support the conclusion that virtually complete command of cellular mobility systems by in cyto medical nanorobots should be feasible. Such control may be effected by either biochemical or mechanical means.

The “amoeboid” movement (Section 9.4.3.7) of leukocytes and fibroblasts relies primarily upon temporary extensions of the cell surface rather than major movements of cytoplasm. These temporary extensions, such as the small filopodia and the larger lamellipodia, are surface evaginations involving bundles of actin filaments or microtubules in parallel array, often comprising a large fraction of the total surface area of the cell. Motion is made possible by an internal system of bracing which is built or modified according to the needs of the cell. Immediately below the plasma membrane is a region of densely packed and interconnected actin microfilaments called the cell cortex (Section 8.5.3.11). A leukocyte moves by breaking down actin filaments and rebuilding them in the proper orientation, pushing the cell in a new direction and allowing the posterior regions to retract.

Fiber bundles isolated by microlaser surgery from a glycinated fibroblast have been shown to contract with the addition of ATP. Focal attachments between substrate and cell are modulated both by external chemotactic gradients on the substrate and by internal biochemical regulatory pathways (second messenger molecules; Section 8.5.4.1); similar signal transduction pathways control the polymerization and disassembly of filaments and microtubules comprising the podial extensions. Regulatory proteins that bind actin monomers and filaments, some of which function in response to transmembrane signaling, include ADF, adseverin, calsedion, cap Z, cap 100, cofilin, gelsolin, MCP, phospholipase C-γ, profilin, and severin. Manipulation of the internal concentrations and spatial distributions of appropriate intracellular messenger molecules and ATP should permit in cyto medical nanorobots to command these biochemical pathways and to regulate localized polymerization processes, thus controlling the speed and direction of motile cell locomotion.

* Treadmilling and dynamic instability control microtubule length in vivo.

** PDGF concentration in human serum is 50 ng/cm³ (Appendix B); optimum concentration is 1-20 ng/cm³, or ~1 molecule/micron², for chemotactic response of neutrophils and monocytes during wound repair (Chapter 24). Fibroblasts need the higher concentration in this range (~20 ng/cm³); PDGF stimulates a full mitogenic response.
As another example of an intracellular mediator system that may be subject to direct nanorobotic control, consider the Ras superfamily of small guanosine triphosphatases (GTPases). One member of this superfamily, known as Rho, can be activated by the addition of extracellular ligands (such as lysophosphatidic acid), leading to the assembly of contractile actin-myosin filaments (stress fibers) and associated focal adhesion complexes.\(^{1538}\) Rho acts as a molecular switch to control a signal transduction pathway linking plasma membrane receptors to the cytoskeleton.\(^{1539}\) Rac, another member of the Rho subfamily, can be activated by the association of distinct guanine nucleotide exchange factors (guanine exchange factors or GEFs) with the plasma membrane receptor, leading to the activation of the small GTPase Rac. Rac activation induces the formation of lamellipodia and filopodial surface protrusions that are also associated with distinct actin-rich filamentous structures.\(^{1540,1541}\) Members of the Rho GTPase family appear to be key regulatory molecules linking surface receptors to the organization of the actin cytoskeleton.\(^{1539}\) By 1998, about 10 GTPase-activating proteins and 3 guanine nucleotide dissociation inhibitors (GDI-s) were known to be potential down-regulators of GTPase activity, and at least 20 cellular effector targets with relevant protein-protein interaction domains, had been described.\(^{1539}\) Local cytoplasmic injections or extractions of nanomolar concentrations of these second-messenger substances should permit full control of cytovegetable direction and speed.

Less subtle biochemical methods of cytoskeletal manipulation are also available to nanorobotic pilots. For example, the drugs colchicine and taxol disrupt microtubule function in distinctly different ways. Colchicine binds to tubulin, strongly inhibiting its polymerization or depolymerization reactions unless they are actively suppressed. \(^{936}\) Taxol has a much smaller effect: it reversibly inhibits microtubule assembly. \(^{936}\) (Microtubules are always losing and reacquiring tubulin dimers, and polymerization of actin microfilaments.\(^{939}\) Drug molecule diffusion through the micrometer-thick cellular plasma membrane delays, a purely biochemically-mediated control system is probably limited to the maximum speeds of actin polymerization or osmotic inflation, or ~1-10 microns/sec. However, purely mechanical cytovegetable control systems can be computed at much higher speeds, at least in tension. Individual actin microfilaments will tolerate a maximum tensile stress of ~108 pN (Section 8.5.3.11); the binding force between cytoskeleton and cell cortex at focal adhesions is of similar magnitude. Let us assume that the piloting nanorobot manipulates just a single fiber to drive each ~1 micron diameter lamellipodium, and that this single fiber is not driven beyond 10% of its theoretical failure strength, a very conservative limit. Using Eqn. 9.73 to crudely approximate the viscous drag force on the protuberance as it moves through interstitial fluid, each lamellipodium can be flexed at speeds up to ~500 microns/sec. Manipulating more fibers adds to the margin of safety. (Cyclosis in mature plant cells such as *N. tabacum* and *Ch. vulgaris* produces circular motions of some cell components at fastest 100 microns/sec.\(^{1538}\) Similar manipulation of the much stronger intermediate fiber (IF) cytoskeleton (Section 8.5.3.11) would in theory permit speeds up to ~0.1 m/sec, but such aggressive forces might rip the IF fibers loose from their intracellular moorings. Another potential concern is that rapid cycling will stimulate unwanted peripheral actin polymerization or depolymerization reactions unless they are actively suppressed.

Electric fields may also drive leukocyte motion (Section 4.9.3.1). For example, a mild electric current induces lymphocytes to travel in the same direction of the current ("electrotaxis") at speeds up to ~0.3 microns/sec. Small electric fields have been shown experimentally to stimulate leukocyte diapedesis (Section 4.9.3.1), though multiple nanorobots would probably be required to generate the necessary currents.

### 9.4.7.5 Navigation and Sensing

The nanorobotic pilot can directly access the high-resolution extracellular acoustic navigational grid (if such has been installed in the patient; Section 8.3.3) to establish absolute position and orientation of the cytovehicle within the human body, or may employ alternative positional or functional navigational cues as may be provided by the attending physician (Chapter 8). Chemosensory data (including tissue- and cell-recognition events) may be accessed by the attending physician (Chapter 8). Chemosensory data (including tissue- and cell-recognition events) reaching the exterior surface of the piloted cell trigger natural internal biochemical cascades which are readily monitored and interpreted by the cyto nanorobot. Extracellular and ECM-related mechanosensory data and plasma membrane stretch sensor data produce similar recognizable internal biochemical and cytomechanical effects.

Eavesdropping on natural chemotactic signals is particularly effective. For instance, after leaving the blood vessels, neutrophils recognize biochemicals produced by bacteria and migrate toward a rising concentration. Laboratory demonstrations of neutrophil chemotaxis often use a concentration gradient of FMLP (n-formylmethionine-leucine-phenylalanine), a peptide chain produced by some bacteria, at ~0.1 nanomole/cm\(^3\). Chemotaxis also occurs among human T-lymphocytes in the presence of a 0.01-1 picomole/cm\(^3\) gradient of Met-enkephalin or β-endorphin. Chemotactic cytokines for leukocytes, macrophages, and fibroblasts—many of which are tissue-specific (and thus additionally serve as an aid to
navigation) or condition-specific (e.g., inflammation)—are well-studied. At least two cytoplasmic signaling proteins, CheA and CheW, are known to interact with chemosensors in vitro. Various intracellular mediator protein molecules convey chemoreceptor signals by diffusion (across one cell length in ~200 milliseconds) to the flagellar motors, which drive the bacterium forward in 20-40 micron/sec bursts. MCPs respond to attractants and repellants in the extracellular environment. All these signals can be detected in cyto by an eavesdropping "piggybacking" nanorobotic pilot equipped with appropriate (membrane-penetrating) cytoplasmic chemosensors.

The nanorobotic pilot may also deploy loosely-tethered nanosensor remotes and insert them into the cytovehicle's plasma membrane, during or after cytopenetration. These nanosensors may employ simple mechanical plasma membrane locking (Section 9.4.5.2) or a stationary solvation wave pattern (Section 9.4.5.3) to straddle the plasma membrane, mimicking natural transmembrane integral proteins (Fig. 8.37), or could anchor one end of their transmembrane structure to the underlying cellular cortex. Such sensors could gather whatever extracellular thermal, mechanical, acoustic, chemical, or electromagnetic data might be required by the pilot for proper navigation and steering. The nanosensors are retrieved when the nanorobotic pilot exits the cytovehicle at the destination.

When macrophages and other leukocytic cells become infected, they express B7 (a co-stimulator molecule) on their plasma membrane surface which can be recognized by a T-cell CD28 receptor protein, triggering an immunologic response in the presence of surface-displayed antigen. Nanorobotic pilots should inspect the cell surface of all prospective cytovehicles for B7 and similar patho- surface-displayed antigen. Nanorobotic pilots should inspect the cell surface of all prospective cytovehicles for B7 and similar pathogenic flags prior to cytopenetration, so as not to unintentionally choose an infected cell for cytocarriage which could then spread the infection or become subject to immunosystem attack. If the cell subsequently becomes infected and begins expressing B7 or other warning substances during the journey, the nanorobotic pilot should abort the mission and steer the cytovehicle to a nearby disposal site, or implement immediate cytotherapeutic measures. Failing this, the pilot should abandon the vehicle at once.

9.4.7.6 Cytovehicular Behavioral Control
Besides simple steering and chemotaxis, many stereotypical cell behaviors may be triggered biochemically as well. Diapedesis, ECM brachiation with a preference for a particular ECM ligand, particle engulfment, transitions between migratory and profibrotic phenotypes in fibroblasts, cell volume and shape changes, cell cycling and growth, and epithelial-mesenchymal transformations may be initiated, amplified, or suppressed by regulating the flows and distributions of various cytoplasmic signal molecules which may include proteins or nonproteins (such as Ca++ or GTP). Motility-stimulating factors such as autocrine stimulating factor and migration stimulating factor are produced by fibroblasts and act in an autocrine (Section 7.4.5.1) fashion. Leukocytes keep rolling over vascular endothelium until a neutrophil activator (e.g., a selectin) is encountered, whereupon they stick tightly and undergo the shape change characteristic of neutrophil diapedesis—the process that nanorobotic pilots should be able to monitor and control. Mammary epithelial cells receive mechanical signals from the ECM via a tyrosine kinase (biochemical) signal transduction pathway through the β1 integrin receptor leading to the activation of elements in the promoter region of the β-casein gene. Release of Sr++ inside cytolytic T lymphocytes induces degranulation—the emission of cytotoxic granules containing granulysin that can inhibit growth in a broad spectrum of pathogens, including bacteria, fungi, and parasites, and can directly kill the tuberculosis bacterium.

Most integrins bind to their ECM ligands via the tripeptide RGD (Arg-Gly-Asp), so the nanorobotic pilot can release RGD as an antagonist to moderate or to prevent cell-ECM interactions. Alternatively, ECM-integrin binding can be enhanced by specific cytoplasmic molecules such as the cell adhesion regulator (CAR), a myristoylated protein that can connect the plasma membrane to the cytoskeleton. ECM molecules interact with their receptors at the cell surface, transmitting signals directly or indirectly to second messengers which in turn unravel a cascade of events leading to the coordinated expression of a variety of genes involved in cell adhesion and release, migration, proliferation, differentiation, and death.

By intercepting and regulating specific second messenger pathways, or by artificially triggering them mechanically, the pilot can exercise considerable control over cell behavior.

The nanorobotic pilot must inhibit certain natural cell behaviors that might distract the cytovehicle from its imposed mission. For example, emission of Rap1-GTP protein into the cytoplasm of a Jurkat T cell renders the cell anergic (unresponsive) and incapable of transcribing the gene encoding interleukin-2 when stimulated with antigen. Intracellular kinesin molecular motors (Fig. 9.32) are inhibited by adenosulfate-2 (AS-2; MW = 738 daltons), a molecule that targets the kinesin motor domain. Activated macrophages themselves secrete chemotactant factors that may draw a crowd of leukocytes, fibroblasts, and other macrophages toward the flight path, impeding mobility and possibly triggering a natural granulomatous (encapsulation) reaction. These processes may be desirable to stimulate at the destination, but they should be suppressed en route.

Unwanted cytovehicular responses due to nanorobot intrusion should also be suppressed, much as pathogens such as M. tuberculosis resist or inhibit oxidative killing, attack by defensins, phagosome-lycosome fusion, and formation of the electron-transparent zone that impairs lysosomal enzyme diffusion, thus ensuring their survival inside macrophages for long periods of time.

Cell behavior can also be mechanically regulated by altering cell shape, because cells can detect when a mechanical strain is placed upon them. In one experiment, stretching a fibroblast by 1% for 180 sec triggered large-scale changes in cell shape, including retraction of pseudopods and cell elongation. Other experiments show that very flat cells, with their cytoskeletons stretched, sense that more cells are needed to cover the surrounding substrate (as in wound repair) and that cell division is needed. A rounded shape indicates that too many cells are competing for space on the matrix and that cells may be excessively proliferating; some must die to prevent tumor formation. Normal tissue function is established and maintained between these two extremes. Certain mechanical stresses on cells, such as might be generated by a nanorobotic pilot inside the cytovehicle or dense cytovehicular traffic near the destination, may trigger unwanted reactions that should be avoided.

9.5 Ex Vivo Locomotion
In this final Section of Chapter 9, our attention turns to locomotion across, through, and around the external surfaces of the human body. The most important of such surfaces are the teeth (Section 9.5.1) and the skin (Section 9.5.2). Another important consideration is aerial locomotion near the human body (Section 9.5.3). These discussions are necessarily incomplete; more com-
prehensive treatments are beyond the scope of this introductory text.

### 9.5.1 Dental Walking

The exposed surfaces of natural dentition consist of a ~500 micron thick layer of enamel coating the exposed crown of each tooth inside the mouth. Enamel is a composite structure containing ~96% inorganic material similar to hydroxyapatite, a dense bone mineral, embedded in an organic matrix of glycoprotein and a keratin-like protein. The mineral forms crystals oriented into hexagonal cylinders with a horseshoe-shaped cross-section wrapping around the dentin, forming enamel prisms ~5 microns in diameter stacked in parallel congeries.646,854,1571 These prismatic columns are directed vertically at the summit of the crown and horizontally along the sides of the tooth, pursuing a generally wavy course. There are no nutritive channels permeating the enamel and leading into the dentinal layer below, although dentinal tubules (nutritive channels) do extend from the pulp chamber into the dentin which may get extremely close to the occlusal (enamel) surface of the tooth, and may be exposed by receding gums. However, ~0.4-micron-wide fissures separating enamel prisms are filled with an interprismatic substance of much higher organic content. The entire enamel surface is pockmarked by numerous micropores of irregular but generally oval shape, ranging from 0.1-2 microns (average ~0.5 micron) in diameter and spaced ~1-2 microns apart. Thus a dental-walking micron-size nanorobot encounters a spongelike topography with holes and fissures comparable in size to the nanorobot diameter, and with gently rolling hills ~2 device diameters in height and ~5 diameters across. (There is also a thin layer of salivary proteins coating the tooth enamel surface.)3669,3670

Nanorobots may traverse tooth enamel using ambulatory techniques previously described in other contexts (Section 9.4), including legged walking, rolling, or amoeboid locomotion using footpads for anchorage. Teeth may be clenched with a force up to 300-500 N, although forces of ~50 N are more common during normal mastication of soft foods. Assuming ~1 cm² of dental contact surface, shear forces at that surface during chewing or clenching are 0.5-5 x 10⁶ N/m², close to the limit for noncovalent anchorage (Section 9.4.3.3). Nanorobots seeking further refuge from shear forces may congregate in the valleys between enamel prisms or can seat themselves in the irregular micropores. E. Reifman notes that ~90% of U.S. adults have multiple occlusal surface restorations composed of many different materials such as resin composites of various hardness coefficients, porcelain or gold crowns, amalgam fillings, composites, and the relatively newer but very popular glass polymer crowns, representing a substantial percentage of all exposed dental surfaces with which oral nanodevices must cope; yield strength of these materials at 0.1% strain deformation typically ranges from 110-10,000 atm.3271

Are the forces of dental grinding sufficient to crush medical nanodevices? Tests of sized jeweler’s grinding powders by the author confirm that even irregularly-shaped diamondoid particles ~3 microns and smaller apparently roll smoothly out of the way when ground between the teeth, whereas particles larger than ~3 microns cannot roll sufficiently and retain a sensible grittiness. In the case of a trapped nanorobot-size diamond block slowly being squeezed between opposed enamel surfaces (static force), the Young’s moduli for enamel and diamond are 7.5 x 10¹⁰ N/m² and 1.05 x 10¹² N/m², respectively (Table 9.3), so the enamel deforms ~14 times more than the diamond. Crystalline diamond can deform at least ~5% before fracturing, diamond composites (Chapter 11) much more; a 50% strain in each enamel surface (an indentation equal to the entire diamond block radius, probably causing the enamel to fail) produces a 3.6% deformation across the diameter of the block, slightly below the conservative tolerance limit for diamondoid fracture strain. In the case of a 50-gram jaw moving at the maximum 0.1 m/sec clenching velocity that is suddenly brought to a halt via enamel deformation to a depth equal to a ~1 micron nanorobot radius across a 1 cm² enamel contact area, the jaw decelerates at ~1000 g’s producing a kinetic force of ~500 N and a contact force of 5 x 10¹⁰ N/m² (~50 atm), well below the ~10¹¹ N/m² failure strength of solid diamond.

However, nanorobot crushing strength is an explicit design parameter which is normally significantly lower than for solid diamond if the device has large unbraced internal voids or diameters that are weak in compression. Such nanorobots may not be regarded as solid diamond structures as assumed above. For example, the example shown in this context. For example, the Euler buckling force for a solid cylindrical rod of outer radius R is proportional to R⁴ (Eqn. 9.44), but for a hollow cylinder with inner radius r the buckling force is proportional instead to (R⁴−r⁴) during axial compressions.364 Hence a trapped, thin-walled cylindrical nanorobot of radius R with r/R = 0.90 buckles at only ~34% of the force needed to buckle a solid rod of radius R, and thus could probably be crushed by static compression of the teeth. The oral cavity appears to be the only place in the human body where thin-walled medical nanorobots might plausibly be destroyed by mechanical grinding. Such destruction is avoided by nanorobots with composite diamondoid shells ≥10% R thick at a maximum allowable strain ≲10%, or by nanodevices possessing maps of mastication contact spots that actively avoid those spots during nanorobotic perambulations.

### 9.5.2 Epidermal Locomotion

The epidermis is a modified cellular surface, so most of the techniques described for cytoambulation (Section 9.4.3) are applicable to epidermal locomotion as well. There are so many special situations and obstacles to motility on skin that we can only mention a few of them here:

A. **Flaky Corneum**—Desiccated stratified epidermal cells crack, chip, or flake off entirely when rubbed or jostled with too much force. Nanorobots traversing a dead cell could find themselves flaked off into the air or onto the floor, especially if the surface is rubbed by the patient.

B. **Steep Topography**—Many physical features of the skin are huge in scale compared to a micron-size nanorobot and thus must be carefully circumnavigated. Dermal fingerprint ridges are ~500 microns wide and 20-50 microns deep. Hair follicle shafts are 50-100 microns in diameter. Scent-secreting apocrine glands are ~200 microns in diameter; eccrine (sweat) glands are only 20 microns wide. Skin-dwelling fungi (e.g., athlete’s foot) grow into irregular globs 5-20 microns in size. The tongue surface includes taste buds ~150 microns wide and tiny papillae measuring 30 microns wide and 80 microns long. Other surface features are described in Section 8.6.1. At the micron scale, the topmost layer of the epidermis is littered with obstacles and protrusions, with electron micrographs showing a surface that closely resembles a flaky puff pastry in appearance.1571

C. **Constant Motion**—The skin is almost always in motion at size scales up to ~1000 microns, folding and unfolding, stretching and tightening, twisting and curving, with normally distant surfaces being brought into and out of contact as the patient moves about normally.
D. Dust Mites and Bacteria—In eyelashes and eyebrows, the hairs of the upper lip, in the ears, and elsewhere on the body, dust mites measuring 150 microns x 300 microns survive by chewing up dead skin cells. Epidermal-walking nanorobots must be able to identify mite surfaces, and avoid climbing aboard them by mistake. Everyone inhales a few mites now and then; mite excrement (to which some patients are allergic) is concentrated into fecal packets so small and light that they float in the air and may settle back onto the skin. Many other surface-dwelling ectoparasites and microfauna similarly must be avoided.253 Bacterial density over most of the body surface is quite low. -10^3/cm²; the count ranges from -10^2/cm² on the palms and dorsa of the hands, up to 10^4-10^5/cm² on the hairy axilla, scalp, perineal regions, and beneath the distal end of the nail plate.360

E. Obstructions—The skin may be covered with dirt, or grease, or cooking oils and fats, or sebaceous gland oils, which may be many tens or even hundreds of microns thick. Another potential obstacle is the sweat wash—the skin can pour out up to 2 liters/hour of perspiration. Nanorobots can easily crawl underneath the tightest-fitting clothing or rubber gloves, although a patch of superglue stuck to the skin could provide a formidable obstacle requiring overpassage.

F. Itching/Crawling Sensations—Tickling sensations attributable to isolated nanorobots traversing the skin are unlikely. Skin-crawling ~1-2 mm ants are readily detected; 100-micron mites are not. Absolute epidermal pressure stimulus thresholds, as measured by a laboratory esthesiometer, range from a low of 2000 N/m² at the tongue- and epidermal pressure stimulus thresholds, as measured by a laboratory to isolated nanorobots traversing the skin are unlikely. Skin-crawling is ~45 cm/sec. Skin sensor frequency response is << KHz (Table in this book."

Additional discussion of the weight of a (100 micron)³ nanorobot, if distributed across a (100 micron)² contact surface, is only ~1 N/m², quite undetectable on the skin. At a velocity of 1 cm/sec, the inertial force required to propel a 100-micron nanorobot is ~1 nN (Section 9.4.2.1). This additional force, if distributed over 10 footpads each of area 100 micron², gives a shear pressure of ~1 N/m² across all contact surfaces, also undetectable; the sensible threshold velocity for such a nanorobot is ~45 cm/sec. Skin sensor response frequency is << KHz (Table 7.3), whereas skin walkers ≥ 10 microns in size may employ υleg ≥ KHz leg motions; the weight of a 1 micron² nanorobot produces only ~0.01 N/m² of contact pressure.

Nanorobots of large size can also step over most of the obstacles described above. Consider an Rnano = 50 micron walking nanorobot with Nleg = 10, Lleg = 100 microns, Rleg = 15 microns, and υleg = 1 cm/sec as defined for Eqn. 9.82, with υleg = υleg / 2 Lleg ~ 50 Hz. From Eqn. 9.75, the viscous force per leg is Fleg = Fnano3 ~ 5 nN in a 20°C water bath, 0.1 nN in 20°C air. From Eqn. 9.73, Fnano3 ~9 nN in water, 0.2 nN in air. Thus the total driving force Ftotal = 59 nN in water, 1.2 nN in air, assuming a still medium.

Another way to avoid obstacles is to hop over them in the manner of a jumping flea, an insect that subjects itself to an acceleration of ~200 g’s while leaping ~130 times its own dimension.739 Diamondoid saltators could tolerate much higher accelerations, although it is important not to disturb the takeoff surface; also, airborne ex vivo trajectories may be only poorly controlled. The approach is interesting1982,2385,2386 but will not be considered further in this book.

What about dislodgement forces? Assuming a headwind of v ~ 1 m/sec and a d ~ 10 micron boundary layer, then from Eqn. 9.58 the shear stress (and shear force) on a 100-micron wide nanorobot is F/A ~ 100 N/m² (1000 nN) in water, 2 N/m² (20 nN) in air. A single micron² footpad with a conservative adhesive pressure of 10^2 N/m² (Section 9.4.3.3) provides a 1000 nN anchorage force, allowing the nanorobot to remain adhered to the skin even in the strongest water currents (e.g., white water rapids, showers, or spas at ~1 m/sec per anchored footpad used) and air currents (e.g., up to ~50 m/sec (~110 mph) winds per anchored footpad used). A finger rubbed hard against the skin with the objective of dislodging adhered epidermal-resident nanorobots may apply up to 10 N of force over a ~1 cm² area, giving a ~10⁻³ N/m² shear force (~1,000,000 N) which may be resisted by ten 100 micron² footpads each having a maximum adhesive pressure of ~10⁴ N/m² (Section 9.4.3.3).

What about epidermal penetration? Complex and subtle low-speed methods are readily imagined (Section 9.4.4), but let us consider here a simple brute-force approach. Failure strength of human skin is ~10⁻⁷ N/m² (Table 9.3), so a 1 micron² rod may be driven through the skin with a force of ~10,000 nN. A 100-micron long telescoping manipulator arm with stiffness ~75,000 N/mm can apply 10,000 N to the rod at 1 cm/sec producing an elastic deflection in the arm of less than 1 mm, drawing ~100,000 pW of power (manipulator arm power density ~10⁻⁶ watts/m³) at the moment of penetration (Section 9.3.1.4). Bruising and other biocompatibility issues related to dermal penetration are deferred to Chapter 15.

The surfaces of toenails and fingernails have 2-20 micron features and longitudinal striations with 10-15 micron valleys, and are usually littered with dirt, bits of skin, and other surface debris as small as 1-2 microns in size. The nail surface is heavily keratinized, with a failure strength about 20 times higher than that of skin (Table 9.3). Similar considerations apply as with the skin, in regard to nanorobot locomotion.

What happens if the nanorobot falls off, or is otherwise removed from, the epidermal surface? Solem1982 has analyzed the case of nanorobots walking on walls and ceilings. Setting the electrostatic image force (Eqns. 9.11) equal to the gravity force ρnano Lnano³ g for a cubical conducting-surface nanorobot of edge Lnano the voltage required to cling to the ceiling without falling off is:

\[
V_{plate} = z_{sep} \left( \frac{2 \rho_{nano} L_{nano} g}{\varepsilon_0 \kappa_{e}} \right)^{1/2} \text{ (volts)} \tag{Eqn. 9.83}
\]

Taking z_{sep} = 0.5 microns, ρ_{nano} = 200 kg/m³, g = 9.81 m/sec², \varepsilon_0 = 8.85 x 10⁻¹² farad/m, \kappa_e = 1, and L_{nano} = 10 microns, then V_{plate} ~ 0.1 volt. Similarly, the largest nanorobot that can cling to the ceiling using an adhesive contact surface of strength S has a characteristic dimension:1982

\[
L_{max} = \frac{S}{\rho_{nano} g} \tag{Eqn. 9.84}
\]

For S = 10-100 N/m² (typical for weak adhesives such as Post-It notes), then L_{max} ~ 500-5000 microns. By wetting the contact area with water having surface tension \gamma ~ 73 x 10⁻³ N/m at room temperature (Section 9.2.3), a cubical nanorobot clinging to the ceiling may have a maximum characteristic dimension of:1982

\[
L_{max} = \left( \frac{4 \gamma}{\rho_{nano} g} \right)^{1/2} \sim 3900 \text{ microns} \tag{Eqn. 9.85}
\]
9.5.3 Nanoflight

Perhaps surprisingly, aerobiotics has many useful applications in nanomedicine. Examples include monitoring and diagnostic systems, since flying nanorobots can transit the length of a patient’s body in ~1 sec or less (Chapter 18); sterile fields (Chapter 19); personal defensive systems (Chapter 21); and accident recovery systems (Chapter 24). Unfortunately, in 1998 the aerodynamics of micron-scale flight systems was only lightly studied and aerial robotics was still considered an exotic sport or a military curiosity.

9.5.3.1 Nanoflight and Reynolds Number

It has been observed that a 30-cm paper airplane will glide slowly and stably, but a 3-cm paper airplane made from thinner paper requires a much higher velocity/size ratio to remain airborne, and with a noticeable lack of stability—sometimes the plane flies well, sometimes not. If size is further reduced into the millimeter range, the plane almost cannot fly.

As in the case of swimming, this transition can be explained in terms of the Reynolds number $N_R$ (Section 9.4.2.1), the ratio of inertial to viscous forces acting on a body that is passing through a fluid such as air. Generally speaking, microscopic organisms (e.g., $N_R < 10^2$) or flying nanorobots with $N_R < 1$ move by utilizing viscosity, while macroscopic objects such as aircraft (e.g., $N_R = 10^6$) with $N_R > 1$ use inertia to generate lift. Millimeter-size airfoils with $N_R \approx 0.1-100$, as typified by flying insects, occupy a transitional regime. Insects make use of both inertial and viscous forces, often employing unusual wing flapping patterns and elastic energy storage systems to remain aloft. The smallest known flying insect can make any use of aerodynamic lift (inertial forces are the four-winged parasitic chalcid wasp Encarsia formosa, which has a total wing span of ~1.4 mm. Aerobic machines with wingspans smaller than ~100 microns probably must make almost exclusive use of viscous propulsive forces.

A nanorobot of dimension $L$ flying at velocity $v$ through 20°C sea-level dry air ($\rho_{\text{air}} = 1.205 \text{ kg/m}^3$, absolute viscosity $\eta_{\text{air}} = 0.0183 \times 10^{-3} \text{ kg/m sec}$)$^{1577}$ has a Reynolds number $N_R = 66,000 v L$ (Eqn. 9.65). Viscous forces dominate when $N_R < 1$, or when:

$$v \leq \frac{15}{L_{\text{micron}}} \quad \text{(m/sec)} \quad \text{(Eqn. 9.86)}$$

where $L_{\text{micron}}$ is characteristic nanorobot size expressed in microns. Thus a 1-micron nanorobot remains in the viscous regime up to ~15 m/sec flight velocity (34 mph), a sufficient speed for most medical applications. Note that formulas involving viscous forces may not be applicable to aerobic airfoils with $L \leq \lambda_{\text{gas}}$ (Eqn. 9.23), indicating transitional or ballistic flow (Section 9.2.4).

The main implication of Eqn. 9.86 is that conventional aerodynamics technologies such as rigid wings and jets are usually inappropriate for micron-size flyers. Instead, aerial nanorobots may more profitably employ nanotaxis mechanisms as described in Section 9.4.2.5, including surface deformations (e.g., flexible oars or wings, cilia, invaginating doughnuts, rotating spheroids, traveling waves), inclined planes (e.g., screw drives, corkscrews, flagella), volume displacement, and viscous anchoring. Specific aeromotive mechanisms are of great interest but will not be considered further in this Volume. In many cases, a viscous-regime nanorobot may exit the bodily fluids and enter the atmosphere, or vice versa, using the same propulsive mechanisms, though operated at a modified speed or pitch angle.

9.5.3.2 Nanoflight and Gravity

In addition to forward progression through the medium, atmospheric flight requires active and continuous support against the pull of gravity. For small spherical objects of radius $R = R_{\text{nano}}$ in the laminar flow (low $N_R$) regime, the rate of fall in a still medium is approximated by Stokes’ law for Sedimentation (Eqn. 3.10).$^{**}$ For nanorobots near the Earth’s surface that are falling in air, with $g = 9.81 \text{ m/sec}^2$, $\rho_{\text{particle}} - 1000 \text{ kg/m}^3$, $\rho_{\text{fluid}} = \rho_{\text{air}}$, and $\eta = \eta_{\text{air}}$, then terminal velocity $v_t = 1.2 \times 10^8 R_{\text{nano}}^{-2} \text{ m/sec}$. An $R_{\text{nano}} = 1$ micron nanorobot falls at $v_t = 120 \text{ microns/sec}$; an $R_{\text{nano}} = 10$ micron nanorobot falls at $v_t = 1.2 \text{ cm/sec}$, requiring a rather modest power expenditure of only $P_{\text{nano}}/e\% = (0.5 \text{ pW})/e\%$ to remain aloft (Eqn. 9.74).

Note also that in free atmosphere only the largest particles ever settle out by gravity alone. Thermal and dynamic turbulence keeps most of the smaller particles in suspension long beyond the period indicated by Eqn. 3.10. For example, nanorobot thermal velocity $v_{\text{thermal}}$ (Eqn. 3.3) exceeds nanorobot terminal velocity $v_t$ (Eqn. 3.10) for a nanorobot radius $R_{\text{nano}} \leq 2$ microns in sea level air at 20°C with $\rho_{\text{nano}} = 1000 \text{ kg/m}^3$. The energy required to raise a 1-micron nanorobot a height of ~1 micron in a 1 g gravity field equals ~80 Joules. Random indoor atmospheric eddies of 1-10 cm/sec caused by the movements of people, animals, doors, and heating and air conditioning systems equal or exceed the Stokes terminal velocity for particles of $R \approx 10-30$ microns.

9.5.3.3 Buoyant Nanoballoons

Gravity may also be overcome by reducing density relative to the surrounding medium. An object has neutral buoyancy when its density equals that of the medium in which it is suspended, becoming “weightless” within that medium. For example, a nanorobot whose interior volume consists 90% of vacuum has a $v_t$ ~10 times smaller than a completely solid object of equal size.

What is the tiniest possible lighter-than-air balloon? J.S. Hall notes that for a one-atom-thick graphene shell the out-of-plane bending stiffness of the C-C bond is much lower than the in-plane stretching stiffness. This is why hollow fullerene of submicron diameter are experimentally observed to collapse (and remain collapsed due to van der Waals forces) even when their interiors are not evacuated. Simple internal bracing sufficient to stabilize an evacuated structure outweighs the lift. For example, applying the Euler buckling formula (Eqn. 9.44) to three diamondoid orthogonal diametral stiffening rods inside a spherical evacuated nanoballoon gives a scale-invariant total beam mass ~6 times the mass of the displaced air, hence net lift is impossible for any device radius using this minimal crossbeam design although macroscale geodesic trusswork-stabilized vacuum balloons cannot be ruled out. (See also Section 10.3.5.)

Hall also observes that pressurizing nanoballoons to atmospheric pressure removes most shell stress. This strategy also eliminates the need to thicken the single-atom shell walls, up to a nanoballoon radius of at least $\sigma_{\text{w}} \tau_{\text{wall}} / \Delta p = 100$ microns, taking wall thickness $\tau_{\text{wall}} = 0.17 \text{ nm}$, a conservative diamondoid wall working stress $\sigma_{\text{w}} = 10^{10} \text{ N/m}^2$ (Table 3.3), and allowing a maximum environmental

* From Eqn. 9.14, the capillary force will restrain a 1-micron diameter nanorobot with a force of ~440 nN in pure water at the liquid-air interface due to surface tension, but this attraction may be reduced to less than ~4 nN by discharging small aliquots of the appropriate surfactants into the local aqueous environment; Section 9.2.3.

** In sea level pressure dry air, particles of 100 nm diameter will fall about twice as fast as estimated by Eqn. 3.10; 10-nm diameter particles fall about 12 times as fast as Stokes’ sedimentation formula predicts.
pressure fluctuation of Δp = 0.17 atm (vs. -0.002 atm for 140 dB sound waves (Section 4.9.1.5), -0.1 atm normal barometric variation (Section 4.9.1.6), and -2 atm maximum sound pressure in air). Bursting strength of pressure shells is briefly treated in Section 10.3.1.

The smallest atmospheric-pressurized atomic-walled nanoballoon that can achieve neutral buoyancy has radius $R_{\text{min}} = 3 \rho_{\text{wall}} \rho_{\text{gas}} / (\rho_{\text{air}} - \rho_{\text{gas}})$, where $\rho_{\text{wall}}$ is wall density, $\rho_{\text{air}}$ is air density ($\rho_{\text{air}} = 1.2929 \text{ kg/m}^3$ for dry air at STP), and $\rho_{\text{gas}}$ is the density of the filling gas. Thus $R_{\text{min}} = 1.6$ micron for hydrogen gas with $\rho_{\text{gas}} = 0.0899 \text{ kg/m}^3$; $R_{\text{min}} = 1.7$ micron for STP helium gas ($\rho_{\text{gas}} = 0.1785 \text{ kg/m}^3$) which, in conjunction with a slightly thicker nondiamondoid (e.g., sapphire) shell, eliminates flammability concerns at all device number densities. Diffusion leakage must also be addressed (Section 10.3.4).

Nanoballoons of radius $R > R_{\text{min}}$ can carry payloads of mass:

$$M_{\text{payload}} = \frac{4}{3} \pi \rho_{\text{gas}} R^3 \left( 1 - \frac{\rho_{\text{wall}}}{\rho_{\text{gas}}} \right) - 4 \pi \rho_{\text{wall}} R^2$$  \hspace{1cm} \text{(Eqn. 9.87)}

For instance, a helium-filled R = 2.2-micron fullerene sphere can lift a $10^{-11}$ kg payload mass, representing a payload volume of $0.01$ micron$^3$ at $\rho_{\text{payload}} > 1000 \text{ kg/m}^3$. Expanding nanoballoon radius to $R = 6.8$ microns increases payload volume to ~1 micron$^3$. Pressurized buoyancy-based lift systems may be useful either in early-generation aerial nanodevices that must rely upon primitive energy supplies, or in default-float applications. However, the modest power expenditure needed to overcome gravity in micron-size devices (Section 9.5.3.2) suggests that nonbuoyant active-propulsion designs will normally be preferred.

### 9.5.3.4 Nanoflyer Force and Power Requirements

What forces must be applied, and what power must be expended, in order for a flying nanorobot to maintain continuous and controlled progression through the air? An exact calculation requires a detailed knowledge of the mode of locomotion, the shape and dimensions of wing or body, airfoil surface characteristics, and many other factors. However, the force needed to drive a spherical flyer of mass $m_{\text{nano}}$ through the air at a velocity $v_{\text{nano}}$ may be approximated by summing the three drag components of skin friction ($F_{\text{viscous}}$), pressure drag ($F_{\text{inertial}}$), and induced drag due to lift ($F_{\text{induced}}$), as:

$$F_{\text{nano}} = F_{\text{viscous}} + F_{\text{inertial}} + F_{\text{induced}}$$ \hspace{1cm} \text{(Eqn. 9.88)}

$$F_{\text{viscous}} = 6 \pi \eta_{\text{air}} R_{\text{nano}} v_{\text{nano}}$$ \hspace{1cm} \text{(Eqn. 9.89)}

$$F_{\text{inertial}} = C_D \pi R_{\text{nano}}^2 \rho_{\text{air}} v_{\text{nano}}^2 / 2$$ \hspace{1cm} \text{(Eqn. 9.90)}

while conservatively taking the dimensionless coefficient of drag $C_D = 2$. Induced drag, associated with the shedding of transient vortices from each wing tip, is an important consideration in macroscale winged vehicles but may be ignored here (e.g., $F_{\text{induced}} = 0$) for microscale flyers operating in the viscous or transitional flight regimes.

The nanorobot aeromotive energy plant, of efficiency $e\%$, must develop a continuous power of:

$$P_{\text{nano}} = \frac{\nu_{\text{nano}} F_{\text{nano}}}{e\%}$$ \hspace{1cm} \text{(Eqn. 9.91)}

giving a whole-device power density of:

$$D_{\text{nano}} = \frac{3 P_{\text{nano}}}{4 \pi R_{\text{nano}}^3}$$ \hspace{1cm} \text{(Eqn. 9.92)}

These formulas give somewhat conservative results (e.g., over-estimates of required force and power) because they do not take into account special body shapes, dynamic wing motions (e.g., delayed stall, rotational circulation, and wake capture$^{3268}$), virtually complete pressure recovery in creeping flow, and so forth that may be used to improve flight efficiency in an optimized design. On the other hand, flight efficiency may be reduced by the accretion of water molecules and other environmental substances on working surfaces, and by the power expended in shearing air (dominant at the micron scale) or in imparting net kinetic energy to the air (at higher Reynolds numbers).

Table 9.5 shows force, power, and power density for spherical flying nanorobots of various sizes and airspeeds, computed using the above relations. Note that power scales as $-\nu_{\text{nano}}^2$ in the viscous regime, $-R_{\text{nano}}^2$ in the transitional regime. (The maximum sustainable velocity for translational-regime insects is ~11 m/sec$^2$.)

For an $R_{\text{nano}} = 1$ micron aerial nanorobot, taking $\eta_{\text{air}} = 0.0183 \times 10^{-5}$ kg/m·sec and $\rho_{\text{air}} = 1.205 \text{ kg/m}^3$ in dry air at 20˚C and at 1 atm, with $v_{\text{nano}} \sim 1000 \text{ kg/m}$ and $e\% = 0.10$ (10%), a $v_{\text{nano}} = 1 \text{ cm/sec}$ airspeed requires $F_{\text{nano}} \sim 4 \text{ pN}$ and a $P_{\text{nano}} \sim 0.4 \text{ pW}$ powerplant ($D_{\text{nano}} \sim 82,000 \text{ watts/m}^3$). At $v_{\text{nano}} = 1 \text{ m/sec}$, then $F_{\text{nano}} \sim 350 \text{ pN}$ and $P_{\text{nano}} \sim 3500 \text{ pW}$ powerplant ($D_{\text{nano}} \sim 8 \times 10^8 \text{ watts/m}^3$). A much larger $R_{\text{nano}} = 1 \text{ mm}$ nanorobot flying at $v_{\text{nano}} = 1 \text{ m/sec}$ requires $F_{\text{nano}} \sim 4 \mu \text{N}$ and $P_{\text{nano}} = 41 \mu \text{W}$ ($D_{\text{nano}} = 10,000 \text{ watts/m}^3$). Again, an optimized design might reduce some of the power figures by a factor of 10 or more.

Aerobot power density $D_{\text{nano}}$ scales as $-\nu_{\text{nano}}^2$ in the viscous regime. In the transitional regime, power density scales as $-R_{\text{nano}}^{-2}$ at low velocities and $-R_{\text{nano}}^{-4}$ at high velocities. Thus, to minimize power density and therefore conserve energy, circuomcoral clouds of aerial nanorobots may coalesce into progressively larger but fewer tightly-packed clumps as the collective velocity of the cloud moves to higher airspeeds, assuming that the aeromotive mechanism design is largely scale-invariant over the full size range of the progressive aggregations. This strategy is most effective in the viscous regime.

For example, consider a cloudlet consisting of 1 million nanorobots, each of size $R_{\text{nano}} = 1$ micron. With individual nanorobots traveling at $v_{\text{nano}} = 30 \text{ cm/sec}$, the cloudlet consumes ~0.4 milliwatts and operates at a power density of $-10^8 \text{ watts/m}^3$. Now assume that the cloudlet must speed up to 10 m/sec to track a fast-moving object around which it is stationkeeping, or to compensate for a heavy wind. If the individual nanorobots comprising the cloudlet simply increase their airspeed to 10 m/sec, then power density in each nanorobot increases to $-10^4 \text{ watts/m}^3$ and cloudlet power consumption rises to ~400 milliwatts (a 1000-fold increase). However, if the nanorobots temporarily aggregate into a single collective approximating a single device of $R_{\text{nano}} = 100$ microns,
The nanoflyer of similar size to impart momentum to a nearby mass of air by speeding up that mass of air to a higher velocity. The wings sweep accelerates with constant acceleration $a_{nano} = \frac{v_{nano}^2}{2X_{accel}}$ from a spherical nanorobot of radius $R_{nano}$ that uses circumferential wings are not considered further here but may be useful. into the downwind slipstream of a rapidly-moving tracked object, (Other power-conserving behaviors, such as preferential migration wide range of velocities without significantly increasing power. design-dependent but a crude generalization may be made. Consider a standing start to a final velocity $v_{nano}$ in a time $t_{accel} = \frac{v_{nano}}{a_{nano}}$ and kinetic energy imparted to the swept-out air is $KE_{air} = \frac{1}{2} M_{air} v_{air}^2$ and to the nanorobot is $KE_{nano} = \frac{1}{2} M_{nano} v_{nano}^2$, with $KE_{total} = KE_{air} + KE_{nano}$. Neglecting drag losses (e.g., ~3500 pW for $R_{nano} = 1$ micron and $v_{nano} = 1$ m/sec; Table 9.5) and assuming negligible losses within the air-accelerating mechanism itself and any drag on the air used as reaction mass, then propulsive efficiency is $e\% = \frac{KE_{nano}}{KE_{total}}$ and total power consumption is $P_{nano} = P_{drag} + (KE_{total} / e\% \cdot t_{accel})$. In the examples below we take $R_{nano} = 1$ micron, $v_{nano} = 1$ m/sec, $P_{drag} = 3500$ pW (Table 9.5), $\rho_{air} = 1.205$ kg/m$^3$ for dry air at 20˚C and 1 atm, and $\rho_{nano} = 1000$ kg/m$^3$; here the Reynolds number is near unity. A 1 micron$^3$ high-density (~10$^{12}$ W/m$^3$; Chapter 6) powerplant develops $P_{nano} = 1$ million pW. At this high power level, a nanorobot can accelerate at $a_{nano} = 12,000$ g's for $t_{accel} < 9$ microsec, reaching $v_{nano} = 1$ m/sec after crossing a running distance of $X_{accel} = 4.4$ microns with an efficiency of $e\% = (0.016) 1.6\%$ and $v_{air} = 64$ m/sec. As an alternative approach, note that a 5-microsec gas discharge from a simple pressure-release pump (Section 9.2.7.1) of volume $V_{reservoir} = 0.1$ micron$^3$ having a pressure differential of 1000 atm (storing ~10$^8$ J/m$^3$; Section 6.2.2.3) produces a mean discharge power of $P_{nano} = 2$ million pW, which allows $a_{nano} = 15,000$ g's of acceleration with $e\% = (0.01) 1\%$ for a maximum gas exit velocity of $v_{max} = 70$ m/sec using a nozzle of length $l_{tube} = 1$ micron and radius $r_{tube} = 10$ nm (Eqn. 9.36). At a more modest $P_{nano} = 4500$ pW, the nanoflyer accelerates at $a_{nano} = 1100$ g's for $t_{accel} = 96$ microsec, reaching $v_{nano} = 1$ m/sec after crossing a running distance of $X_{accel} = 48$ microns with an efficiency of $e\% = 0.15 (15\%)$ and $v_{air} = 6$ m/sec. Attitude control of flying machines is a problem often mentioned in MEMS aerobotics work;~1576 airborne nanorobots may employ

---

### Table 9.5. Conservatively Estimated Force, Power, and Power Density for Hovering and Flying Spherical Nanorobots as a Function of Size and Velocity

<table>
<thead>
<tr>
<th>$(R_{nano})$ Nanorobot Radius</th>
<th>$(v_{terminal})$ Terminal Velocity</th>
<th>(Force $F_{nano}$)</th>
<th>(Power $P_{nano}$)</th>
<th>(Power Density $D_{nano}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1 µm</td>
<td>1.2 µm/sec</td>
<td>$F_{nano}$ (N)</td>
<td>$P_{nano}$ (W)</td>
<td>$D_{nano}$ (W/m$^3$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$4.1 \times 10^{-17}$</td>
<td>$3.5 \times 10^{-13}$</td>
<td>$3.5 \times 10^{-10}$</td>
</tr>
<tr>
<td>1 µm</td>
<td>120 µm/sec</td>
<td>$F_{nano}$ (N)</td>
<td>$P_{nano}$ (W)</td>
<td>$D_{nano}$ (W/m$^3$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$4.1 \times 10^{-14}$</td>
<td>$3.5 \times 10^{-12}$</td>
<td>$3.5 \times 10^{-10}$</td>
</tr>
<tr>
<td>10 µm</td>
<td>1.2 cm/sec</td>
<td>$F_{nano}$ (N)</td>
<td>$P_{nano}$ (W)</td>
<td>$D_{nano}$ (W/m$^3$)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$4.2 \times 10^{-11}$</td>
<td>$3.5 \times 10^{-10}$</td>
<td>$3.5 \times 10^{-10}$</td>
</tr>
</tbody>
</table>

* aerodynamic lift forces available for hovering at this size scale

---

then the power consumption of the collective can be held to the original 0.4 milliwatts and power density remains constant at ~10$^8$ watts/m$^3$. Facultative aggregation may permit stationkeeping over a wide range of velocities without significantly increasing power. (Other power-conserving behaviors, such as preferential migration into the downwind slipstream of a rapidly-moving tracked object, are not considered further here but may be useful.)

How fast can nanoflyers accelerate? The answer is highly design-dependent but a crude generalization may be made. Consider a spherical nanorobot of radius $R_{nano}$ that uses circumferential wings of similar size to impart momentum to a nearby mass of air by speeding up that mass of air to a higher velocity. The wings sweep air from a circular cross-section of radius $2R_{nano}$. The nanoflyer accelerates with constant acceleration $a_{nano} = \frac{v_{nano}^2}{2X_{accel}}$ from a standing start to a final velocity $v_{nano}$ in a time $t_{accel} = \frac{v_{nano}}{a_{nano}}$ and a running distance $X_{accel}$. The mass of the swept-out air is $M_{air} = 4 \pi r_{tube} X_{accel} R_{nano}^2$ and the mass of the nanorobot is $M_{nano} = \frac{4}{3} \pi R_{nano}^3$. To conserve momentum, $M_{air} \cdot v_{air} = M_{nano} \cdot v_{nano}$, hence $v_{air} = \frac{M_{nano} \cdot v_{nano}}{M_{air}} = \frac{M_{nano} \cdot v_{nano}}{3 \cdot \rho_{air} X_{accel} v_{air} < v_{sound} = 343$ m/sec in dry air at 20˚C and 1 atm to remain subsonic (Section 9.5.3.5). The kinetic energy imparted to the swept-out air is $KE_{air} = \frac{1}{2} M_{air} v_{air}^2$ and to the nanorobot is $KE_{nano} = \frac{1}{2} M_{nano} v_{nano}^2$, with $KE_{total} = KE_{air} + KE_{nano}$- Neglecting drag losses (e.g., ~3500 pW for $R_{nano} = 1$ micron and $v_{nano} = 1$ m/sec; Table 9.5) and assuming negligible losses within the air-accelerating mechanism itself and any drag on the air used as reaction mass, then propulsive efficiency is $e\% = \frac{KE_{nano}}{KE_{total}}$ and total power consumption is $P_{nano} = P_{drag} + (KE_{total} / e\% \cdot t_{accel})$. In the examples below we take $R_{nano} = 1$ micron, $v_{nano} = 1$ m/sec, $P_{drag} = 3500$ pW (Table 9.5), $\rho_{air} = 1.205$ kg/m$^3$ for dry air at 20˚C and 1 atm, and $\rho_{nano} = 1000$ kg/m$^3$; here the Reynolds number is near unity. A 1 micron$^3$ high-density (~10$^{12}$ W/m$^3$; Chapter 6) powerplant develops $P_{nano} = 1$ million pW. At this high power level, a nanorobot can accelerate at $a_{nano} = 12,000$ g's for $t_{accel} < 9$ microsec, reaching $v_{nano} = 1$ m/sec after crossing a running distance of $X_{accel} = 4.4$ microns with an efficiency of $e\% = (0.016) 1.6\%$ and $v_{air} = 64$ m/sec. As an alternative approach, note that a 5-microsec gas discharge from a simple pressure-release pump (Section 9.2.7.1) of volume $V_{reservoir} = 0.1$ micron$^3$ having a pressure differential of 1000 atm (storing ~10$^8$ J/m$^3$; Section 6.2.2.3) produces a mean discharge power of $P_{nano} = 2$ million pW, which allows $a_{nano} = 15,000$ g's of acceleration with $e\% = (0.01) 1\%$ for a maximum gas exit velocity of $v_{max} = 70$ m/sec using a nozzle of length $l_{tube} = 1$ micron and radius $r_{tube} = 10$ nm (Eqn. 9.36). At a more modest $P_{nano} = 4500$ pW, the nanoflyer accelerates at $a_{nano} = 1100$ g's for $t_{accel} = 96$ microsec, reaching $v_{nano} = 1$ m/sec after crossing a running distance of $X_{accel} = 48$ microns with an efficiency of $e\% = 0.15 (15\%)$ and $v_{air} = 6$ m/sec. Attitude control of flying machines is a problem often mentioned in MEMS aerobotics work;~1576 airborne nanorobots may employ
gravity-based dynamic stabilization (Section 8.3.3), gyrostabilization (Section 9.4.2.2), or other means. In 1998, the energetics and control of steering maneuvers in highly-maneuverable insects was being investigated.1581,1582

9.5.3.5 Hovering Flight

In large helicopters, thrust is produced by imparting a downward velocity to the mass of air flowing through the rotor, with lift proportional to the change in momentum. In 1997, engineers at the Institute for Microtechnology in Mainz, Germany, constructed a 3-cm long, 1-cm tall mini-helicopter weighing 0.3 gm with two blades turning at ~1700 Hz, and flew it to a hovering altitude of 13 cm, then landed it safely. I. Kroo at Stanford University has built and flown a “mesicopter” whose motor is 3 mm in diameter and 5 mm long, weighing ~0.3 gm, that turns the oddly-shaped rotors at ~50,000 rpm.3246

As helicopter size shrinks through the transitional regime, drag grows at the expense of lift and the rotorblade microhelicopter becomes increasing inefficient. However, the viscous-lift helicopter (Fig. 9.22) is estimated to be able to hover by rotating its four wheels at a frequency of $v_{wheel} = 3 \rho_{nano} \rho_{wheel} / 32 \nu_{air}$ = 1 KHz, taking $\rho_{nano} = 2000$ kg/m$^3$, $g = 9.81$ m/sec$^2$, and $\rho_{wheel} = 10$ microns.1982

In the general case, the terminal velocity $v_{terminal}$ of a compact nonaerodynamic body of any size falling through fluid may be approximated by setting:

$$F_{viscous} + F_{inertial} = \frac{4}{3} \pi R_{nano}^3 \rho \nu (\rho_{air} - \rho_{nano})$$

[Eqn. 9.93]

and solving the resulting quadratic for $v_{terminal}$ ($= v_{nano}$), using $F_{viscous}$ and $F_{inertial}$ for spherical bodies from Eqs. 9.89 and 9.90. This velocity may then be used in Eqs. 9.88 through 9.92 to conservatively estimate $F_{nano}$ and $D_{nano}$ for hovering. Representative values are given in Table 9.5, assuming flight in dry sea-level air and powerplant efficiency $\eta_0 = 0.10$ (10%). Hovering power $P_{hover}$ scales as $\sim R_{nano}^{-3}$ for $R_{nano} \approx 10$ microns and as $\sim R_{nano}^{-5}$ in the $R_{nano} \sim 0.1-1$ mm range; for $R_{nano} \approx 1$ mm, aerodynamic lift forces are available, the computation of which is beyond the scope of this book. For $R_{nano} = 1$ micron in air, $v_{terminal} = 120$ microns/sec (Section 9.5.3.2) and $P_{hover} = \sim 0.0005$ pW ($\eta_0 = 0.10$ (10%)); for $R_{nano} = 10$ microns, $v_{terminal} = 1.2$ cm/sec and $P_{hover} = \sim 5$ pW ($\eta_0 = 0.10$ (10%)).

In the special case of flapping winged insects in the transitional flight regime (e.g., 100 microns $\sim R_{nano} \sim 10$ cm) during steady-state hovering, T. Weis-Fogh1578,1583 calculates that the average aerodynamic power needed for small winged objects to remain airborne is:

$$P_{hover} = \frac{2}{3} \pi R_{air}^2 \nu \tau C_D \nu_{wing} \psi^3 \nu_{wing} \nu_{wing}^{-4}$$

[Eqn. 9.94]

where $\rho_{air} = 1.205$ kg/m$^3$ at 1 atm and 20°C, $\tau$ is a shape factor that equals 0.5 for a rectangular wing, 0.1 for a triangular wing attached at the base, and 0.4 for a triangular wing attached at the apex; $C_D = 0.07-0.36$ for bats, birds, and insects with wing length of $L_{wing} = 0.25-13$ cm and wing width (chord) of $w_{wing} = 0.7-55$ mm (flight mass $M = 0.001-20$ gm); wingstroke frequency $\nu_{wing}$ ranges from 15 Hz at $L_{wing} = 13$ cm (large hummingbird) to 600 Hz at $L_{wing} = 0.25$ cm; and stroke angle $\psi$ is the angle subtended by each flapping wing in the flapping plane during a complete stroke cycle, typically 2-3 radians (120°-180°).

Thus for example, the common honeybee Apis mellifera (NR = 1900) has $M = 100$ milligrams, $L_{wing} = 1.0$ cm, $w_{wing} = 0.43$ cm, $\tau = 0.27$ (half-ellipse), $C_D = 0.09$, $\nu_{wing} = 240$ Hz, and $\psi_{wing} = 2.09$ rad, giving $P_{hover} = 1$ milliwatt with a dynamic efficiency of $\eta_0 = 30\%$; during level flight at peak speeds near ~6 m/sec,739 experimentally-measured honeybee metabolic demand is 20-60 milliwatts.1580 Honeybees can also carry a ~40 milligram payload of honey. At the lower extreme of the transitional flight regime, the parasitic chalid wasp Encarsia formosa (NR = 15) has $M = 25$ µg, $L_{wing} = 620$ microns, $w_{wing} = 230$ microns, $\tau = 0.50$ (rectangle), $C_D = 3.20$, $\nu_{wing} = 400$ Hz, and $\psi_{wing} = 2.36$ rad, giving $P_{hover} = 0.4$ microwatt. Wingspeed $\nu_{wing} > 1.5$ m/sec during the downstroke.1578

From the conservative assumption that biological wing muscles are limited to ~200 watts/kg (~300,000 watts/m$^3$), Weis-Fogh1577 estimates than no flying animal with a mass larger than ~100 grams can hover continuously by means of wing flapping and wing twisting alone.

For artificial nanorobotic flyers near the edge of the design envelope, maximum wingbeat frequency $\nu_{max} \leq \nu_{wing} / L_{wing}$. To avoid supersonic turbulence, $\nu_{wing} < \nu_{sound} \sim 343$ m/sec in air at 20°C and 1 atm, giving $\nu_{max} \leq 10$-$100$ MHz for $L_{wing} = 3$-$30$ microns and $\leq 100$ KHz for $L_{wing} = 3$ mm. For each small-wing cycle at 100 MHz, the boundary layer after an impulsive start is at most (modified from Prandtl1684) $\delta_{layer} = \delta_{air} / \rho_{air} \nu_{max}^{1/2} < 0.4$ microns - 0.2 microns, the mean free path of air molecules at 1 atm (Eqn. 9.23); faster cycling would allow insufficient time for mechanical coupling to the medium. Note also that $\nu_{sound}$ is significantly slower than the torsional deformation propagation velocity $v_{torsion} = (G / \rho_{wing})^{1/2} = 12.000$ m/sec along the wing structure, taking $G = 5 \times 10^{11}$ N/m$^2$ and $\rho_{wing} = 3510$ kg/m$^3$ for diamondoid materials. For comparison, the slowest insectile wingbeat frequency is ~5 Hz for the swallowtail butterfly (Papilio machaon); the fastest is ~1046 Hz for the tiny midge Forcipomyia in natural conditions, up to ~2200 Hz at 310 K in laboratory experiments with truncated wings.739,1033

9.5.3.6 No-Fly Zones

Most circumcorporal aerial nanorobots should take care to avoid certain well-defined zones near the human body. Perhaps the most important medical issue is that particles under 5 microns in diameter are dangerous to inhale. Concentrations of $>10^{12}$ m$^{-3}$ of free silica particles <5 microns in size are usually considered hazardous.1572 Airborne nanorobots may be deployed in number densities of $\sim 10^{12}$ m$^{-3}$ or higher (Section 7.4.8), so anti-inhalation, inhalate-safe, and post-inhalation extraction protocols are essential in this application. (Particles larger than 2-5 microns that are inspired through the nose become trapped on the nasal mucus membrane and do not reach the lower airway; Section 8.2.2.)

At rest, inhalation velocity at the trachea is $v_{inhale\,rest} \sim 0.3$ m/sec. During the heaviest exercise, turbulent airflow into the trachea peaks near ~3.3 liters/sec at a maximum inhalation airspeed of $v_{inhale\,max} \sim 5$ m/sec (Section 8.2.2 and Eqn. 9.30). Nanorobots capable of flying at $R_{nano} \approx v_{inhale\,max}$ can outrun even the fastest inspired air by simply reversing course. Nanorobots restricted to slower speeds can also avoid inhalation by promptly initiating lateral motion immediately upon entering either of two hemiellipsoidal no-fly zones which are very conservatively estimated to be of major-axial radius:

$$R_{clear} = \frac{R_{oral} \nu_{inhale\,max} \nu_{nano}}{2 \nu_{nano}} \text{ (meters)}$$

[Eqn. 9.95]

* A few other dynamic efficiency figures for hovering flight include the hornet wasp (31%), hummingbird (51%), mosquito (70%), fruit fly (95%), and butterfly (97%).1578
centered on the vestibule of the mouth and on the anterior nares of
the nose, where $R_{oral} \approx 2.5 \text{ cm}$ is the radius of the oral orifice. Thus
a nanorobot with maximum airspeed $v_{nano} = 1 \text{ m/sec}$ must observe
a perifacial no-fly major-axial radius of $R_{clear} \approx 6 \text{ cm}$.

The second most important no-fly zone is the surface of
the skin, including (in special circumstances) the oral mucous (tissue)
membrane, tongue, soft palate, larynx, trachea, and nasal passages.
It is theoretically possible for diamondoid nanorobots with rigid
protrusions that accidentally impact the epidermis at high speeds to
irritate or even tear the skin. Consider a nanorobot of density $\rho$
and dimension $L$ with a rigid appendage of dimension $qL$, that impacts
an epidermal surface of tearing strength $\sigma_{tear}$ at a velocity $v$, coming
to a halt in a distance $X_{scratch}$. If the impact energy is transferred
exclusively through the appendage, then the skin tears if:

$$X_{scratch} \leq \frac{\rho L v^2}{2 q^2 \sigma_{tear}} \text{ (meters)} \tag{Eqn. 9.96}$$

Taking $\rho = 2000 \text{ kg/m}^3$, $L = 100 \text{ microns}$, $v = 10 \text{ m/sec}$, $q/L = 0.1$, and $\sigma_{tear} = 10^7 \text{ N/m}^2$, then $X_{scratch} \approx 100 \text{ microns}$—potentially
leaving up to a 100-micron long, 10-micron deep scratch on the
skin. Long scalp hair, which often waves randomly in the wind,
should also be avoided (or actively managed; Chapter 28) by the
aerial nanorobot cloud. Nanorobots with 10-100 micron-long wings
operated at 0.1-1 MHz (avoiding subharmonics to ensure buzzless
performance) can have transverse wingtip velocities up to 1-10 m/sec,
which could damage soft tissues upon impact.

Another possible no-fly zone is the interior of the auditory
canal, especially immediately adjacent to the tympanic membrane. In
theory (Section 7.4.8), a very dense nanorobot cloud using acoustic
communications at high bandwidth could generate a pressure
intensity at the eardrum that exceeds the threshold of pain, although
this is unlikely in practice. (More interestingly, such a cloud could enter
the ear canal and emit coordinated audible subharmonics in the
100-1000 Hz range, thus "speaking" directly to the user; Section 7.4.6.3.)

Optical communication intensities are equally unlikely to
exceed safe limits in the vicinity of the human eye (Section 7.4.8).
However, entirely aside from the epidermal no-fly zone (above),
general-purpose aerial nanorobots should especially avoid all contact
with the surface of the cornea, the conjunctiva, and the inner surfaces
of the eyelid, in order to prevent irritation, or serious scoring and goug-
ing (e.g., corneal ulcer, conjunctivitis, or superficial punctate keratitis),
or embedment in these much softer, exposed mucosal tissues.
This final Chapter describes a miscellany of important technical capabilities that may prove useful in some or all medical nanodevices, in various scenarios or theaters of operation. Any one of these subjects deserves an entire chapter to itself, but unfortunately there is only space in this introductory text for a brief survey of each area. The most important of these topics is computation (Section 10.2), including nanomechanical, nano-electronic, and biological computing, as well as nanoscale data storage technologies. However, the fields of organic and fullerene nanoelectronics, biocomputing, and quantum computing are advancing so fast that whatever is written here will quickly become obsolete. Thus our coverage in this Volume is limited to a broad overview. Interested readers are strongly advised to consult the current literature for the latest results.

Other important basic capabilities covered in this Chapter include timers and nanoclocks with long-term nanosecond stability (Section 10.1); high-pressure materials storage (Section 10.3); defensive cytoidal and antiviral weaponry to allow the efficient destruction of pathogenic intruders and unrepairable tumor cells (Section 10.4); and the effects of very hot or very cold temperatures on nanorobot materials and operations (Section 10.5).

10.1 Nanochronometry

Nanorobots will use clocks in many applications. Computer gating, navigation, and high-speed sensor applications may require repeatable timing in the 1-1000 nanosec range, or 1-1000 microsec for lower-speed chemical and chemotactic sensing, although long-term clock stability is not especially critical for computing. Neuron-mediated signals and muscle motions are monitored and controlled on a 1-1000 millisec timescale, while conscious human action and most human biorhythms occur in the 1-100 sec range.

This Section briefly introduces human chronobiology (Section 10.1.1), then describes possible nanoscale oscillator systems that could be useful in nanorobotic clocks (Section 10.1.2), basic principles of pre- and post-infusion nanorobot chronometer synchronization (Section 10.1.3), and dedicated chronometer organs (Section 10.1.4).

10.1.1 Human Chronobiology

The human body incorporates numerous biological clocks. The best-known and most-studied example is the daily (24-hour) endogenous circadian oscillator. This internal clock is normally reset by natural sunlight, which is much brighter than indoor lighting. The clock, in turn, sets the cadence for most of the other 24-hour body rhythms—for example, sleep/wakefulness cycles (e.g., melatonin), urine production, body temperature, blood cortisol and ACTH cycles (Appendix B), and the diurnal rhythm of mitosis in epidermal epithelium (e.g., greatest during sleep or inactivity, least during wakefulness or activity). The circadian clock stability (variation in free-running period) is ~1% in Drosophila and ~0.2% in humans.

Mammalian circadian rhythms are regulated by a master pacemaker within the suprachiasmatic nuclei (SCN) of the hypothalamus (just above the point at which the optic nerves from each eye cross in mid-brain). In humans, the SCN is comprised of ~10,000 special cells that send out electrochemical signals in a 24-hour timing pattern. Entrainment of the clock to light-dark cycles is mediated by photoreceptors in the retina, with light information conveyed directly from ganglion cells of the retina to the SCN via the retinohypothalamic tract. Mammals and other vertebrates have another independent circadian clock in each retina, possibly driven by the retinal cryptochromes CRY1 (also made in the SCN) and CRY2, producing rhythms in local physiology such as the diurnal renewal cycles of rods and cones. CRY1 is also abundant in skin tissues. Clock-resetting photoreceptors have been found in the human popliteal region (behind the knees), and such photoreceptors may exist in many different tissues throughout the human body.

The current model for a core circadian oscillator comprises, in part, a transcription/translation-based negative feedback loop in which clock genes are rhythmically expressed, giving rise to cycling levels of clock RNAs and proteins (negative elements). The proteins then feed back, after a lag, to depress the level of their own transcripts, perhaps by interfering with positive elements that increase transcription of the clock genes. In 1998 this genetic network was still in the discovery process. One ~100,000-base-pair gene, aptly named “clock,” was known to produce an 855-residue protein (mCLOCK in mice, ~115.7 kD) that serves as a major regulator of the ~10 other, still unidentified, genes thought to affect circadian rhythm. (One portion of mCLOCK is the same as in a related 1023-residue clock protein, dCLOCK, found in Drosophila fruit flies). Circadian rhythms will almost certainly be made up of many interconnected feedback loops that interact with the CLOCK-related core pathway. It is believed that this interconnected
ensemble will ultimately determine all of the classical circadian properties—period length, temperature compensation, and resetting by light or temperature. In 1998, most chronobiologists believed that many of these outer loops would be organism specific and that only the core loop would be more universal.

In addition to the 24-hour circadian oscillator, many other biological clocks have been identified in humans. For example, the 28-day menstrual cycle is timed by two ovarian clocks, each delimiting a period of 14 days. One of these two clocks is the Graafian follicle and its production of estradiol, while the other clock is the corpus luteum and its secretion of progesterone. Both are obligatorily dependent upon the proper functioning of a third clock located in the arcuate region of the hypothalamus, known as the GnRH pulse generator, which ensures the rhythmic production and release of the gonadotropin hormones into the peripheral circulation with a cyclical period of 1-1 hour (±25% stability), coincident with a 1-hour cycle of hypothalamic electrical activity (±12% stability). A 90-minute clock paces the development of the somites, blocks of tissue that form in regular arrays along the spinal cord of vertebrate embryos. One regulatory gene, named “chaire,” undergoes repeated 90-minute activity cycles and is known not to require protein synthesis; gene expression follows the repeating pattern even when protein synthesis is biochemically blocked. A much longer half-weekly (circasemiseptan) pattern has been observed in mitotic activity in cancer patients.

Other biological clocks abound. The stomach produces rhythmic contractions of ~3-minute duration during digestion. Rhythmic segmentation contractions of the bowel at ~0.2 Hz (Section 8.2.3) are driven by a thin layer of specialized pacemaker cells called the intestinal cells of Cajal, which demonstrate electrical oscillations at about the same rate as the contractions. Cardiac pacemaker cells in the sinoatrial node emit a cyclical pulse that triggers a heartbeat at a ~1 Hz basal rate. The respiratory centers in the medulla oblongata produce an autonomic contraction cycle of the diaphragm at ~0.3 Hz. The human eye produces many spontaneous rhythmic dilations and contractions, including high-frequency ocular microtremors present in all people (mean frequency ~88 Hz, accommodation microfluctuations (~1-2 Hz) with the pupil varying <300 microns in diameter, and low-frequency pupillary oscillations (e.g., ~0.2 Hz “hippus”). Electrical rhythms generated by the brain that are detectable by an electroencephalograph include alpha waves predominantly in the occipital region at 8-13 Hz, beta waves in the frontal and central areas at 18-30 Hz, delta waves at <4 Hz during deep sleep or abnormal function, and theta waves in the temporal and parietal areas at 4-7 Hz.

Individual cells display a number of oscillatory biochemical pathways as well, including calcium ion oscillations, cyclic AMP signalling, and various other cell cycles. For instance, the glycogenic enzyme oscillator, controlled primarily by phosphorylkinase, may mediate various rhythmic physiological behaviors such as slow waves of contraction in smooth muscle, electrical activity in neurons, and insulin release from β islet cells of the pancreas. In one experiment, this oscillator mediated a 55 mV voltage change in guinea pig cardiomyocytes at 0.010 Hz (~10% stability) in an almost perfectly sinusoidal pattern, although a few cells displayed oscillations with irregular phase, amplitude, or both. In another experiment, fluorescently-labeled puffs of tumor-cell cytoplasm were observed being released through plasma membrane ruffles, regularly at ~0.05 Hz during a neutrophil-mediated cytolytic attack, matching the nicotinamide-adenine dinucleotide phosphate and superoxide release periods for neutrophils. By 1998 there was a growing belief that many, if not most, of the cells in an animal may possess individual biological clocks. Fruit flies were already known to have clocks distributed throughout their wings, legs, and abdomen.

The human brain also possesses an interval timer that allows a person to gauge the passage of seconds or minutes to a mean accuracy of ±15%. The interval timer functions like a stopwatch and resides in the basal ganglia, a region of the brain that coordinates voluntary muscle movements. A population of neurons in the substantia nigra releases regular pulses of dopamine into an accumulator, called the caudate-putamen, a major part of the basal ganglia. These pulses are then read via neural pathways to the cerebral cortex which are known as striato-cortical loops. Other internal “alarm clocks” that reliably allow waking from sleep at desired times have been described.

In vivo nanorobots should be able to eavesdrop on most of these active chronobiological channels, gaining complete knowledge of the oscillator frequencies and phase settings of virtually all of the human body’s biological clock systems. For example, blood plasma cortisol typically peaks in the morning (e.g., 6-23 x 10^3 gm/cm^3 at 8 AM), then declines throughout the day (e.g., 3-15 x 10^3 gm/cm^3 at 4 PM, then ~50% of 8 AM value at 10 PM). Melatonin (also detectable in saliva) cycles in the opposite direction, normally peaking at 6.7 x 10^-11 gm/cm^3 at night and then falling to 1.4 x 10^-11 gm/cm^3 during the day (Appendix B). Longer cycles are easily tracked too. For instance, blood viscosity in healthy young women shows the rhythmic variation of the menstrual cycle, probably due to changes in serum fibrinogen and globulin levels which can be measured more directly. These nanorobot-accessible readings may be used either to measure “body time” or to control it, and in some cases can recalibrate relevant onboard nanochronometers.

Environmental time cues may also be detected by in vivo nanorobots (e.g., temporal macrosensing), such as the level of illumination penetrating the epidermis (sometimes permitting inference of day or night) or the all-pervasive 60/50 Hz (U.S./Europe) electromagnetic hum generated by alternating current electrical supply systems in common use throughout the world, a crude frequency standard whose voltage can sometimes display time-of-day load-related fluctuations.

10.1.2 Artificial Nanoscale Oscillators

Nanorobot clocks must perform at least two critical functions: (1) Interval timing (measuring elapsed time between two or more sequential events) and chronometry (maintaining an onboard calendrical “standard time” in good calibration with an external time standard such as Coordinated Universal Time). Complete clock design, including clutches, feedback mechanisms and governors, dampers and filters, triggers and strikers, amplifiers and mixers, phase detectors and counters, regenerative frequency dividers and frequency multipliers, and antivibration housings, is very complex and quite beyond the scope of this text. However, both timing and chronometry require a high-frequency oscillator of known and stable frequency, whose oscillations can then be counted for clocking purposes. If frequency multipliers can be built, then the primary oscillator need not be high frequency; the requirement for frequency stability is more fundamental. Timing accuracies of 1-1000 nanosec are typically needed for sensor (Chapter 4), navigation (Section 8.3.3), and

*Long ago, native American Indians of the Sioux tribe learned, after carefully controlled experiments, that they could use a full urinary bladder as a kind of alarm clock. By drinking a certain amount of water at bedtime, they found they would awaken in a specific number of hours—drink more fluid, and they would awaken earlier from the urge to urinate; drink less, and they would awaken later.*
computational (Section 10.2) applications, with operating frequencies up to \( v_{osc} = 1 \) GHz and durations ranging from 1 microsec up to 10^9 sec (~1 day) between timed events or between clock recalibrations (Section 10.1.3). This section presents a brief survey of several useful nanoscale oscillator systems.

10.1.2.1 Mechaenochemical and Photochemical Oscillators

A chemical oscillator involves transitions between two well-defined chemical, energetic, or conformational states. For example, repetitive room-temperature protein folding such as \( \alpha \)-helical coiling may occur in ~10^{-6} sec (e.g., \( v_{osc} = 1 \) MHz).\(^{467}\) Turnover number (molecules of substrate converted to product per active site per unit time) for catalase, one of the fastest known enzymes, is \( k_{cat} = v_{osc} = 40 \) MHz.\(^{759}\) Reversible gas phase reactions may occur at up to 5 GHz.\(^{390}\) However, chemical-based clocks are expected to display poor frequency stability. Most chemical reactions display \( \sim \) \( e^{-t/\tau} \) rate dependence, giving a frequency stability of \( \Delta \nu / \nu \sim 1\% \) for a ~1 K temperature (T) variation in the oscillator near 310 K. Chemically-driven mechanical oscillators such as biological cilia display a similar temperature dependency and frequency stability.\(^{695}\) Time measurements that depend on the rates of thermally activated bulk reactions also suffer Poisson noise in the reaction rate from the statistics of interaction of many random independent molecules.

Some molecular transitions are even faster. Molecular dynamics simulations of carbonmonoxy-myoglobin show periodic openings and closings of internal voids on a time scale of ~100 picosec\(^{1693}\) and \( \sim 0.5 \) nm physical motion detectable by displacement nanosensors.\(^{1694}\) Time measurements that depend on the rates of thermally activated bulk reactions also suffer Poisson noise in the reaction rate from the statistics of interaction of many random independent molecules.

Numerous alternative mechanical oscillators are readily conceived:

A. Tuning Fork — A diamondoid tuning fork with tines of length \( L_{fork} = 100 \) nm, half-thickness \( R_{fork} = 10 \) nm, density \( \rho_{fork} = 3510 \) kg/m\(^3\) and Young’s modulus \( E_{fork} = 1.05 \times 10^{12} \) N/m\(^2\) has a natural vibrational frequency of:\(^{1697,1698}\)

\[
u_{osc} = \frac{R_{fork} E_{fork}^{1/2}}{\pi^{1/2} P_{fork}^{1/2} L_{fork}^{1/2}} \sim 6 \text{ GHz}
\]

[Eqn. 10.1]

In 1998, Sandia National Laboratories began offering a product line of electrostatically-driven ~1 MHz polysilicon “tuning fork” microresonators.

B. Helical Spring — A helical spring with spring constant \( k_{s} = 10 \) N/m and mass \( m_{spring} = 2 \times 10^{-19} \) kg (~40-nm-edge diamond cube) has a natural vibrational frequency of:

\[
u_{osc} = \frac{1}{2 \pi \left( \frac{k_{s}}{m_{spring}} \right)^{1/2}} \sim 1 \text{ GHz}
\]

[Eqn. 10.2]

ignoring gravity and spring mass.

C. Circular Membrane — A flexible, thin circular membrane of radius \( R_{membrane} = 100 \) nm, uniform density \( \rho_{membrane} = 3510 \) kg/m\(^3\), and thickness \( h_{membrane} = 10 \) nm, if clamped at its boundary and stretched by a tension \( F_{membrane} \) (the force per unit length anywhere in the membrane), has characteristic frequencies of vibration of:\(^{1698}\)

\[
u_{osc} = \frac{\beta}{2 R_{membrane}^{1/2} h_{membrane}^{1/2} \rho_{membrane}^{1/2}} \sim 8 \text{ GHz}
\]

[Eqn. 10.3]

where \( F_{membrane} \sim 100 \) N/m for a 10-nm thick diamondoid sheet stretched to near a conservative working stress of \( -10^{10} \) N/m\(^2\), and \( \beta \) is a constant of order unity, derived from the roots of Bessel functions for various diametral and circular vibrational modes.

D. Torsional Pendulum — Another oscillator is the torsional pendulum, wherein a diamondoid rod of radius \( R_{rod} = 35 \) nm, length \( l_{rod} = 1900 \) nm, density \( \rho_{rod} = 3510 \) kg/m\(^3\), and shear modulus \( G_{rod} = 5 \times 10^{11} \) N/m\(^2\), is clamped at one end and the other end twists around the longitudinal axis in vacuo at a characteristic oscillation frequency: \(^{1694}\)

\[
u_{osc} = \left( \frac{G_{rad}}{4 \pi^{2} L_{rod}^{1/2} \rho_{rod}^{1/2}} \right)^{1/2} \sim 1 \text{ GHz}
\]

[Eqn. 10.4]

If the end of this rod twists through an amplitude \( \Delta X \), then the time averaged power loss due to shear radiation\(^{10} \) is \( P_{rad} = \pi R_{rod}^{6} \Delta X^{2} G_{rod}^{3/2} / 256 L_{rod}^{6} \rho_{rod}^{1/2} \) ~ 26 pW for \( \Delta X \sim 95 \) nm, a rod strain of \( \Delta X / L_{rod} \leq 5\% \) and an oscillator power density of \( D_{rod} = P_{rad} / \pi R_{rod}^{2} L_{rod} = 4 \times 10^{8} \) watts/m\(^3\). The stored torsional energy is \( E_{rad} = (1/2) k_{torison} \theta^{2} = \pi G_{rod} \Delta X^{2} R_{rod}^{-2} / 4 L_{rod}^{3} \), so the characteristic decay time for the oscillation in the exemplar system is:

\[
t_{decay} = \frac{E_{rad}}{P_{rad}} = \frac{64 L_{rod}^{3/2} G_{rod}^{1/2}}{R_{rod}^{4} \rho_{rod}^{1/2}} \sim 0.1 \text{ sec}
\]

[Eqn. 10.5]
giving plenty of access time for clock resonance driver mechanisms. Aside from exogenous noise sources, which can be considerable in magnitude but may be filtered out by good design, one important endogenous source of uncertainty in $v_{osc}$ is the elastic longitudinal displacement ($\Delta L$) for the end of a thermally excited rod that randomly alters rod length, slightly changing the frequency. For bulk diamond at $T \sim 300$ K, and ignoring entropic contributions, the $\Delta L \sim 10^{-10}$ $L_{rod}^{1/2}$ / $R_{rod}$ (Tables 5.8 and 5.16 in Drexler10). Since $v_{osc} \sim L_{rod}^{-1}$, then $\Delta v / v \sim \Delta L / L_{rod}$ and so:

$$\frac{\Delta v}{v} = \left( \frac{k_{start}}{L_{rod} R_{rod}^2} \right)^{1/2} \sim 1 \times 10^{-6} \quad [Eqn.10.6]$$

for $k_{start} \sim 2.5 \times 10^{-33}$ m$^4$, $L_{rod} = 1900$ nm, and $R_{rod} = 35$ nm. Another systemic source of uncertainty is rod length variation due to the slowly temporally- and spatially-varying ambient human body temperature. Given a coefficient of volume expansion $\beta = 3.5 \times 10^{-10}$ K$^{-1}$ for diamond,567 a typical variation of $\Delta T \sim 3$ K per $t_{circ} \sim 60$ sec circulation time for bloodborne nanorobots, and a thermal recalibration time of $t_{recalib} \sim 1$ sec (to correct for temperature dependence), then $\Delta v / v \sim t_{recalib} \Delta T / L_{rod} \sim 2 \times 10^{-7}$. J. Soreff notes that yet another limitation on oscillator accuracy is the 1/Q resonance width, which is a function of the design details.

Clock accuracy may be defined by a time measurement error over a single observation cycle ($N_{obs} = 1$) of $\Delta v_{error} = t_{error} - t_{clock}$, where $t_{error} = (v_{rod} + 1)$ and $t_{clock} = v_{osc} (1 - 1/2 / v)$. Time measurement errors over multiple cycles (e.g., $N_{obs} = v_{osc} t_{obs} > 1$), spanning a total observation time $t_{obs}$ between clock recalibrations, are randomly distributed around zero but do not sum to zero, constituting instead a random walk with a maximum excursion of:

$$t_{error} - N_{obs}^{1/2} \Delta v_{error} = \frac{t_{obs}}{v_{osc}} \left( \frac{\Delta v}{v} \right) \left( 1 + 1/2 / v \right) \quad [Eqn.10.7]$$

This assumes that the errors are uncorrelated—e.g., for thermal fluctuation errors, individual observation times must lie at least $t_{EQ}$ apart, where $t_{EQ}$ is the time required for thermal equilibration with the environment (Eqn.10.24). For the exemplar oscillator with $t_{error} = 1$ nanosec between clock recalibrations, are randomly distributed around zero but do not sum to zero, constituting instead a random walk with a maximum excursion of:

1. changes in the velocity of sound due to thermal variations, and;
2. changes in acoustic path length due to elastic longitudinal displacements of thermally excited transmission rods.

First, the speed of transverse sound waves in an isotropic elastic medium having Poisson’s ratio $c_{Poisson}$ (-0.1 for diamond) is given by:

$$v_{sound} = \left( \frac{E}{\rho} \right)^{1/2} \left[ \frac{1 - c_{Poisson}}{(1 + c_{Poisson})} \right]^{1/2} \sim 17,300 \text{ m/sec}$$

where Young’s modulus $E = 1.05 \times 10^{12}$ N/m$^2$ and density $\rho = 3510$ kg/m$^3$ for diamond. In addition to the thermal dependency of $E$, $\rho$ varies as $(1 + \beta_{thermal} T)^{-1}$ where $T$ is temperature, because volume changes according to the volume coefficient of thermal expansion $\beta_{thermal} = 3.5 \times 10^{-8}$ K$^{-1}$ for diamond, 1.56 x 10$^{-5}$ K$^{-1}$ for sapphire. In an uncorrected oscillator system, $\Delta v / v \sim \Delta v_{sound} / v_{osc} = (1/2) \beta_{thermal} \Delta T / L_{rod} \sim 10^{-5}$ for diamondoid transmission lines, assuming that $\Delta T = 6$ K temperature variations are typically encountered inside the human body (Table 8.11). Correcting oscillator timing using independent temperature sensors accurate to $\Delta T_{min} / T \sim 10^{-6}$ (Section 4.6), and ignoring other possible sources of frequency instability (which may be significant), could reduce measurement $\Delta T$ to -310 microkelvins at $T \sim 310$ K, thus improving $\Delta v / v$ significantly for this source of frequency instability.

Second, longitudinal displacements $\Delta L / L_{rod} \sim 10^{-4} - 10^{-5}$ at 300 K for rods of length $L_{rod} = 1.73-17.3$ microns and cross-section $A_{rod} \sim 30$ nm$^2$ (Figure 5.8 in Drexler10). J. Soreff observes that $\Delta L / L_{rod} = (kT / E L_{rod} A_{rod})^{1/2} \times 10^{-6}$ for -0.01 micron$^3$-volume systems. For the zeroth longitudinal vibrational mode,10 these displacements occur on a timescale $v_{osc}^{-1} = 4 L_{rod} (\rho / E)^{1/2} = 0.4-4$ nanosec, comparable to signal transit times, taking $\rho$ and $E$ for diamond as above and $L_{rod} = (1730)$ nm with $n = 1, 2, ..., n_{line}$. This may restrict $\Delta v / v \sim 10^{-6}$ unless transmission lines can be further rigidized by end-clamping, sheathing, or latticed bracing at intervals, all of which, in effect, enlarge $A_{rod}$.

10.1.2.4 Quartz Resonators

Piezoelectric quartz crystals, when cut in a predetermined manner and mounted so that two opposing faces have electrical contacts, will compress or expand when an oscillating electric field is applied across the faces.701,1702 The amplitude of crystal vibration falls off very rapidly as the applied electric field frequency deviates from the resonant frequency of the crystal. (Cyclical mechanical compressions at the resonant frequency also induce oscillating electric fields of maximum amplitude.) In 1998, most high-precision timing devices were based on quartz crystal resonators, with ~2 billion resonators manufactured annually worldwide for oscillator, clock, and filter
applications. Commercially-available off-the-shelf GHz devices (e.g., Micro Networks’ M101 resonator) typically displayed ~10 ppm stability, or \( \Delta \nu / \nu = 10^{-5} \), readily improved to ~10^{-7} in temperature-controlled ovens.

The fundamental mode resonant frequency of a quartz crystal plate of thickness \( d_{\text{quartz}} \approx 1 \text{ micron} \) (~2000 molecular layers thick), density \( \rho_{\text{quartz}} = 2650 \text{ kg/m}^3 \) and elastic (Young’s) modulus \( E_{\text{quartz}} = 1.1 \times 10^{11} \text{ N/m}^2 \) (Table 9.3) vibrating in thickness mode is: \( \nu_{\text{osc}} \approx \frac{1}{2d_{\text{quartz}}} \left( \frac{E_{\text{quartz}}}{\rho_{\text{quartz}}} \right)^{1/2} \sim 3 \text{ GHz} \) \[\text{Eqn. 10.9}\]

The resonant frequencies of the highest-quality, lowest-noise quartz microresonators can be measured with a precision of ~1\( \times \)10^{15} GHz. 

10.1.3 Nanorobot Synchronization

Ideally, the clocks of all nanorobots present in a patient’s body will be synchronized to some universal time. This time setting may be defined by the patient, by the physician, or by reference to some external standard. As part of a chronometry-critical nanorobot installation procedure prior to infusion, an initializing signal can be transmitted throughout the infusate volume, allowing each individual nanorobot present therein to receive the signal and set its clock.

For the highest pre-infusion synchronization accuracy, optical (Section 10.7.3) or rf (Section 10.7.1) pulses will be employed. Electro-magnetic energy travels at \( c = 3 \times 10^8 \text{ m/sec} \) and so the container volume until intercepted by a nanorobot optical sensor, giving \( \Delta \nu_{\text{error}} / \nu_{\text{signal}} = 0.05 \text{ msec} \). Allowing 100 green photons (360 zJ/photon at 550 nm) per nanorobot at a nanorobot number density of \( 10^{12} \text{ cm}^{-3} \), the flash injects \( -40 \text{ J/m}^3 \) of energy into the container, raising the water temperature by ~10 microkelvins after this energy is fully thermalized. A 1-nsec flash produces a peak intensity of \( -4 \times 10^{10} \text{ watts/m}^2 \), comparable to monochromatic optical tweezers which are tolerated by biological macromolecules.

Some nanorobots inevitably will miss the signal and may later synchronize, prior to infusion, via close or direct physical contact with others who received the signal, or else multiple synchronization pulses can be used.

Onboard nanorobot clocks can also be synchronized post-infusion. For example, a pressure cuff inflated around the patient’s arm can administer acoustic synchronization pulses to ~microsec precision as bloodborne nanorobots pass through the

In 1998, the miniaturization of rubidium AFS for space applications was being actively studied along with newer approaches such as optically-pumped cesium AFS using solid state diode lasers (thus eliminating the bulky magnets) diode laser-pumped rubidium AFS in which the Rb discharge lamp is replaced with a ~100% efficient diode laser tuned to the correct transition frequency, \( Hg^+ \) “optical clocks.” Optical pumping methods using diode lasers defined the NIST-7 atomic-beam standard of \( \Delta \nu / \nu = 5 \times 10^{-15} \), starting in 1993.

A detailed analysis of micron-size atomic clocks is beyond the scope of this book. The possibility cannot be ruled out, but in 1998 the feasibility was unknown. From Eqn. 4.50, the minimum sensor capable of detecting a spin transition requires \( N_{\text{min}} \approx 7 \times 10^6 \text{ Cs atoms} \) (~0.001 micron \( \text{m}^3 \)) to display bulk \( \Delta \nu / \nu = 9.192,631,770 \text{ GHz} \) and electron spin angular momentum \( L_{\text{electron}} = L_{\text{proton}} \). However, many problems must be overcome including interactions involving spontaneous decay, measurement-induced transitions, phase changes due to gas molecule collisions, spatial confinement effects, the use of symmetries and cancellation of couplings. J. Soreff [personal communication, 1998] suggests consideration of an oscillator structure in which a phosphorus atom is covalently bound to a tetrahedral support of carbyne rods extending to the ends of an evacuated chamber. The fifth valence electron on the phosphorus atom should have a hyperfine interaction with the \( 31P \) nucleus, but it is presently unknown how tightly coupled the thermal vibrations in the rods are to the hyperfine state transition, since many coupling modes may vanish by symmetry.
blood vessels within; after signaling for several mean blood circulation times (e.g., a few minutes), most bloodborne nanorobots should be properly synchronized. Whole-body acoustic transmissions at 0.1-1 MHz may be generated by operating tables, vibrating chairpads, wristwatch transmitters and the like (Section 6.4.1). Such methods may produce synchronization errors of -100 microsec over 15-cm anteroposterior path lengths, reducible to ~1 microsec error using \( N_{\text{obs}} = 10,000 \) repetitions of the calibration signal (Eqn. 10.7). For higher precision, the physician transmits \(-\)MHz radio wave pulses into the body; such pulses may be detected by appropriate onboard receivers (Section 6.4.2). Such waves have only ~75% attenuation over 15-cm anteroposterior path lengths (Eqn. 6.32), producing 0.5 nsec synchronization accuracy. Optical photons applied to the skin surface are subject to scattering, producing signal pulse broadening which increases synchronization error, and to absorption, which makes signal reception difficult or impossible beyond tissue depths of a few centimeters (Section 4.9.4).

Another approach is to inject a relatively small number of chronocytes—mobile communicators (Section 7.2.6) modified to include an onboard nanoclock of high precision (e.g., a portable frequency standard) and the ability to transmit clocking synchronization signals to calibrate neighboring nanorobots. These signals may be transmitted acoustically at close proximity to the nodes of a mobile communications network (Section 7.3.2), subsequently allowing all nanorobots within 100 microns of a calibrated node to synchronize to within \( \leq 67 \) nsec. Averaging algorithms can be employed to synchronize uncalibrated nodes in the system that were not recently visited by a chronocyte. Bloodborne chronocytes will normally be carried from the capillary vasculature to within ~100 microns of most tissue locations, thus allowing \( \leq 67 \) nsec single-pulse acoustic synchronization even in the absence of an installed communications network (Section 7.3).

### 10.1.4 Dedicated Chronometer Organs

Just as the hypothalamus is the physical locus of the natural circadian clock (Section 10.1.1), it may be convenient to establish artificial chronometer organs inside the human body for various purposes. Dedicated nano-organs have been described previously in connection with power (Section 6.4.4), communication (Section 7.3.4), and navigational (Section 8.3.6) systems.

Chronometer organs, which would likely be millimeter-scale or smaller, could be used as endogenous clocks to regulate the precise administration of time-release substances or devices. Such organs could serve as highly accurate timers to improve the natural human interval timing sense from ±15% accuracy in biological systems (Section 10.1.1) to parts per million accuracy, or better, using artificial systems. If linked to the patient or user through various outmessaging channels (Section 7.4.6), chronometer organs can provide a continuously-available, consciously-accessible unfailing "time sense" of extraordinary precision for time of day, calendar date, and other time-related information.

Dedicated chronometer organs can be used to synchronize chronocytes or mobile communication networks at the nodes (Section 7.3.2). For example, chronocytes with lower-stability crystal oscillators could be repeatedly recalibrated as they passed near larger embedded chronometer organs during each blood circulation cycle (e.g., once per minute). Such chronometer organs might incorporate higher-stability onboard atomic frequency standards. Alternatively, chronometer organs could serve as transdermal data ports through which timing synchronization signals can be injected into fiber-based in vivo communication networks (Section 7.3.1) from external timing sources, similar to the autoclock pulse (containing time and date information) that is added to most contemporary television broadcast signals. Hard connectors, centimeter-scale rf wireless antennas, and other types of links are possible (Sections 6.4.2 and 7.2.3). In 1998, a desk clock could be purchased for $50 that automatically recalibrated itself by picking up WWVB radio signals anywhere in North America, several times a night, using a few centimeters of rod antenna. \(^{1711}\) WWVB’s carrier frequency is regulated to \( \Delta \nu / \nu \approx 5 \times 10^{-12} \), ultimately providing a time synchronicity at the receiver of ~100 microsec/day. With a slightly larger in vivo antenna, satellite GPS signals may be received, providing \( \pm 20 \) nanosec time error. \(^{1711}\) If buildings and vehicles are rewired to permit dissemination of time synchronization pulses and other useful information via continuous rf or IR channels, chronometer organs could employ smaller in vivo antennas and still achieve very precise results.

Although it has been informally speculated \(^{3022}\) that nanorobots could use biological nerve fibers as rf antennas to receive electromagnetic time recalibration signals from outside the human body, such as from WWV, this concept is probably unworkable for several reasons. First, the axoplasm is only \(~10^{-3}\) as electrically conductive as a metal wire of equivalent size, because axoplasmic charge carrier density and mobility are much lower than for electrons in a wire. \(^{799}\) Second, passively conducted axonal currents are quickly attenuated by leakage through ion channels in the membrane, which is a very poor insulator. From cable theory applied to neurons, \(^{799,3022}\) the length constant \( \lambda = d_{\text{axon}} \) is defined as the distance over which an applied potential depolarizes to \( \frac{1}{e} \) (~37%) of its maximum value: \( \lambda = d_{\text{axon}} \left( \frac{r_{\text{mem}}}{4 \pi r_{\text{axo}}} \right)^{1/2} \approx \left( 0.04 d_{\text{axon}} \right)^{1/2} \), where \( d_{\text{axon}} \) is axonal diameter in meters, \( r_{\text{mem}} \) is the specific membrane resistance (~0.2 ohm-m\(^2\) in human neurons) and \( r_{\text{axo}} \) is the specific resistance of the axoplasm (~1.25 ohm-m in human neurons). For a typical human axon with \( d_{\text{axon}} \approx 1 \) micron, then \( \lambda \approx 200 \) microns (~internodal distance between nodes of Ranvier) and an external signal is >99% attenuated in ~1 mm of longitudinal travel. Third, short pulses (e.g., rapidly-oscillating rf signals) are severely distorted and attenuated by the electrical capacitance of the cell membrane. The capacitive time constant in human nerves and muscle cells ranges from 1-20 millisecond, \(^{799}\) thus limiting any possible electromagnetic reception to frequencies under 50-200 Hz, although pulsed microwaves directed through the brain can induce auditory effects in animals and humans. \(^{3473-3481}\)

### 10.2 Nanocomputers

Many important medical nanorobotic tasks will require computation during the acquisition and processing of sensor data, the control of tools, manipulators, and motility systems, navigation and communication, and during the coordination of collective activities with neighboring nanorobots. Ex vivo computation has few theoretical limits, but computation by in vivo nanorobots will be subject to a number of constraints such as physical size, power consumption, onboard memory and processing speed.

The memory required onboard a medical nanorobot will be strongly mission dependent. Recognizing and manipulating molecules is fundamental. A very simple mission might demand only the identification or handling of perhaps 1-10 different molecules. For example, basic respiratory gas transport nanorobots such as respirocytes \(^{400}\) (Chapter 22) may require the operation of fixed-shape receptors that bind simple molecules such as \( \text{O}_2, \text{CO}_2, \text{H}_2\text{O}, \) and glucose. Identifiers for such receptors may need only a few bits, hence this memory requirement should be negligible. A toxin removal device (Chapter 19) similarly may require keeping track of only a few types of fixed-shape receptors. On the other hand, a survey or assay mission might need to recognize \( N = 100-1000 \) distinct proteins. A spherical
1-micron nanorobot can have $>10^4$ fixed-shape receptors on its surface; if these will suffice, then we require $N \log_2 (N) \sim 10^6$ bits to identify each of $N = 1000$ different receptor types. This seems more efficient than using more advanced reconfigurable receptors (Section 3.5.7.4), which might need $>10^9$ bits per receptor-pattern to specify each binding site geometry to the necessary atomic-scale resolution (Section 3.5.7.5), thus imposing a total memory requirement of $>10^7$ bits for an onboard library of $N = 1000$ different receptor types. Consequently, simple missions involving basic process control with limited motility may require no more than $10^3-10^6$ bits of memory, comparable to an old Apple II computer (including RAM plus floppy disk drive). At the other extreme, a complex cell repair mission might require the onboard storage of a substantial fraction of the patient's genetic code, representing $10^7$ bits of memory including perhaps $0.2 \times 10^8$ bits of linear sequence data for all 100,000 protein types found in the human body, again assuming 300 amino acids per protein. (Most amino acids are folded into the protein's interior and are not readily accessible to surface probing unless the protein is unfolded, which usually is not desirable or convenient. On the other hand, binding sites for large molecules should be physically easier to construct than small-molecule receptors; Section 3.5.9.) An onboard memory of $10^8-10^{10}$ bits would be in the same range as the 1985 Cray-2 (2 x $10^{10}$ bits) or the 1989 Cray-3 (6 x $10^8$ bits) supercomputers.\(^1\)

Computational speed will also be strongly mission dependent. However, extremely simple process control systems in basic factory settings may only require speeds as slow as $10^4$ bits/sec (Chapter 12). Individual natural biocomputational devices (as opposed to multiple such bio-devices operating in parallel) generally do not exceed this speed. Examples include mRNA translation during protein manufacture at ~15 Hz (~75 bits/sec assuming 5 bits/protein);\(^{997}\) transcription from DNA by RNA polymerase at ~40 Hz (~80 bits/sec at 2 bits/nucleotide);\(^{997}\) DNA replication at ~800 Hz (~1600 bits/sec at 2 bits/nucleotide),\(^{997}\) typically 5-100 Hz (bits/sec) for neural electrical discharges (Section 4.8.6); ~1000 Hz (bits/sec) for excitatory cholinergic synapses, and gated ion channels at ~$10^2$ Hz (bits/sec) (Section 7.4.5.6). At the other extreme, a processing speed of $10^8$ bits/sec allows a $10^9$ bit genomic information store to be processed in ~1 sec, the small-molecule diffusion time across an average 20-micron wide cell. In 1998, personal desktop computers capable of $10^9$-10$^{11}$ bits/sec ($10^8$ operations/sec) were commonly available.

This Section describes possible nanomechanical (Section 10.2.1) and nanoelectronic (Section 10.2.2) computers, biocomputers (Section 10.2.3), and briefly examines the ultimate limits to computation including reversible and quantum computing (Section 10.2.4). Computer architectural issues are not explicitly addressed here. (For example, advanced architectures, especially large CPU systems, may employ distributed clocks and a distributed computing network.)

### 10.2.1 Nanomechanical Computers

Electronic computers were evolutionarily preceded by purely mechanical computational devices, starting with the venerable abacuses (hand-operated movable beads on rods) in ~3000 BC.\(^{1726,1727}\) the Pascaline (the first hands-on algorithm-executing machine) in 1642,\(^{1728}\) and the Difference Engine (the first hands-off algorithm-executing machine) designed in 1821 by Charles Babbage\(^ {1728,1731,1744}\) A 2000-part working subsection of the brass-geared Engine was demonstrated in 1832; an entire working Difference Engine was reconstructed by historians in 1991, proving that Babbage’s design was sound.\(^ {1732}\) In the 1840s, Thomas Fowler built and exhibited a calculating device using sliding rods made of wood instead of metal.\(^ {1733}\) Whereas Babbage’s engines used the familiar 0-9 decimal system with each number represented by a discrete position of a rotating gear wheel, Fowler’s machine was more fully digital, using as its active element not rotating wheels but sliding “trinary” rods which could occupy only one of three positions at any time, the first known example of “rod logic.” (By reducing the number of distinct physical states, parts could be made less precisely.)

By 1834, Babbage had also conceived detailed plans for his Analytical Engine, intended as a general-purpose programmable computing machine but based entirely on 19th century mechanical technology. The Analytical Engine was to have a random-access memory consisting of 1000 words of 50 decimal digits each (~175,000 bits), with separate memory and central processing unit (CPU), stored program control, data entry via punched metal cards, and even an output printer.\(^ {1728,1730,1732}\) This ambitious device, though well specified, was never built.

The mechanical computing tradition was not entirely abandoned. Vannevar Bush built his analog mechanical computer, the Differential Analyzer, in 1930 at MIT.\(^ {1736}\) In 1954, M. Minsky and R. Silver\(^ {289}\) used hydraulic logic elements to build a mechanical “hydroflip computer” which was operated at ~30 Hz and powered by a 3-inch high column of water. (Section 9.2.7.6 describes the basis for a mechanical fluidic computer operating at ~5000 Hz.) In 1975, D. Hillis and B. Silverman\(^ {1738}\) built a special-purpose all-mechanical computer -2 meters in size, entirely out of Tinkertoys, and powered by a hand crank, that was able to play tic-tac-toe. In 1990, University of Minnesota engineers\(^ {1734}\) fabricated a complete family of micromechanical digital logic devices, including electrostatically-actuated linear-sliding -30-micron mechanical logic elements confined to a one-dimensional track, forming NAND and NOR gates suitable for low-speed radiation-hard digital functions “in environments hostile to electronic devices.” Sandia’s pin-in-maze microlocks\(^ {2356}\) could also be used to make a mechanical computer, albeit at the above-micron scale. In 1996, J. Gimzewski and colleagues\(^ {1735}\) at IBM Zurich used a scanning tunneling microscope (STM) probe to repeatedly reposition spherical C$_{60}$ fullerene molecules ~1 nm in diameter along a terraced copper substrate that constrained the buckyballs to move only in a straight line, operating the mechanical array like beads on an abacus. In 1997, Stoddart’s group\(^ {1740}\) described a mechanical XOR gate based on their “molecular shuttles.”

Perhaps the best-characterized (though not yet built) mechanical nanocomputer is Drexler’s rod logic design.\(^ {10,2282}\) In this design, one sliding rod with a knob intersects a second knobbed sliding rod at right angles to the first. Depending upon the position of the first rod, the second may be free to move, or unable to move. This simple blocking interaction serves as the basis for logical operations. Figure 10.1 shows a nanomechanical implementation of a Boolean NAND “interlock” gate, using clock-driven input and output logic rods 1 nm wide which interact via knobs that prevent or enable motion, all encased in a housing, allowing ~16 nm$^2$/interlock. (Any logic function, no matter how complicated, can be built from NAND or NOR gates alone.\(^ {1736}\) Figure 10.2 is an exploded diagram of the moving parts of a thermodynamically efficient class of register capable of mechanical data storage, using rods ~1 nm in width with 0.1-nanosec switching speeds, allowing ~40 nm$^3$/register. The activity sequence depicts a simplified version of Drexler’s register, omitting the reading mechanism.\(^ {1743}\) Initially the register shows a 0 (black ball position in A) or a 1 (B). Then the barrier is lowered and the ball wanders freely (C); entropy increases. The register is then reset to 0 by spring-rod compression, converting -KT ln(2) joules of work into heat (D). To write a 0, the barrier is raised (E). To write
a 1, the input rod at right is first extended and then the barrier is raised (F); the input rod does work compressing the ball into the spring, but this energy can be retrieved when the spring rod is retracted. Finally, the spring rod is retracted, returning the device to state (A) or (B).

Figure 10.3 illustrates a programmable logic array (PLA) using knobbed rods (omitting several drive and spring systems for clarity) to implement nanomechanical logic, which, in combination with rod-based registers, can be used to build much of the control circuitry for a CPU in a ≥1 GHz clocked computer system. For details of operation, see Drexler.10 While a PLA system requires three successive rod displacement cycles to compute a set of Boolean functions, the functions AND, OR, and NOT can also be computed in a single displacement cycle by employing a linkage in which any input rod can displace the output rod without displacing any other input rod; Figure 10.4 illustrates this alternative approach for an asynchronous-input OR gate.

Drexler’s benchmark mechanical nanocomputer design has 10^6 interlock gates, 10^5 logic rods, 10^4 registers, an energy-buffering flywheel and other components with total cubic volume (~400 nm)^3, mass ~10^-16 kg, and total power ~60 nW, giving a power density of ~10^12 watts/m^3.10 Power dissipation per logic operation is ~0.013 zJ per gate per cycle, giving (including register dissipation) ~2 x 10^8 operations/sec-m^3.10 Processing speed is ~10^9 operations/sec (~1 gigaflop) or ~10^18 operations/sec-m^3; assuming one bit per register, processing speed is ~10^13 bits/sec or ~10^30 bits/sec-m^3. In 1998, the typical desktop PC operated at ~10^8 operations/sec. Cooling is provided by a refrigerant fluid flowing through an integral fractal plumbing system at near the speed of sound.10

The most primitive 4-bit Intel 4004 microprocessor, introduced in 1971 for early-generation pocket calculators, had 2300 transistors. A comparably simple early mechanical nanocomputer with 2000 interlock gates, 100 registers, and a 1 KHz clock speed could have a volume as small as 36,000 nm^3, a (~33 nm)^3 cube, assuming 16 nm^3/gate and 40 nm^3/register, and could process ~10^5 bits/sec.

One major disadvantage of mechanical nanocomputers is that they will almost certainly be slower than nanoelectronic systems,
because a device that depends on moving heavy nuclei (10^{-27} kg) will necessarily be slower than a device that depends on moving less-massive electrons (10^{-30} kg). Mechanical signals travel near the speed of sound (10^4 m/sec in diamond), whereas electronic signals may travel near the speed of light (10^8 m/sec). Thus while nanomechanical computers may be limited to switching speeds of ~50 picoseconds, electronic devices may be ~10^{13}-10^{15} times faster.

On the other hand, mechanical computers are more EMP (electromagnetic pulse) resistant, conceptually easier to understand and to model, and mechanical designs scale readily from the macroscale to the nanoscale, unlike most nanoelectronic designs. In 1996, a ~50 nm reversible mechanical latch was fabricated and physically cycled using an AFM; the similarity to Drexler's mechanical OR gate (Figure 10.4) was explicitly noted in the paper.1737

What about mechanical data storage? In the MEMS world, Halc1742 fabricated a nonvolatile microelectromechanical memory cell consisting of a longitudinally-stressed surface-micromachined ~10 micron bridge that mechanically buckles into one of two states, making a bistable storage device. Moving to the nanoscale, data could be mechanically stored using compact three-dimensional arrays of diamondoid register rods (~10^7 bits/micron^3, ~10^{10} bits/sec access speed; Sections 7.2.1.1 and 7.2.6, and Drexler16). Another theoretical mechanical storage medium is spooled hydrofluorocarbon memory tape (~10^{10} bits/micron^3, ~10^8 bits/sec access speed; Sections 7.2.1.1 and 7.2.6 and Drexler16), or polymer or diamond surface-bound patterns of H and F atoms that could be read using an optimized (CH_{3})_{3}PO scanning probe tip with a maximum raw error rate of 6 x 10^{-8} per read.1200 (Linear DNA molecules achieve ~10^9 bits/micron^3 data storage, by comparison.) Direct AFM reading and recording of 10-nm pits on a polycarbonate surface (~10^6 bits/micron^2, up to ~10^7 bits/sec access speed)1739 and other means of directly reading and writing data at the atomic level1749,2182 have been studied, and the fullerene abacus1735 described earlier is one of the first experimental efforts to mechanically store numerical information using individual molecules at room temperature.

It has been speculated that individual atoms or atomic vacancies could serve as information units.278,1200,1739,1745,2711 For example, Y. Mo1741 used an STM to store information in the reversible rotational states of individual antimony dimers deposited on a silicon substrate. In theory, a block of diamond with data encoded as single-atom lattice vacancies could store up to ~176 bits/nm^3, or nearly ~2 x 10^{11} bits/micron^3, although cryogenic temperatures might be required to reduce diffusion effects and structural lability to acceptable levels consistent with long-term information storage. Phonon probes or other nondestructive means of lattice interrogation could be employed for readout.

10.2.2 Nanoelectronic Computers

In 1998, the metal-oxide-semiconductor field-effect transistor (MOSFET) was still the most common type of transistor in general use. These solid-state devices are characterized by an electrical source, an electrical drain, and a gate that controls the flow. Since its inception, the FET has scaled well to lower sizes. Individual working transistors with 40-nm gate lengths have been demonstrated in silicon,1751 and 25-nm gate-length transistors have been fabricated in gallium arsenide.1752 However, as feature sizes shrink below ~100 nm, a number of obstacles to stable operation begin to appear and the cost-effective downsizing of dense circuitry may not persist.1753-1757

A number of solid-state replacements for the bulk-effect semiconductor transistor have been suggested that could overcome these obstacles by taking advantage of nanometer-scale quantum mechanical effects.141 Such replacements include quantum dots (QDs),1784-1787 resonant tunneling devices (RTDs),1787-1791,1747,1790 electron turnstiles,709 and single-electron transistors (SETs).1791 In general, while hybrid RTD-FET circuits1791 and SETs1791,1814,1815 have been successfully switched at room temperature, many of these devices must be operated at cryogenic temperatures1848 and other practical challenges remain.1747,1784,1789-1791,1794 such top-down approaches to nanoelectronics are not considered further here.

In contrast, molecular nanoelectronics uses primarily covalently bonded molecular structures, electrically isolated from a bulk substrate, producing wires, switches and devices composed of individual molecules and nanometer-scale supramolecular structures.1747 While it is relatively difficult and expensive to sculpt trillions of identical nanoscale structures in bulk materials (e.g., using 2-nm resolution electron-beam lithographic techniques1819), individual molecules are natural nanometer-scale structures that can be manufactured to be exactly the same in near-mole quantities.1806 The great versatility of organic chemistry offers more options for designing and fabricating nanoelectronic devices than are available in silicon.1755-1778 Investigators are designing, modeling, fabricating, and testing individual molecules1760-1769,1871 and nanometer-scale supramolecular structures1811,1817 that act as electrical switches and even exhibit some of the same properties as small solid-state transistors.1768 For example, three- and four- terminal devices with ~0.3-nm feature sizes have been examined computationally.1879

Molecular electronics is a rapidly growing area in the literature.1760-1776,1832,1924 J.M. Tour expects experimental demonstration of molecular-sized transistor devices by 2000-2001, and commercial high-level computers using molecular electronics by 2008-2013.1517 A current list of major research groups is maintained by the International Society for Molecular Electronics and Biocomputing. Some of the following discussion is appreciatively drawn from an extensive 1997 review by MITRE Corporation.1747

10.2.2.1 Molecular Wires

Before one can seriously discuss molecular electronic devices embedded in single molecules, it must be asked whether a small single molecule can be made to conduct electrical current at all. A
series of difficult and sensitive experiments\textsuperscript{1765,1767,1800} and theoretical investigations\textsuperscript{1758-1761} during the 1990s answered this question affirmatively. Single molecule conduction was demonstrated by Tour\textsuperscript{1766,1767} using a conducting molecule characterized by repeating structures—in this case, a series of benzene-like rings connected by acetylene linkages—each part of which is linked to the next by bonds including many π-electrons above and below the plane of the structure. These electrons in the π orbitals\textsuperscript{1812} conjugate with each other, or interact, to form a single large orbital throughout the length of the wire to permit mobile electrons to flow.\textsuperscript{1778,1831} Thiol (-SH) functional groups at either end adsorb well to gold surfaces and act as “alligator clips” for attaching molecular electronic units to metal substrates.\textsuperscript{1819,1764,1867-1870} Tour’s polyphenylene wires can carry ~nanoampere currents. Such molecular wires also have the desirable property that they can be made quite long, if necessary, because they can be lengthened systematically using chemosynthetic methods.\textsuperscript{1766,1820} Another experiment using a single benzene ring attached between two gold electrodes through thiol groups across a 0.846-nm gap found a single-ring threshold resistance of \(-20\) megohm and a capacitance on the order of \(-0.1\ e^2/kT\) \(-10^{-19}\) farad.\textsuperscript{1811,1812,1837}

Fullerene carbon nanotubes (Section 2.3.2) or “buckytubes”\textsuperscript{1821} are hollow cylindrical tubes, essentially rolled-up sheets of graphite, measuring \(\geq 1\) nm in diameter that (in many chiral and compositional forms; Section 10.2.2.4) can conduct electricity. For instance, both individual 10-20 nm diameter carbon nanotubes\textsuperscript{1822,1844} and atomic wires pulled from an unraveling carbon nanotube\textsuperscript{643} can conduct currents of \(-1\)-to-10 microamps (vs. nanoamps via tunneling). Maximum nanotube current density is \(-10^{10}-10^{13}\) amperes/m\(^2\) over a 0.2-6.0 volt range,\textsuperscript{1844,1857} superior to the \(-10^9\) amperes/m\(^2\) typically achieved in superconductors. Nanotube resistivity measures \(-8\)-to-20 ohm-m for straight tubes, \(-38\)-to-49 ohm-m for slightly curved (5-30°) tubes, and \(>100\) ohm-m for highly curved (65-80°) tubes.\textsuperscript{1844} Interestingly, electronic transport in carbon nanotubes is ballistic (e.g., like electron waveguides, permitting only a few propagating modes\textsuperscript{1306}), with all waste heat being dissipated in the leads to the nanotube element and none in the nanotube itself.\textsuperscript{1857} Buckytubes are stiff enough to serve as supports for molecular circuit elements, and if filled with conducting metal can create one of the structurally strongest nanowires that is chemically possible. (See also Section 10.2.2.4.)

Other 0.6-3 nm molecular wires,\textsuperscript{1767,1801,1802,1864} 3-20 nm semiconductor wires,\textsuperscript{1843} and nanoscale metallic wires\textsuperscript{1740,1862,1863} have been studied as well.

### 10.2.2.2 Electromechanical Molecular Switching Devices

In 1998, the best-studied category of molecular electronics is probably the electromechanical molecular switch. This class of switch employs electrically or mechanically applied forces to turn a current on and off, either by changing the conformation of the molecule\textsuperscript{1766} or by moving a switching molecule or group of atoms in the manner of a gate.\textsuperscript{1805,1806} These switches are promising because they could be laid down in a dense network on a solid substrate, or in some cases in three-dimensional arrays; unfortunately, no integrated CPU-scale designs have yet been attempted, although plans for a half-adder\textsuperscript{1760} and a molecular shift register\textsuperscript{1750} have been published. Some examples:

**A. Electromechanical Amplifier** — Joachim and Gimzewski\textsuperscript{1768} measured conductance through a single buckyball held between an STM tip and a conducting substrate. By pressing down harder on the STM tip, the buckyball deforms, tuning the conductance off-resonance and reducing current \(-50\%\). Although the STM tip could be replaced with a small in-situ piezolectric gate, the ideal actuator would be a molecular-scale electromechanical actuator akin to a logic rod; the major fundamental limit to the speed of this switch is the natural mechanical vibrational resonance frequencies of a buckyball, \(-10^{13}\) Hz.

**B. Linear Atomic Relay** — Researchers at Hitachi Corporation\textsuperscript{1823} simulated a two-state electronic switch wherein a mobile atom that is not firmly attached to a substrate moves back and forth between two terminals (Figure 10.5). If the switching atom is in place, the device conducts electricity; if the switching atom is displaced from the two wires, the resulting gap greatly reduces the current that can flow through the atom wire. A small negative charge placed on the third atom wire “gate” near the switching atom moves the switching atom out of its place in the wire; the switching atom is pulled back into place by a second “reset” gate after each use of the switch. Actual experiments approximating this design created a bistable atom switch using a xenon switching atom moving back and forth between an STM tip and a substrate,\textsuperscript{1807,1808} demonstrating that the movement of a single atom can be the basis of a nanometer-scale switch. In a more mature device, relays could be \(-10\ nanometers\)^2 in size with a speed limited only by the intrinsic vibrational frequency of atoms, typically \(-100\) THz, but it seems likely that atom relays could only operate at very low temperatures because not much energy would be required to evaporate a switching atom off the substrate and out of the plane of the atom wires, destroying the switch.\textsuperscript{1747}

**C. Rotational Molecular Relay** — A more reliable two-state device based on atom movement uses the rotation of a molecular group to affect an electric current. The switching atom is part of a rotating group, or “rotamer,”\textsuperscript{1824} which itself is part of a larger molecule, perhaps affixed to the same surface as the atom wires. Figure 10.6 is a conceptual diagram illustrating this approach,\textsuperscript{1747} wherein the electric field of a nearby gate forces the switching atom to rotate in or out of the atom wire. When the switching atom is “in,” wire conductance is high and the switch is “on”; when the switching atom is “out,” a second group takes its place, hindering current flow and turning the switch “off.” A third large group may provide resistance to thermal free rotation. Alternatively, hydrogen bonding might provide enough rotational resistance to stop the rotamer in the conducting position, but not so much that reversing the gate voltage would be insufficient to turn the rotamer. Controlled rotational state switching of a Pt-adsorbed oxygen molecule by STM using 0.15-volt 20-microsec pulses was demonstrated in 1998.\textsuperscript{1874}

While a rotating switch with a methyl-like rotamer group has three distinct switch positions, a hinging switch that rotates back
and forth between two distinct states might be preferable. Cyclohexane, a simple example of this kind of molecule, can bend into two different forms, commonly known as the “boat” and “chair” conformations.\textsuperscript{1809,1810} As shown schematically in Figure 10.7, a voltage on a nearby gate might force the cyclohexane switch into one of its two conformations, affecting the conductivity of a nearby atom wire. The cyclohexane-type molecule attaches to a molecular framework while the remaining ring carbons are replaced by groups tailored to use steric repulsions or van der Waals attractions to reduce undesired switching caused by thermal energy. Switch speed is limited by molecular rotational and torsional frequencies, typically \(\sim 1-1000\) GHz, which is slower than the atom relay but more reliable. Also in contrast to the atom relay, molecular relays could be packed in three dimensions, possibly up to \(-1\) \(\text{nm}^{-3}\) densities.

D. Molecular Shuttle Switch — A rotaxane “shuttle switch” has been synthesized by F. Stoddart’s group\textsuperscript{1805} that consists of one ring-shaped molecule that encircles and slides along a shaft-like chain molecule (Figure 10.8). Two large terminal groups at either end of the shaft prevent the shuttle ring from slipping off the shaft. The shaft contains two other functional groups, a biphenol group and a benzidine group, that serve as natural stations between which the shuttle moves. The shuttle molecule contains four positively charged functional groups that cause it to be attracted to shaft sites with extra negative charge, so the shuttle spends 84% of its time at the benzidine station, which is a better electron donor than the biphenol station. Removing an electron from the benzidine station forces the shuttle to switch to the other station; switching can be controlled chemically or electrochemically.\textsuperscript{1805} A charged gate could be added at one or both ends to force the shuttle to move between stations, with switch state probed by arranging for the ring to complete an electric circuit in one of its two positions.\textsuperscript{1747} Switching speed is slow, limited by the speed of electron transfer to and from the benzidine station and the sluggish motion of the ring which is \(-10^6\) times heavier than an electron, but three-dimensional packing densities of \(-0.01\) \(\text{nm}^{-3}\) appear plausible. Other rotaxane shuttle systems have been studied.\textsuperscript{1845,1846,2487,2522,2529,2530} In 1999, the first chemically-assembled rotaxane-based logic gate was demonstrated,\textsuperscript{3541} although the gates could be opened only irreversibly.

10.2.2.3 Field-Controlled Molecular Switching Devices

A second category of molecular electronics device that appears promising for the development of molecular nanoelectronic digital computers is the electric-field controlled molecular switching devices, including molecular quantum-effect devices. These are most closely descended from the solid-state microelectronics and nanoelectronic devices described earlier, and promise to be the fastest and most densely integrated among all of the current alternatives. A group at Purdue University has used self-assembly to fabricate and demonstrate functioning arrays of molecular electronic quantum confinement structures connected by molecular wires.\textsuperscripts{1811,1812} As another example, J.M. Tour proposes embedding a quantum well in a molecular wire by inserting pairs of barrier groups that break the sequence of conjugated \(\pi\)-orbitals (Figure 10.9), forming a two-terminal molecular RTD;\textsuperscript{1747} other structures for three-terminal molecules have been suggested.\textsuperscript{1825} In February 1999, Tour reported the demonstration of RTD effects in the manner proposed above; at press time, the paper was still in review (J.M. Tour, personal communication, 1999.)\textsuperscript{1747} Nondissipative (adiabatic) Thouless electron pumping has also been described at 0.33 K.\textsuperscript{3189}

Merkle and Drexler\textsuperscript{1097} consider a hypothetical “helical logic” device in which a single electron can switch another single electron, \(\text{fig. 10.9}\).
with 1 or 0 represented as the presence or absence of individual electrons. The electrons are constrained to move along helical paths, driven by a rotating immersive electric field (normal to the helical axis) that confines each charge carrier to a fraction of a turn of a single helical loop, moving it like water in an Archimedean screw. A logic operation involves two helices, one of which splits into two “descendant” helices. At the point of divergence, differences in the electrostatic potential resulting from the presence or absence of a carrier in the adjacent helix control the direction taken by a carrier in the splitting helix; the sequence is reversible, allowing two initially distinct helical paths to merge into a single outgoing helical path without forcing a dissipative transition. Energy dissipation at ~10 GHz is $-10^6$ zJ/gate-cycle at an operating temperature of 1 K.

### 10.2.2.4 Other Molecular Electronic Devices

Other classes of molecular electronic devices have been studied, including picosecond photoactive or photochromic molecular switching devices\(^1\) using light, and electrochemical molecular devices\(^1\) using electrochemical reactions, to change the shape, orientation, or electron configuration of a molecule in order to switch a current. However, photoactive devices in a dense network would require an NSOM-like emitter (Section 4.8.4) to switch individually, since optical photon pathways are not easily confined at length scales very much below optical wavelengths of ~500-1000 nm; electrochemical molecular devices might require immersion in solvent to operate.\(^1\)

Fullerenes exhibit a wide range of quantized conductivities that may be useful in nanoelectronic computers.\(^2\) For example, chemisorption alters properties: 3 alkali dopant atoms per C\(_{60}\) buckyball gives high conductivity, while none or 6 dopant atoms gives an insulator.\(^3\) Buckytube conductivity also varies with mechanical stress and torsion,\(^4\) tube diameter, and other geometrical parameters.\(^5\) Depending upon the direction of the nanotube chiral vector, carbon nanotubes may be either metallic or semiconducting.\(^6\) Suitable, deep flexed, stressed, or chiralized fullerenes could be nested to make insulated wires\(^7\) or joined to make nanoelectronic computer components\(^8\), including resistors with quantized “staircase” resistance of nh/2e, where n = 1, 2, ... (e is electronic charge, h is Planck’s constant).\(^9\) Two-terminal devices such as diodes,\(^10\) three- or four-terminal devices such as heterojunctions,\(^11\) TUBEFET,\(^12\) P-channel or thin-film transistors, and gates that introduce electric fields,\(^13\) A single-molecule transistor using semiconducting buckytubes has been demonstrated experimentally at room temperature.\(^14\)

Carbon nanotubes have been cut to length, size-sorted by field flow fractionation, their open ends derivatized with thiol groups and then tethered to 10-nm gold particle “junction boxes” which AFM images show can connect together at least two separate tubes.\(^15\) If the tethering process can be spatially controlled, the current associated with the charge carrier is transferred without the need for direct electrical contact.\(^16\) Fullerenes exhibit a wide range of quantized conductivities that may be useful in nanoelectronic computers.\(^17\)

### 10.2.2.5 Molecular Electrostatic Field Computers

In 1998, bulk semiconductor technology typically required ~16,000 electrons to transfer one bit of information.\(^18\) SET and related techniques described above promise to reduce this to ~1 electron per bit. But a “traditional” molecular electronics nanocomputer with a mean gate density of ~1 nm\(^3\) operating at ~10 GHz and using single electrons to transport, indicate, fetch, or represent a binary digit would move ~1 coul/sec of electrons (~1 amp at ~1 volt, or ~1 watt) through a 1 micron\(^3\) CPU, producing an untenable ~10\(^{18}\) watt/m\(^2\) power density compared to ~10\(^{12}\) watts/m\(^2\) for the nanomechanical CPU described in Section 10.2.1.

In order to reclaim this factor of 10, Tour and Seminario\(^19\) conclude that a molecular CPU must operate by electrostatic interactions rather than by electron currents, using intermolecular “talk” to efficiently transfer a bit using only ~10\(^{-6}\) electrons.\(^20\) This charge-density approach is already well-known in biology, where enzyme-ligand recognition is mediated by hydrogen bonding or van der Waals interactions, working through the electric fields of the molecules even though no formal charge transfer occurs. Similarly, in an electrostatic nanocomputer the input signal is initiated by bringing up a charge to one end of the molecular device, thus reshaping the local charge density. That molecular device, in turn, perturbs the electrostatic field of an adjacent molecular device, rearranging its electron density in ~10\(^{-15}\) sec.\(^21\) and so the electrostatic field perturbation is propagated without charge transfer: “These movements of electrons (wave-particle entities) in the molecular orbitals (or circuits) are performed without any dissipation of energy (stationary states).”\(^22\) Perturbation pulses may have an energy magnitude similar to those of van der Waals interactions (Section 3.5.1). The hypothetical electrostatic molecular logic device shown in Figure 10.10 serves as an OR gate if the output is active (as a potential) or as a NOR gate if the output is passive (as an impedance) if positive logic is used, or an AND or NAND gate if negative logic is used.\(^23\)

Gates at input/output interfaces must still be able to receive or to drive signals from standard electronic circuits, but the vast majority of molecular gates would lie within the molecular CPU and could process information by controllably affecting the electrostatic potential of neighboring molecules. Such devices would likely shift the hardware/software equilibrium towards hardware, with massively wired-logic supplanting programmed-logic computing.
and with large molecular arrays allowing computation to take place inside the CPU with minimal need for main or auxiliary memories.1981

Tour expects the first laboratory demonstration of electrostatic field information transfer by 2001-2003.1517 J.C. Ellenbogen [personal communication, 1998] notes that a charge density nanocomputer achieves low dissipation by operating very close to equilibrium at all times, but notes that clock speeds of at least 1-10 KHz should still be obtainable with this approach; Tour [personal communication, 1999] claims that 1-10 THz switching is possible. Current-free polarization gates have also been proposed3216 and fabricated.3217

10.2.3 Biocomputers

Biocomputers use natural organic materials, thus offer the possibility of self-assembly or construction by natural or engineered biological organisms. At least three classes of biocomputers are distinguishable—biochemical, biomechanical, and bioelectronic.

10.2.3.1 Biochemical Computers

The Systron-Donner Analog Computer familiar to engineering students in the 1960s and 1970s used potentiometers and high-gain DC amplifiers, called operational amplifiers, to allow the creation of electrical circuits comprised of adders, multipliers, and integrators arranged in feedback loops capable of solving complex time-dependent differential equations modeling actual physical systems. In 1998, Motorola and others were producing programmable analog chips containing arrays of operational amplifiers and other analog components. Related control systems are commonplace in human physiology.71

Biochemical feedback loops exemplified by enzyme activity in the living cell exhibit similar analog computational functionality.2695,3509,3543,3544 To start with, the net flow of carbon through an enzyme-catalyzed reaction may be influenced by

1. the absolute quantity of enzyme present,
2. the pool of nonenzyme reactants and products available, and
3. the catalytic efficiency of the enzyme.996

All three influences may themselves be subject to recursive enzymatic control. For instance, the absolute quantity of enzyme present is determined by competing rates of synthesis and degradation; both of these processes are controlled by other enzymes. Product synthesis may be triggered by an inducer, or may be repressed by an end product. Local concentrations of coenzymes or metal ions can regulate catalytic efficiency. Inhibited activity of an enzyme in a biosynthetic pathway by an end product of that pathway, such as the much-studied inhibition of aspartate transcarbamoylase by cytidine triphosphate (CTP),3689-3691 is called feedback inhibition.3689-3693 Typically the feedback inhibitor, acting as a negative allosteric effector, binds to the sensitive enzyme at an allosteric site that is spatially distinct from the catalytic site. Further fine control is provided by additional multiple feedback loops, exhibiting such features as cumulative feedback inhibition (e.g., the inhibitory effect of two or more end products on a single regulatory enzyme is strictly additive),3694-3696 multivalent feedback inhibition (e.g., complete inhibition occurs only when two or more end products both present in excess)3697,3698 and cooperative feedback inhibition (e.g., single end product present in excess inhibits the enzyme, but inhibitory effect of two or more end products present in excess is “multiplicative”).3699,3700 Similarly, parallel chemical computers consisting of open, bistable coupled reaction systems3542 capable of pattern storage and pattern recognition have been proposed and simulated to model a neural network,1911 and experiments have been performed with a system of 16 coupled bistable chemical oscillators using the chlorite-iodide reaction.1912

Networks of interacting enzymatic reactivities can also be used to create biochemical digital computers.2851 An enzymatic Turing machine was first proposed by C.H. Bennett,296 and Liberman3471 described the living cell as an analog-digital molecular stochastic parallel computer. More recently, T. Knight of the MIT Artificial Intelligence Laboratory suggested that normal cellular activities can be co-opted to implement digital logic gates, a process he calls “cellular gate technology.”1962,3544 In this scheme, protein concentrations are used as binary signals. The signal is “1” if the marker protein is being synthesized, thus reaching some detectable threshold equilibrium value where the rates of synthesis and degradation are equal. The signal is “0” if the synthesis of the marker protein is being inhibited by some other protein, and thus the concentration is trending toward zero.

This principle may be used to make logical circuit elements. For example, an inverter system consists of two different proteins, one of which is inhibiting the synthesis of the other. A NOR gate requires a system in which either of two proteins can inhibit the synthesis of a third protein; multivalent feedback inhibition in which two proteins are both required to inhibit production of a third makes a NAND gate. This is sufficient to implement any Boolean function, given enough gates and signals. Two NORs crossed in a loop makes a bistable multivibrator, or flip-flop, a fundamental component in data registers, shift registers, rectangular pulse generators, or pulse delay elements. The logic circuitry (the arrangement of “gates” and
“wires”) comprises a set of designed DNA sequences that determine which proteins are being synthesized and which combinations of proteins can inhibit the synthesis of others.

In the most optimistic case, we assume operation near concentrations of $c_{prot} \sim 10^4 \text{g/m}^3$ (similar to most complement factors and some clotting factors in human blood; Appendix B), marker proteins and gating enzymes of mass $MW_{prot} \sim 10$ kilodaltons, and a minimum of $N_{prot} \sim 100$ protein molecules per gate, giving a typical logic element size of $V_{gate} = N_{prot} \cdot MW_{prot} / N_A \sim 0.02 \text{micron}^3$ where $N_A$ is Avogadro's number. Flip-flop registers may store $\sim 50 \text{bits/micron}^3$. In theory, the DNA specifications for this computer could be grafted onto the genome of an animal cell (~20 micron$^3$ = up to 1000 gate volumes) or a bacterial cell (~2 micron$^3$ = up to 100 gate volumes), taking care to avoid any unwanted interactions with natural biochemical pathways. Minimum diffusion-limited gating time across a biochemical gate of characteristic dimension $\Delta X \sim V_{gate}/\tau$ is given by Eqn. 9.80 as $\tau = \Delta X^2 / 2D = 0.3 \text{millisecond}$ assuming protein marker diffusivity $D \sim 10^{-10} \text{m}^2/\text{sec}$ (Table 3-3), consistent with -kHz maximum operating frequencies when DNA is localized near the gates. For nucleoplastic DNA with cytoplasmic computation, diffusion-limited maximum switching frequency falls to ~0.1-1 Hz. For comparison with natural systems, genes controlling the fibroblast cell proliferation program have been observed to switch states in ~900 sec,$^{2683}$ implying computation at ~0.001 Hz.

Other classes of biochemically-modulated biomolecular switches have been investigated. For example, Matthews$^{1929}$ incorporated an artificial molecular on-off switch into an enzyme by adding a pair of thiol groups into the active site of native lysozyme, replacing two amino acids on opposite side of the active site with cysteine, an amino acid with a thiol-containing side chain. Thiol may form covalent bridges (-SS-) or broken bridges (-SH HS-) under suitable chemical conditions. Cycling the solution composition causes the cysteines to form a bridge across the active site (inactivating the enzyme) or to break the bridge (activating the enzyme); the process is completely reversible. Another experiment$^{1930}$ involved genetic modification to the bacterial phage lambda, a virus with two unusual DNA motifs (Section 2.3.1), using stable-branched DNA enzymes to implement a DNA-based Turing machine, using only commercially available restriction enzymes and ligases for every operation with states represented as sequences of bases around a plasmid.$^{1886}$ and Warren Smith has another Turing machine design using the chemistry of guide RNAs in Trypanosome kinetoplasts.$^{701}$ Practical challenges include finding fast, efficient, and low-noise input and output techniques.$^{1902,2331}$

DNA computation is in theory energetically efficient, requiring only ~12 KT per ligation operation at room temperature, with massively parallel operations although each ligation reaction requires ~1 sec to complete.$^{1896}$ DNA data representation allows up to ~1 bit/nm$^3$ storage density.$^{1895}$ However, a practical mass storage device would probably require 10-100 nucleotide-long words to avoid ambiguities during the recall process,$^{1903}$ thus reducing maximum density to $10^{-10}$ bits/micron$^3$, though of course some of this reduction could be offset by reducing read rate requirements (e.g., by using multiple read heads with majority logic).

10.2.3.2 Biomechanical Computers

Components of mechanical computers could be constructed entirely of biologically-derived materials which in theory might be fabricated using synthesis pathways accessible to engineered microbiota. For example, Drexler's mechanical nanocomputer,$^{10}$ assumes diamondoid logic rods with a Young's modulus of $E = 5 \times 10^{11}$ N/m$^2$, but the Young's modulus for human bone hydroxyapatite and fluorapatite crystals is almost as high, with $E \approx 1-2 \times 10^{11}$ N/m$^2$ (Table 9.3). Ignoring the significant problems inherent in noneutactic rod logic systems, apatite-based mechanical computer structures might be laid down using engineered osteocytes, perhaps assisted by chemotactic or contact guidance techniques. Failure strength of apatites is ~100 times poorer than for diamond, so apatite logic rod force and acceleration should be reduced by ~100 and rod vibrational energy by ~10$^4$, increasing minimum switching time by a factor of ~10 (to ~1 nanosec) and decreasing maximum rod speed to ~1 m/sec.

N. Seeman has constructed molecular building blocks from unusual DNA motifs (Section 2.3.1), using stable-branched DNA molecules with the connectivity of a cube$^{1914}$ or a truncated octahedron.$^{1915}$ E. Winfree$^{1916}$ has proposed using arrays of DNA crossover molecules$^{1924}$ in DNA-based computing, requiring the ability to build periodic backbones with bases differing from one unit cell to another. In addition to branching topology, DNA also allows control of linking topology. DNA-based topological control has led to the construction of Borromean rings which could be used in DNA-based computing applications.$^{1916,1918}$ In Borromean rings, the linkage between any pair of rings disappears in the absence of the third. Rings can be designed with an arbitrary number of circles; the integrity of a link could represent the truth of each of a group of logical statements.$^{1916}$

DNA structural transitions could be used to drive nanomechanical devices via branch migration.$^{1916}$ Application of torque to a cruciform leads to the extrusion or intrusion of a cruciform.$^{1921}$ A synthetic branched junction with two opposite arms linked can relocate its branch point in response to positive ethidium-induced supercoiling, representing the first experimental research has shown that the approach can be applied to a much wider class of digital computations.$^{1908-1909}$ The well-known problem of combinatorial explosion (e.g., even a small protein of 300 amino acids has a sequence space of $10^{390}$) leading to huge-volume DNA libraries during a complex computation may be avoided using a recursive selection approach,$^{1899}$ crudely analogous to genetic algorithms. Some work has been done using plasmids and restriction endonucleases to implement a DNA-based Turing machine, using only commercially available restriction enzymes and ligases for every operation with states represented as sequences of bases around a plasmid.$^{1886}$ and Warren Smith has another Turing machine design using the chemistry of guide RNAs in Trypanosome kinetoplasts.$^{701}$ Practical challenges include finding fast, efficient, and low-noise input and output techniques.$^{1902,2331}$
step in developing DNA structural transitions that can achieve a nanomechanical result.\textsuperscript{1922} The possible use of the B-Z transition (e.g., from right-handed B-DNA to left-handed Z-DNA) in nanomechanical devices is being explored.\textsuperscript{2409}

In 1998, Winfree and colleagues in Seeman’s laboratory\textsuperscript{1970} described a simple, predictable, highly precise technique for arranging DNA molecules into two-dimensional crystals. At the same time, a paper by Heath and others\textsuperscript{1980} described a defect-tolerant computer architecture called the Teramac, a massively parallel experimental computer which apparently could operate normally with at least half of its most critical components failed. The authors claimed this architecture showed it would be “feasible to chemically synthesize individual [molecular] electronic components with less than a 100% yield, assemble them into systems with appreciable uncertainty in their connectivity, and still create a powerful and reliable data communications network.”

One can also imagine a network of interlocking enzymelike conformational switches, possibly embedded in a semirigid two- or three-dimensional scaffolding of DNA\textsuperscript{1904,1913-1916,1919,1920} or protein\textsuperscript{1905,1913} molecules. Enzymes may display multivalent feedback inhibition (making a NAND gate; Section 10.2.3.1) or may have multiple metal-ion binding sites influencing enzymatic action.\textsuperscript{1928} An enzyme requiring two conformational changes to inhibit its action could serve as a biochemical NAND gate if tethered to other similar enzymes in an appropriately structured protein array.\textsuperscript{1905} Single-device error rates will be high but might be reduced to acceptable levels by employing short multiply-redundant pathways for each digital computational operation, or a sufficiently parallel architecture.\textsuperscript{1980} Site-directed mutagenesis optimization of antibodies has shown that natural proteins can be made significantly stiffer and more stable by the mutation of just a few amino acids. Systems of reciprocating actomyosin molecular motors (e.g., as found in muscle fibers; Section 6.3.4.2) might serve as interdigitating components of a push-pull mechanical logic architecture resembling rod logic or a three-dimensional loom. In early 1999, Viola Vogel (University of Washington Center for Nanotechnology) and colleagues discovered, via computer simulation, a tension-activated fully-reversible bio-mechanical switch comprised of a single strand of fibronectin protein—like untying a shoelace, a slight mechanical tug unravels a folded segment, switching off the protein’s biochemical activity.\textsuperscript{3249}

Mechanical coupling of intracellular Ca\textsuperscript{2+} ion release channels allows coordinated voltage gating across the surface membrane of the sarcoplasmic reticulum inside muscle cells.\textsuperscript{1964} R. Bradbury [personal communication, 1999] suggests that site-directed mutagenesis could produce a family of proteins that could be activated by the release of various mono- (e.g., K\textsuperscript{+}), di- (e.g., Ca\textsuperscript{2+}), and tri-valent (e.g., Al\textsuperscript{3+}) ions. This would allow multivalued logic (e.g., K = 1, Ca = 2, Al = 4), or increased parallelism or bandwidth (e.g., Ca = ![K & K], Al = ![Ca & Ca], etc.); once a protein exists that can be activated by a single ion, then varying its architecture to allow substitution of an ion of different size or charge is a relatively straightforward engineering exercise. A biomechanical Turing machine design has also been reported by Shapiro.\textsuperscript{3247}

### 10.2.3.3 Organic and Bioelectronic Computers

Many organic materials may be useful in electronic computing. Electrical conduction in synthetic organic molecules is well known,\textsuperscript{1761,1906-1909,3125-3128} and already exploited in liquid-crystal displays of laptop computers.\textsuperscript{1925} Doped polycrylen is a conducting plastic with a carbon backbone; undoped, it is an insulator. There are many other known families of conducting polymers such as polyaniline, which includes nitrogen atoms in the backbone, and polythiophene, which includes sulfur.\textsuperscript{1933} Three-dimensional interconnect arrays of organic conducting polymers have been fabricated,\textsuperscript{1839} conductivity is higher along a chain than between chains, so sheets of oriented chains can produce relative conductivity anisotropies as high as 100-1000.\textsuperscript{1922} An all-plastic transistor has been built.\textsuperscript{1935} Organic transistors have been used as the active semiconducting element in thin-film transistors fabricated with organic-material thicknesses ranging from 5-150 nm,\textsuperscript{1842} and organic microscale transistors and other organic devices may exploit bulk-effect electron transport just like silicon-based semiconductor devices.\textsuperscript{1925} Conjugated polymer poly(1,6-heptadiester) was employed to make an optical correlator with a peak processing rate of 3 x 10\textsuperscript{16} operations/sec.\textsuperscript{1841} Spin-transition polymers could serve as thermal sensors or memory devices with the theoretical ability to store one bit in a ~4 nm cube (~2 x 10\textsuperscript{7} bits/micron\textsuperscript{3} storage density) with GHz addressing speeds at room temperature.\textsuperscript{1840} Other organic polymers possibly useful in logic circuits have been investigated.\textsuperscript{2899}

Natural biological materials may also be useful in electronic computing, particularly because of the ability of biological materials to self-assemble. Biomolecular electronics is a subfield of molecular electronics that investigates the use of native as well as modified biological molecules (chromophores, proteins, etc.) in place of the organic molecules synthesized in the laboratory.\textsuperscript{1777} In theory, a three-dimensional DNA or protein scaffolding could be self-assembled, with bacteriorhodopsin, ferritin and magnetite, or related bioelectromagnetic molecules inserted into precise locations within the structure to produce a crystal lattice bioelectronic or biooptoelectronic nanocomputer. DNA already has been used as an atomic scaffold upon which to build silver wires -100 nm in diameter\textsuperscript{1961} and has been decorated with fullerenes.\textsuperscript{3024} Self-assembly can generate membrane-based and biotubule-based devices, and site-directed mutagenesis is a valuable tool for high-resolution protein device engineering.\textsuperscript{1935} Nucleotide sequence-specific electron transfer between metallic electron donor and acceptor complexes covalently or noncovalently intercalated into strands of B-DNA has been demonstrated.\textsuperscript{1934}

Perhaps the best-known bioelectronic (and biochemical) computer is the neuron cell and its congeries—the ganglia, nerve trunks, and brains. In 1997, David Stenger (NRL) and James Hickman (SAIC) were attempting to culture and link together neuronal cells to build a bioelectronic cell-based sensor “computer” for performing complex pattern-recognition tasks.\textsuperscript{1966,1967} In early 1999, W.L. Ditto at Georgia Tech announced the creation of a biological computer comprised entirely of leech neurons, that was capable of performing simple sums.\textsuperscript{3245} We can certainly imagine large numbers of cultured neurons arranged in artificial three-dimensional spatial patterns to produce a synthetic bioneural computer, but such a computer would operate at only ~KHz frequencies and would be very energy inefficient since each neuron consumes ~10\textsuperscript{-8} KJ per discharge (Section 4.8.6.2)—thus offering few advantages.

#### 10.2.4 Ultimate Limits to Computation

A nanosystem designer quickly discovers that total nanorobot power demand (Section 6.5) is often dominated by computational heat dissipation, rather than by energy dissipation from chemical or mechanical processes. Thus important design objectives are to:

1. minimize onboard computation,
2. maximize algorithmic efficiency (e.g., super-Turing computers),
3. minimize power dissipation per computation operation, an
objective which may be met using reversible computing.

It is also interesting to consider the maximum possible information density in physical matter as a calibration for current and projected achievements.

10.2.4.1 Reversible Computers

Computers may be thought of as engines for transforming free energy into waste heat and mathematical work. Early pioneers of computing theory believed that each step in a computer's binary computation required a minimum energy expenditure of $-kT \ln(2) \approx 3 \text{ eV/bit}$ at $T = 310 \text{ K}$. In 1961, R. Landauer argued that it was the erasure of information, not computation per se, that generates waste heat. It is now known that computers can in principle do an arbitrarily large amount of reliable computation per $kT$ of energy dissipated. Following Landauer's insight, Fredkin and Toffoli suggested an idealized "ballistic computer" that could, in theory, compute at finite speed with zero energy dissipation and zero error. A more pragmatic family of models are the "Brownian computers" in which thermal noise pushes system elements in a random walk throughout the entire accessible portion of the computer's configuration space; in these models, energy dissipation trends to zero only in the limit of zero speed. Both ballistic and Brownian computers require that all computations are logically reversible, with no irreversible bit erasures and no machine state having more than one logical predecessor—that is, the output uniquely specifies the input.

One simple implementation of reversible computing is the retractile cascade. In a retractile cascade, all inputs and all intermediate states leading to a result are retained during the course of a computation. After the computation is complete, the final result is copied to an output register, requiring the irreversible erasure of only enough bits in the output register to hold a copy of the final result, which need cost no more than $kT \ln(2)$ per erased bit. The computation is then reversed, step by step, culminating with the original inputs: the slower the reversible steps are performed, the less energy they may dissipate (but the longer the computation takes). Schematics for a retractile AND gate, adder, shifter, and programmable logic array have been published, and in 1998 a reversible processor based on a modern RISC architecture was being designed. A conventional computer architecture, implemented without regard to reversibility, may perform 0.1-1 bit erasures/gate-cycle; by comparison, a retractile computer might average $<10^{-4}$ bit erasures/gate-cycle. During the reversible portion of the computation, Drexler's exemplar rod logic design employs a retractile cascade that reduces room-temperature energy dissipation from the "classical" minimum of $-0.7 \text{eV/kT}$ to $-0.003 \text{eV/kT}$ per gate-cycle. In the limit of slow motion, all identified energy dissipation mechanisms in combinational rod logic systems approach zero.

The Tour-Seminario electrostatic field switch might attain $10^{-5}$ kT per gate-cycle at room temperature; low-temperature helical logic (Section 10.2.2.3) could achieve $10^{-7}$ kT per gate-cycle. Feynman notes that the minimum free energy required for a reversible computation may be made independent of the complexity or number of steps in the calculation, and may be as small as $-kT$ per bit of the output answer.

J.S. Hall suggests two principal design rules for efficient nanocomputers:

1. erase as few bits as possible, and
2. eliminate entropy loss in operations that do not erase bits.

Many reviews of reversible computing have been published.

10.2.4.2 Quantum Computers

Digital computers manipulate discrete units of information, with data strings stored as binary digits, or bits. Any two-state physical system (e.g., high or low voltages in semiconductor circuits) can represent these bits. Similarly, a quantum computer would use quantum states of atomic or molecular systems to store data. For example, the states of a hydrogen atom could be assigned such that a wavefunction corresponding to a state of hydrogen would represent a bit of data. Shining laser light on the atoms would trigger processing the data by inducing transitions between electronic states. Quantum systems, however, exhibit a peculiar phenomenon known as superposition, in which several discrete states can be processed by a single physical system at once, crudely analogous to the musical effect of a single note being made up of many different harmonics. Superposition gives quantum systems the potential to be enormously more powerful than conventional computers. Since more than one state can be supported, quantum bits, or qubits, may store a mixture of many bits of data at the same time. A superposition of $N$ qubits can store $2^N$ binary digits. An operation on these mixtures is massively parallel, effectively performing many calculations simultaneously. For example, photons interacting with an atom in a superposition of states drive all the states in the superposition, producing another superposition corresponding to all solutions of the original states. The advantage of conventional computers is not merely linear, but exponential: A conventional processor operates sequentially on 64-bit numbers, but a quantum computer with 64 qubits would simultaneously operate on the full set of $2^{64}$ binary values. An N-qubit quantum computer could model an N-body system in real time; in contrast, the number of operations required by a conventional computer to perform such a simulation increases exponentially with the number of bodies.

Since the early 1980s, theoretical interest in quantum computing has developed rapidly. Many groups have begun designing quantum computational architectures and constructing physical quantum gates and devices. In 1995, one group at NIST made a controlled NOT gate of two qubits from a cooled beryllium ion in a radio-frequency trap, while a French group used single photons trapped in quantum cavities to control cesium atom transitions. In 1999, experiments began on an analog quantum computer using electrons floating on liquid helium. Others work with NMR techniques, storing data not in electron states but in nuclear spins, which are less susceptible to perturbations. Resonance effects between proton and electron spins in hydrogen atoms can make AND and NOT gates, hence NAND gates, from which all Boolean computers can be made. It has been suggested that by encoding the amplitudes of a couple of thousand electron states in a superposition, large amounts of data could be written onto a single atom (see Section 10.2.4.3). The discovery of certain error-correcting procedures, since verified experimentally, was another important breakthrough. Unlike conventional devices, it is not possible to simply check that qubits are in the right states, because the act of measuring destroys the coherence. But Shor proved that information can be reliably stored and read from a several qubit system even if one of the bits is corrupted, and it has been shown that error correction can be applied during the computation process itself. This opens the door to many-qubit systems; by 1998, two groups had already performed 4-bit sums, and one group had demonstrated a 2-qubit device. Quantum game theory has also been investigated.
10.2.4.3 Bekenstein-Bounded Computation

A fundamental upper limit on the number of possible quantum states in a bounded region (e.g., the maximum number of bits that can be coded in a bounded region)—is given by the Bekenstein Bound\(^{1990}\) as:

\[
I_{\text{Bek}} = \frac{2 \pi E R}{hc \ln(2)} \quad \text{(bits)} \tag{Eqn. 10.10}
\]

within a spherical region of radius \(R\) containing energy \(E = mc^2\) where \(m\) is the enclosed mass, \(c = 3 \times 10^8 \text{ m/sec} \) (speed of light), and \(h = 2 \pi \) where \(h = 6.63 \times 10^{-34} \text{ joule-sec} \) (Planck's constant). Maximum processing speed \(I\) is then bounded by the minimum time for a state transition, which cannot be less than the time required for light to cross the region of radius \(R\), or:

\[
I_{\text{Bek}} = \frac{\pi E}{h \ln(2)} \quad \text{(bits/sec)} \tag{Eqn. 10.11}
\]

Thus in theory, a single carbon atom of mass \(m = 2 \times 10^{-26} \text{ kg}\) and radius \(R \sim 0.15 \text{ nm}\) could be impressed with up to \(I_{\text{Bek}} \sim 10^8 \text{ bits/sec}\) and could process information up to \(I_{\text{Bek}} \sim 10^9 \text{ bits/second}\) in an optimally-designed quantum computer. For a 1 micron\(^3\) computer of mass \(\sim 10^{-15} \text{ kg}\), \(I_{\text{Bek}} \sim 10^{22} \text{ bits maximum storage capacity and}\) \(I_{\text{Bek}} \sim 10^{37} \text{ bits/sec maximum processing capacity}\). These values are not the least upper limits; Schiffer and Bekenstein\(^{1991}\) estimate that the Bekenstein Bound probably overestimates \(I\) and \(h\) by at least a factor of 100, but Likharev\(^{1992}\) and Margolus and Levitin\(^{2319}\) have derived similar limits. In 1998, proposed reversible quantum computer systems would perform significantly below the Bekenstein Bound.\(^{1993,1994}\)

10.2.5 Dedicated Computational Organs

Nanostructured artificial organs specializing in control, information processing, or data storage could be permanently embedded within the human body. Such organs could provide a user-accessible computational facility, or help coordinate and control (or even subsume) the activities of in vivo medical nanorobots or other dedicated nano-organs, including energy (Section 6.4.4), communications (Section 7.3.4), navigation (Section 8.3.6), or chronometer (Section 10.1.4) systems. Computational organs could also serve as embedded data repositories (library nodules), including personal medical and experiential records, customized calendars and personal data banks, or as communications encryption and security interfaces.

In 1998, microchips glued to honeybees\(^{3727}\) or implanted under the skin of livestock and household pets contained owner/address information that could be retrieved by passing a reading wand over the animal (e.g., Pet Trac, manufactured by A.V.I.D. Identification Systems; Norco, California). Read-only microchips with identification numbers for individual cattle were encased in acid-resistant porcelain and inserted into an animal's stomach ten days after birth. These chips could be read by a handheld computer or a stationary reader; Shearwell Data, which sold the chips for $6 each, planned to adapt them to communicate with global positioning satellites so that farmers could remotely track herd movements.\(^{3037}\) Peter Cochrane\(^{2956}\) notes that a colleague (Kevin Warwick) has had a computer chip embedded under his skin,\(^{3324}\) bearing such personal information as medical records and bank account and passport numbers. There were also unconfirmed rumors of a subdermally-implanted bio-powered (4 millamp) 16 mm\(^2\) transmitter chip called Sky-Eye, sold by Gen-Etics for $7,500 as kidnap deterrence for “the rich and famous,” that supposedly emitted a personal tracking signal detectable by satellite.\(^{3762}\)

10.3 Pressure Storage and Ballasting

Storage of gases at high pressure is useful for nanorobots that must perform physiological gas transport functions\(^{1400}\) (Chapters 22 and 26) or emergency reactive functions (Chapter 24), for devices which employ gases in onboard energy systems (Chapter 6), or for devices which require gaseous chemicals for onboard materials processing (Chapter 19). Pressurized gases are useful in buoyancy maintenance to facilitate device ejection from the human body (Section 10.3.6), and possibly in the deactivation of toxic biochemicals or microbial pathogens (Table 10.3). Keeping a eutactic environment\(^{10}\) inside nanodevices requires the ability to draw and maintain a vacuum (Section 10.3.5). This ability also may assist in metamorphic bumper (Section 5.4) and buoyancy control.

10.3.1 Fluid Storage Tank Scaling

The ability of a vessel to store fluids at high pressure is determined by the rupture strength of the valve and pipe systems and by the tensile strength of the tank walls. A symmetric structural shell transmitting only normal stresses in orthogonal directions may employ either a stressed skin or a ribbed design. The stressed skin is more efficient because the same material carries stress in both directions and there is integral resistance to secondary torsional and bending loads.\(^{2023}\) For a stressed skin spherical pressure vessel of radius \(R\) with wall material of density \(\rho_{\text{wall}}\) and working stress \(\sigma_w\) containing fluid at pressure differential \(p_{\text{fluid}}\), the required wall thickness \(t_{\text{wall}}\) is given by:\(^{2023}\)

\[
t_{\text{wall}} = \frac{P_{\text{fluid}} R}{2 (\sigma_w - g \rho_{\text{wall}} R)} \tag{Eqn. 10.12}
\]

where \(g = 9.81 \text{ m/sec}^2\) (acceleration of gravity). For all but the most enormous macroscale tanks, the second term in the denominator is negligible and so the maximum pressure differential that can be restrained by a spherical microscale tank without bursting is:

\[
P_{\text{max sph}} = \frac{2 t_{\text{wall}} \sigma_w}{R} \tag{Eqn. 10.13}
\]

Thus a tank of radius \(R = 1 \text{ micron}\) made of diamondoid walls which are \(t_{\text{wall}} = 5 \text{ nm}\) (-30 carbon atoms) thick with \(\sigma_w = 10^{10} \text{ N/m}^2\) (-0.2 times the failure strength of diamond; Table 9.3) can store fluids up to a maximum pressure differential of \(p_{\text{max sph}} = 1000 \text{ atm}\). Empty tank mass is \(M_{\text{tank}} = 4 \pi R^2 t_{\text{wall}} 3^{67} \text{ nm or it may}\) burst.

\[
P_{\text{max cyl}} = \frac{t_{\text{wall}} \sigma_w}{R} \tag{Eqn. 10.14}
\]

Similarly, for a cylinder or pipe of radius \(R\):

\[
P_{\text{max cyl}} = \frac{t_{\text{wall}} \sigma_w}{R} \tag{Eqn. 10.14}
\]

Note that a hemispherical-capped cylindrical tank, pressurized until it explodes, will usually split its sides before blowing off its ends because \(p_{\text{max sph}} = 2 p_{\text{max cyl}}\). Wall stress for a cylindrical tube with Young's modulus \(E\), inside radius \(R\) and wall thickness \(t_{\text{wall}}\) upon which is imposed a pressure differential of \(P_{\text{cyl}}\) is \(\sim P_{\text{cyl}} R / t_{\text{wall}}^{362}\) giving a
wall strain of $S - p_{cy} R / t_{wall} E$; hence a single-walled carbon nanotube with $R = 20$ nm, $t_{wall} = 0.2$ nm, and $E = 10^{11}$ N/m$^2$ pressurized to $p_{cy} = 1000$ atm stretches by $S \sim 1\%$. (For comparison, aluminum soda pop cans are pressurized to ~2 atm differential, beer cans ~1 atm.)\textsuperscript{3703}

For a toroid tank of small meridional radius $r$ and large circumferential (hoop) direction $R$, the burst pressure in the meridional ($p_{\text{max merid}}$) and hoop ($p_{\text{max hoop}}$) directions are:\textsuperscript{2023}

$$p_{\text{max merid}} = \frac{t_{wall} \sigma_{w}}{r} \quad \text{[Eqn. 10.15]}$$

$$p_{\text{max hoop}} = \frac{2 t_{wall} \sigma_{w}}{r} \quad \text{[Eqn. 10.16]}$$

Hence as pressure is raised, the toroidal tank fails first in the meridional direction, much like the capped cylinder.

Spherical pressure vessels are the most efficient. The maximum mass of gas that can be contained within a spherical tank at ideal gas pressures (Section 10.3.2) is $M_{\text{gas}} = (4 \pi N_A / 3) R_{\text{gas}} T_{\text{gas}} p_{\text{max}} R^3$, where $N_A = 6.023 \times 10^{23}$ molecules/mole (Avogadro’s number), $M_{\text{gas}}$ is the mass per gas molecule, $R_{\text{gas}} = 8.31$ J/mole-K (universal gas constant), and $T_{\text{gas}}$ is gas temperature. Thus we observe that the ratio of gas mass to structural mass ($M_{\text{gas}}/M_{\text{tank}}$) at a given pressure is constant and independent of the size of the tank. This ratio reaches a maximum value at maximum pressurization; the maximum value is independent of both tank size and the maximum pressure selected (within the ideal gas range). In particular, $M_{\text{gas}}/M_{\text{tank}} \sim 2 \sigma_{w} N_A R_{\text{gas}} / 3 R_{\text{wall}} T_{\text{gas}}$ $p_{\text{wall}} - 21$ for nitrogen gas molecules ($m_{\text{gas}} = 4.65 \times 10^{-26}$ kg/molecule) stored at temperature $T_{\text{gas}} = 310$ K. That is, the tank can store up to ~21 times its weight of gas.

Other important microtank design constraints include diffusion leakage and vessel flammability (Section 10.3.4), vapor pressure of gas storage vessels maintained at various pressures at $T_{\text{gas}} = 310$ K. That note that there are significant (near-linear) gains in gas molecule packing density up to 1000 atm, minor (non-linear) gains up to 10,000 atm, and no significant gain at >100,000 atm as density approaches a limiting value for the liquid or solid state. Table 10.2 also shows that pressures calculated using the ideal gas law vary from pressures calculated using the van der Waals equation by ~5% at 100 atm and ~50% at 1000 atm.\textsuperscript{1031}

One early experiment provided evidence of stably trapped (Section 10.3.4) room-temperature gases at ~1300 atm pressure inside carbon nanotubes.

\textbf{10.3.2 The Van der Waals Equation}

The compressibility of gases is most simply described by the well-known ideal gas law:

$$p_{\text{gas}} V_{\text{gas}} = n R_{\text{gas}} T_{\text{gas}} \quad \text{[Eqn. 10.17]}$$

where $p_{\text{gas}}$, $V_{\text{gas}}$ and $n$ are gas pressure, volume, and number of moles, respectively. However, all real gases deviate to some extent from the ideal gas law, even at STP. Most real gases obey the ideal gas law to within a few percent at low pressures—that is, at low pressures (e.g., $\leq 1$ atm) and at temperatures well above their condensation points. For real gases at high pressure, finite molecular volumes and intermolecular attractions cause significant deviation from the ideal gas law. In 1873, J.D. van der Waals deduced an empirical equation of state (subsequently derived from statistical mechanics using suitable approximations\textsuperscript{1031}) that reproduces the observed behavior of real gases with moderate accuracy:

$$p_{\text{gas}} + \frac{A n^2}{V_{\text{gas}}^2} (V_{\text{gas}} - B n) = n R_{\text{gas}} T_{\text{gas}} \quad \text{[Eqn. 10.18]}$$

where $A$ is a measure of intermolecular attraction and $B$ is a measure of finite molecular volumes. The van der Waals gas “constants,” given in Table 10.1 for various gases, are known to vary slightly with temperature. Nevertheless, while Eqn. 10.18 is only one of several expressions commonly employed to represent real gas behavior, it is the simplest to use and to interpret. The van der Waals equation remains approximately valid even at temperatures and molar volumes so low that the gas has become a liquid.\textsuperscript{1031} Note that the critical temperature $T_{\text{crit}}$, the highest temperature at which gas and liquid may exist as separate phases at any pressure, is approximated by $T_{\text{crit}} = 8 c_1 A / 27 B R_{\text{gas}}$, where $c_1 = 1.01 \times 10^5$ J/m$^3$-atm and $R_{\text{gas}} = 8.31$ J/mole-K; critical pressure $p_{\text{crit}} = A / 27 B^2$.\textsuperscript{390} For example, in the case of water, at temperatures and pressures above $T_{\text{crit}} = 647.3$ K and $p_{\text{crit}} = 218.3$ atm,\textsuperscript{763} the vapor and liquid phases become indistinguishable.

The boiling point of a pure liquid as a function of pressure may be approximated by:\textsuperscript{2036}

$$\ln \left( \frac{p_2}{p_1} \right) = \frac{\Delta H_{\text{vap}}}{R_{\text{gas}}} \left( \frac{1}{T_1} - \frac{1}{T_2} \right) \quad \text{[Eqn. 10.19]}$$

where $T_1$ and $T_2$ are the boiling points (in K) at pressures $p_1$ and $p_2$, respectively; $\Delta H_{\text{vap}}$ is the molar heat of vaporization for the liquid (Table 10.8), and $R_{\text{gas}} = 8.31$ J/mole-K. (This formula assumes that $\Delta H_{\text{vap}}$ is constant over the temperature range from $T_1$ to $T_2$.) Taking $T_1 = 373.16$ K, $\Delta H_{\text{vap}} = 40.690$ J/mole, and $p_1 = 1$ atm for water, then at $p_2 = 6.4$ atm the boiling point has risen to $T_2 = 435$ K. Boiling point also is altered by the presence of solute (Section 10.5.3).

Table 10.2 shows the molecular number density achieved inside gas storage vessels maintained at various pressures at $T_{\text{gas}} = 310$ K. That note that there are significant (near-linear) gains in gas molecule packing density up to 1000 atm, minor (non-linear) gains up to 10,000 atm, and no significant gain at >100,000 atm as density approaches a limiting value for the liquid or solid state. Table 10.2 also shows that pressures calculated using the ideal gas law vary from pressures calculated using the van der Waals equation by ~5% at 100 atm and ~50% at 1000 atm.\textsuperscript{763}

One early experiment provided evidence of stably trapped (Section 10.3.4) room-temperature gases at ~1300 atm pressure inside carbon nanotubes.

Fig. 10.11. Phase diagram for water and ice (stable phases only).\textsuperscript{354,2053,2054}
Single-walled carbon nanotubes have been experimentally charged with trapped hydrogen gas up to 5-10% by weight, and with argon gas at ~600 atm. It is only possible here to briefly mention the many pressure-mediated changes of physical properties that may be relevant in nanomedical systems. Perhaps the most familiar is phase change, as illustrated by Figure 10.11 in the case of water. Note that as pressure rises from 1 atm to 2054 atm, the temperature at which water remains liquid dips from 273.2 K to 251.2 K (the Ice I/III/water triple point, 2962), but then the liquidity threshold returns to higher temperatures as pressure rises further and a different ice phase forms; D2O shows similar behavior. At a constant 310 K (37˚C), molecular mass: 5.32 x 10^-26 kg 7.31 x 10^-26 kg 4.65 x 10^-26 kg 6.64 x 10^-27 kg ---

\* For comparison, pressures of ~1000 atm occur naturally in the deepest part of the ocean (the Marianas Trench), ~10,000 atm at the crust-mantle interface (the Mohorovic discontinuity), 3.64 million atm at the Earth’s core, and ~10^9 atm at the center of the Sun.

Table 10.2. Gas Molecules Packed into a Pressure Vessel: Number Density as a Function of Pressure at 310 K, using the Van der Waals Equation

<table>
<thead>
<tr>
<th>Applied Pressure</th>
<th>O₂ Molecules per m³</th>
<th>CO₂ Molecules per m³</th>
<th>N₂ Molecules per m³</th>
<th>He Atoms per m³</th>
<th>Ideal Gas Law, No. per m³</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Solid)</td>
<td>268 x 10²⁶</td>
<td>214 x 10²⁶</td>
<td>221 x 10²⁶</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>(Liquid)</td>
<td>215 x 10²⁶</td>
<td>151 x 10²⁶</td>
<td>174 x 10²⁶</td>
<td>184 x 10²⁶</td>
<td>---</td>
</tr>
<tr>
<td>≥100,000 atm</td>
<td>189 x 10²⁶</td>
<td>141 x 10²⁶</td>
<td>153 x 10²⁶</td>
<td>---</td>
<td>23,700 x 10²⁶</td>
</tr>
<tr>
<td>10,000 atm</td>
<td>177 x 10²⁶</td>
<td>134 x 10²⁶</td>
<td>145 x 10²⁶</td>
<td>---</td>
<td>2370 x 10²⁶</td>
</tr>
<tr>
<td>1000 atm</td>
<td>126 x 10²⁶</td>
<td>111 x 10²⁶</td>
<td>106 x 10²⁶</td>
<td>123 x 10²⁶</td>
<td>237 x 10²⁶</td>
</tr>
<tr>
<td>10 atm</td>
<td>25.5 x 10²⁶</td>
<td>68.9 x 10²⁶</td>
<td>24.5 x 10²⁶</td>
<td>21.7 x 10²⁶</td>
<td>23.7 x 10²⁶</td>
</tr>
<tr>
<td>1 atm</td>
<td>2.38 x 10²⁶</td>
<td>2.47 x 10²⁶</td>
<td>2.38 x 10²⁶</td>
<td>2.34 x 10²⁶</td>
<td>2.37 x 10²⁶</td>
</tr>
<tr>
<td>0.1 atm</td>
<td>0.237 x 10²⁶</td>
<td>0.238 x 10²⁶</td>
<td>0.237 x 10²⁶</td>
<td>0.237 x 10²⁶</td>
<td>0.237 x 10²⁶</td>
</tr>
</tbody>
</table>

| Molecular mass:  | 5.32 x 10^-26 kg  | 7.31 x 10^-26 kg     | 4.65 x 10^-26 kg   | 6.64 x 10^-27 kg | ---                      |
isothermal static compression of pure liquid water causes crystallization from the liquid to solid Ice VI near ~11,500 atm. Unlike water, most materials contract upon freezing, thus freezing point is normally raised by higher pressure. For example, the melting point of paraffin wax is 51.8 K at 1 atm but rises to 323.1 K at 100 atm.1657

At 273 K, gaseous carbon dioxide compressed to 35 atm suddenly liquefies; at 300 K, pressurized CO₂ remains liquid up to ~4500 atm, whereupon it solidifies.2035 However, T_{crit} = 304.2 K for CO₂, so if the gas is stored in pressure vessels at 310 K (human body temperature) > T_{crit}, the physical properties change continuously, showing no sign that the gas has condensed to a liquid, although over ~200 atm the fluid behaves like carbon dioxide liquid.2036 CO₂ compressed to 400,000 atm at 1800 K forms a translucent quartzlike extended covalent solid that is metastable down to ~room temperature at pressures >10,000 atm.3181 Other gases of nanomedical interest such as oxygen (T_{crit} = 154.4 K), nitrogen (T_{crit} = 126.1 K), hydrogen (T_{crit} = 33.3 K), and helium (T_{crit} = 5.3 K) show a similar continuous progression of gas/liquid properties at 310 K with rising pressure. Helium solidifies to ~1000 kg/m³ crystals at 115,000 atm at 297 K.2037 Room temperature isopropyl alcohol solidifies into a glassy state at ~44,000 atm, while a 4:1 methanol/ethanol mixture remains hydrostatic up to 104,000 atm.2051 Hydrogen solidifies at 57,000 atm at 298 K, making 600 kg/m³ crystals;2058 solid deuterium at 300 K compacts from ~620 kg/m³ at 100,000 atm to ~1400 kg/m³ at 1 million atm.2038

Vessel wall materials also are subject to pressure changes. Diamond crystal plastically deforms at a pressure of only ~1 atm at ~1773 K, but at ~1300 K, ductile flow in diamond requires pressures >60,000 atm.2041 At room temperature the diamond [111] surface compressed in the [110] direction is indented by pressures ≥880,000 atm; compression of the [001] surface requires 560,000-1 million atm for indentation, depending upon the direction of the applied load.1597 Colorless diamond takes on a light brown coloration at 1.15 million atm 2052. Liquid environment.10 Gases might also leak out by diffusion through pressure vessel walls. Hydrogen has the highest coefficient of diffusion of all the gases (Table 3.3). H₂ readily diffuses through porous substances such as clay, rubber, and even quartz and silica at elevated temperatures. Gases may diffuse through conventional metals, which have numerous defects, dislocations and grain boundaries, and through silicate glasses, which have open irregular structures that permit substantial diffusion of helium.10 Hydrogen also dissolves in and diffuses through the metals of the nickel, palladium and platinum groups. For example, Pd may dissolve ~900 volumes of H₂ at 293 K and 1 atm, and diffusion through a palladium thimble...
### Table 10.3. Biological and Biochemical Effects of High Pressure

<table>
<thead>
<tr>
<th>Biological Material</th>
<th>Pressure (atm)</th>
<th>Results of Compression to Very High Pressures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cell Division:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eggs (Arbacia)</td>
<td>100-400</td>
<td>Progressive reversible solution of gelated cortical cytoplasm</td>
</tr>
<tr>
<td>Eggs (annelid, echinoderm, vertebrate)</td>
<td>100-500</td>
<td>Cleavage-inhibiting pressure, increases with temperature (5-30˚C)</td>
</tr>
<tr>
<td>Eggs, misc. marine</td>
<td>200-400</td>
<td>Reversible inhibition of cleavage</td>
</tr>
<tr>
<td>Eggs (Urechis)</td>
<td>300-400</td>
<td>Reversible solution of spindles and asters; chromosome movement stopped</td>
</tr>
<tr>
<td>Eggs (Arbacia)</td>
<td>400</td>
<td>Reversible regression of cleavage furrows and cortical plasmagel solution</td>
</tr>
<tr>
<td>Eggs (Ascaria)</td>
<td>800</td>
<td>Exceptionally resistant to pressure; division not inhibited</td>
</tr>
<tr>
<td><strong>Cell Physiology:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amoeba population growth</td>
<td>136-340</td>
<td>Severely depressed; reversible after &gt;1 day at top pressure</td>
</tr>
<tr>
<td>Ciliary movement</td>
<td>200-400</td>
<td>Reversible increase in beat frequency with rise in pressure</td>
</tr>
<tr>
<td>Contraction of pigment cells (fish)</td>
<td>200-400</td>
<td>Contraction phase reversibly inhibited; totally inhibited at higher pressures</td>
</tr>
<tr>
<td>Protoplasmic streaming (Elodea)</td>
<td>200-400</td>
<td>Solation of cytoplasm; streaming slowed, stopped at 400 atm; reversible</td>
</tr>
<tr>
<td>Amoeboid movement (A. proteus)</td>
<td>200-500</td>
<td>Reversible plasmagel solution; pseudopodia collapsed; movement stopped</td>
</tr>
<tr>
<td>Heart rate (cultured tissue, frog)</td>
<td>300</td>
<td>Retardation at lower temperatures; acceleration at higher temperatures</td>
</tr>
<tr>
<td><strong>Cells/Tissues:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythrocytes</td>
<td>500-3000</td>
<td>Rounded</td>
</tr>
<tr>
<td>Carcinoma (Brown-Pearce)</td>
<td>1000</td>
<td>Resistant</td>
</tr>
<tr>
<td>Chick heart (embryonic tissue)</td>
<td>1000-1850</td>
<td>Reduction of subsequent growth in culture</td>
</tr>
<tr>
<td>Sarcoma transplants (rat)</td>
<td>1800</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Carcinoma (Brown-Pearce)</td>
<td>1800</td>
<td>Transplantability destroyed</td>
</tr>
<tr>
<td>Erythrocytes</td>
<td>5000</td>
<td>Disintegrated</td>
</tr>
<tr>
<td><strong>Bacteria/Fungi:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staphylococci and colon bacteria</td>
<td>3000</td>
<td>Unaffected</td>
</tr>
<tr>
<td>Bacillus anthracis</td>
<td>3000</td>
<td>Partial loss of virulence; death of vegetative cells</td>
</tr>
<tr>
<td>Streptococci</td>
<td>3000</td>
<td>Killed or retarded in growth</td>
</tr>
<tr>
<td>E. coli</td>
<td>3800</td>
<td>Inactivated after 10-minute exposure</td>
</tr>
<tr>
<td>Yeast cells</td>
<td>4000-6000</td>
<td>Death, preceded by cytoplasmic flocculation or coagulation</td>
</tr>
<tr>
<td>Bacteria (various species)</td>
<td>5000</td>
<td>Vegetative cells killed</td>
</tr>
<tr>
<td>Bacterial spores</td>
<td>12,000</td>
<td>Killed</td>
</tr>
<tr>
<td><strong>Viruses:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rous sarcoma, Shope papilloma</td>
<td>1800</td>
<td>Tumors delayed</td>
</tr>
<tr>
<td>Staphylococci bacteriophage</td>
<td>2000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Yellow fever (monkey)</td>
<td>3000</td>
<td>Slightly attenuated</td>
</tr>
<tr>
<td>Herpes (rabbit)</td>
<td>3000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Tubercocci necrosis</td>
<td>3000-5000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Rabies (rabbit)</td>
<td>4000</td>
<td>Attenuated</td>
</tr>
<tr>
<td>Papilloma Sarcoma filtrate</td>
<td>4000</td>
<td>No tumors</td>
</tr>
<tr>
<td>Avian pest (fowl plague)</td>
<td>4000</td>
<td>Inactivated, but retained antigenicity</td>
</tr>
<tr>
<td>Foot and mouth disease</td>
<td>4000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Vaccinia (rabbit)</td>
<td>4500</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Rabies (rabbit)</td>
<td>5000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Bacteriophage (B. megatherium)</td>
<td>7000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Bacteriophage (B. subtilis, B. typhosa)</td>
<td>7000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Encephalomyelitis (rabbit)</td>
<td>7000</td>
<td>Inactivated</td>
</tr>
<tr>
<td>Tobacco mosaic virus</td>
<td>7500</td>
<td>Inactivated; coagulated</td>
</tr>
<tr>
<td><strong>Antigens/Antibodies:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tuberculin, Cobra venom</td>
<td>13,500</td>
<td>Not destroyed; lethality unchanged</td>
</tr>
<tr>
<td>Equine tetanus antitoxin</td>
<td>13,500</td>
<td>Gelated; partially inactivated</td>
</tr>
<tr>
<td>Diptheria toxin</td>
<td>13,500</td>
<td>Partially destroyed</td>
</tr>
<tr>
<td>Tetanus toxin</td>
<td>13,500</td>
<td>Greatly attenuated or inactivated</td>
</tr>
<tr>
<td>Diptheria toxin</td>
<td>17,600</td>
<td>Mostly destroyed</td>
</tr>
<tr>
<td><strong>Enzymes:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ribonucleode polymerase</td>
<td>6000</td>
<td>Reversible diminution in activity</td>
</tr>
<tr>
<td>Pepsin, Rennin</td>
<td>6000-7600</td>
<td>Inactivated or largely inactivated</td>
</tr>
<tr>
<td>Chymotryptsinogen, Trypsin</td>
<td>7600</td>
<td>Partially inactivated</td>
</tr>
<tr>
<td>Amylase, Lipase, Sucrase, Lactase</td>
<td>8000-15,000</td>
<td>Some completely, others partially inactivated</td>
</tr>
<tr>
<td><strong>Proteins:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gelatin</td>
<td>2000</td>
<td>Gelation accelerated</td>
</tr>
<tr>
<td>Gelatin gel</td>
<td>3000</td>
<td>Water squeezed out</td>
</tr>
<tr>
<td>Equine serum globulin</td>
<td>3000-13,000</td>
<td>Gelated</td>
</tr>
<tr>
<td>Soy protein</td>
<td>4000</td>
<td>Completely coagulated</td>
</tr>
<tr>
<td>Egg albumin</td>
<td>5000</td>
<td>Slight stiffening</td>
</tr>
<tr>
<td>Egg albumin</td>
<td>7000</td>
<td>Completely coagulated</td>
</tr>
<tr>
<td>Egg albumin</td>
<td>7500</td>
<td>Denatured and coagulated; SH groups exposed</td>
</tr>
<tr>
<td>Carboxyhemoglobin</td>
<td>9000</td>
<td>Coagulated</td>
</tr>
<tr>
<td>Insulin</td>
<td>10,000</td>
<td>Coagulated, but not physiologically inactivated</td>
</tr>
</tbody>
</table>
is often used in the laboratory to purify hydrogen gas. H$_2$ diffusion through bulk metals has been widely studied.\textsuperscript{3216,3217}

However, diffusion leakage through defect-free pressure vessel walls comprised of diamond, graphene, or corundum should be negligible. "It is difficult to see how a molecule can pass through these materials without undergoing a reactive transformation,"\textsuperscript{210} The energy required to move atomic hydrogen from free space into a minimum-energy site in diamond is thermally prohibitive (\(\approx 800\) zJ),\textsuperscript{2053} and the required energy is even higher for molecular hydrogen. Room temperature gases at \(-1300\) atm\textsuperscript{1169} and argon gas at \(-600\) atm\textsuperscript{2034} have been stably trapped inside carbon nanotubes, and helium atoms at \(\sim 1300\) atm\textsuperscript{1169} and argon gas at \(\sim 600\) atm\textsuperscript{2034} have been stably trapped inside carbon nanotubes, and confirmed by computer simulations.\textsuperscript{3212-3215} and helium atoms at energies up to \(-5\) eV \((-800\) zJ\)) cannot penetrate a graphene sheet. Ruby (corundum) is commonly used in diamond anvil experiments in which hydrogen is compressed to pressures \(>10^6\) atm with no apparent leakage.\textsuperscript{2043} Thus it appears that essentially leakproof pressure vessel walls can be constructed, although no computational studies of diamond/sapphire interfaces had been performed by 1998 [D.W. Brenner, personal communication, 1999]. Double-walled structures with an intervening space containing either getter materials or a pumped vacuum may be used, if necessary, for added safety.\textsuperscript{10}

Microscopic pressure vessels comprised of diamond and filled with compressed hydrogen or oxygen are potentially flammable, even explosive. However, the chemical energy contained within a volume scales as \(-L^3\) while the dissipation of that energy during an explosive event (e.g., pressure, heat, light) occurs across a surface which scales as \(-L^2\), hence energy dissipation per unit area (and thus the relative impact of an explosion on the local environment) scales as \(-L^{-1}\) which implies declining severity at smaller device sizes (Chapter 17). For example, a liter of nitroglycerine produces a blast intensity of \(-10^8\) J/m$^2$, while a spherical 1 micron$^3$ of nitroglycerine gives only \(-10^7\) J/m$^2$ upon decomposition.

Heat from operating valves is small and is rapidly conducted away in a diamondoid structure. Corundum employed at all structure/oxidant interfaces virtually eliminates device combustibility. Even for bulk diamond in contact with pure oxygen gas at 1000 atm, the time required to etch a single layer of carbon atoms of thickness \(-0.17\) nm at 310 K may be crudely estimated from Eqn. 9.55 as \(-10^{11}\) years; at 400 K, the atomic-layer etch time is still \(-400\) years. Further discussion of possible nanodevice combustibility is deferred to Chapter 17.

### 10.3.5 Vacuum Pumping and Storage

It will frequently be necessary to establish and maintain vacuum conditions inside nanorobots. Vacuum-enclosing shells were briefly considered in Section 9.5.3.3 and were found incapable of providing vacuum may be surprisingly thin-walled. In the simplest case, consider a circular cylindrical diamondoid tube of inside radius $r_{tube}$ wall thickness $h_{tube}$ Young's modulus $E = 1.05 \times 10^{12}$ N/m$^2$ and Poisson's ratio $\nu_{Poisson} = 0.1$.\textsuperscript{10} When the relative external pressure on the hull is so large that the cylinder becomes neutrally unstable, a small deformation into an elliptical cylinder is possible. The critical pressure of buckling is given by: \textsuperscript{361}

$$P_{crit} = \frac{E h_{tube}^3}{4 r_{tube} (1 - \nu_{Poisson})} \quad \text{[Eqn. 10.20]}$$

Taking $r_{tube} = 500$ nm and $P_{crit} \leq 1$ atm for a vacuum vessel surrounded by normal atmospheric pressure, a hull thickness of $h_{tube} \approx 3.7$ nm will not buckle. Such a hull has a mass \(-60\) times greater than the mass of the displaced air.

How fast can the vacuum be established? Consider a cubical volume of $V_{box} = L_{box}^3 = 1$ micron$^3$ containing gas at molecular number density $n_{gas}$ which is to be evacuated. Molecular sorting rotors (Section 3.4.2) are embedded in the six walls, with each rotor mechanism of area $A_{rotor} = 100$ nm$^2$ having at least one binding site of area $A_{receptor} = 0.1$ nm$^2$ always exposed to the box interior. There are $N_{rotor} = (6 L_{box}^2 / A_{rotor}) = 60,000$ sorting rotors in the box walls. Gas molecules of thermal velocity $v_{thermal}$ (Eqn. 3.3) nearest the walls travel one mean free path $\lambda_{gas}$ (Eqn. 9.23) and impact a rotor mechanism. On average, after $N_{receptor}/A_{receptor}$ impacts a binding site is hit; after $n_{encounter} = 10$ independent hits, the molecule is finally captured by the receptor. Thus the time required for a receptor to capture a gas molecule is $t_{bind} = n_{encounter} \lambda_{gas} A_{rotor} / v_{thermal} A_{receptor}$ and the time required to evacuate the box is approximately:

$$t_{vac} = \frac{n_{gas} V_{box}}{N_{rotor} \lambda_{gas}} t_{bind} = \frac{n_{gas} V_{box} L_{box}^2}{N_{rotor} A_{rotor}^2} \eta_{thermal} A_{receptor} \rho_{gas} \frac{L_{box}^2}{v_{thermal} A_{receptor}}$$

\textsuperscript{[Eqn. 10.21]}

At 1 atm and 310 K, $n_{gas} = 2.4 \times 10^{25}$ molecules/m$^3$ (Table 10.2) and $\lambda_{gas} = 200$ nm; for pressures \(<0.2\) atm, $\lambda_{gas} = L_{box}$ in the ballistic regime. Taking $\lambda_{gas} = 1$ micron and $v_{thermal}$ for O$_2$ at 310 K gives $t_{bind} \approx 20$ microseconds and the box is evacuated in $t_{vac} \approx 10$ milliseconds. Note that for microscale vessels at atmospheric pressure, $t_{vac}$ scales as $\sim L_{box}^{-2}$.

Tank wall material sublimes into the vacuum, the solid establishing equilibrium with its vapor at a temperature-dependent vapor pressure. However, this process produces negligible effluent at moderate temperatures for likely nanorobot building materials. A volumetric number density of one wall-material atom per micron$^3$ of vacuum, representing a minimum detectable contamination in a -1 micron$^3$ box, equals a contaminant partial pressure of $4 \times 10^{-8}$ atm. To reach this vapor pressure, carbon must be heated to 2480 K, aluminum to 1040 K (liquid), silver and corundum (e.g., ruby, sapphire) to 800 K, and even zinc (a soft, low melting point, metal) to 500 K.\textsuperscript{763} At 310 K, the vapor pressure of tank-wall carbon is only $10^{-116}$ atm, aluminum $10^{-39}$ atm, corundum $10^{-19}$ atm, and zinc $10^{-16}$ atm. Many organic materials evaporate or sublime more readily. At a vapor pressure of 0.001 atm, ethanol (C$_{2}$H$_{5}$OH) vaporizes at or above 241.9 K, vs. 253 K for water-ice, 259.2 K for octane (C$_{8}$H$_{18}$), 313.1 K for phenol (C$_{6}$H$_{5}$OH), but 426.8 K for palmitic acid (C$_{17}$H$_{35}$COOH), a typical fatty acid.\textsuperscript{763} A water-ice surface exposed to hard vacuum sublimes at the rate of 6.5 nm/day at 134 K, 1.4 microns/day at 152 K, and 1.2 mm/day at 183 K.\textsuperscript{3220}

### 10.3.6 Buoyancy Control and Nanopheresis

Another design issue that may arise when operating in an aqueous medium is buoyancy, which can readily be controlled by loading or unloading ballast. At the extremes, a diamondoid sphere of radius 500 nm and wall thickness 5 nm may range in density from \(-100\) kg/m$^3$ if filled with vacuum to \(-9000\) kg/m$^3$ if densely packed with diamondoid machinery. Placed in blood plasma, these spheres would rise at \(-0.5\) microns/sec or fall at \(-1\) micron/sec, respectively, near the local gravity field, according to Stokes' Law for Sedimentation (Eqn. 3.10), taking $\eta_{plasma} = 1.025$ kg/m$^3$ and plasma absolute viscosity as $\eta_{plasma} = 1.1 \times 10^{-3}$ kg/m-s at 310 K. This range of speeds lies near the low end of cytonatation velocities (0.1-10 microns/sec; Section 9.4.6) and well
below the maximum sanguinatation velocity (~10,000 microns/sec; Section 9.4.2.6), which suggests that buoyancy control is unlikely to play a significant role during routine in vivo nanorobot locomotion. By comparison, the small difference in density between individual red blood cells (1100 kg/m³) and blood plasma at 310 K causes them to settle out of suspension at a slightly faster 1-3 microns/sec depending on hematocrit (the volumetric percentage of blood occupied by red cells, typically ~46% in humans; Section 9.4.1.2) and degree of RBC aggregation. Free-floating natural erythrocytes appear unhandicapped by their faster settling rate, so active ballast management for free-floating artificial nanorobots is probably unnecessary in normal operations.

However, once a therapeutic purpose is completed it may be desirable to extract artificial devices from circulation. Active ballast control may be extremely useful during nanorobot expulsion from the blood, particularly in the case of simpler devices which are incapable of removing themselves from the body (Chapter 16). Blood to be cleared may be passed from the patient to a specialized centrifugation apparatus analogous to an apheresis circuit such as a cytopheresis or plasmapheresis system. In nanapheresis, acoustic transmitters command passing nanorobots to establish neutral buoyancy. No other solid blood component can maintain exact neutral buoyancy, hence those other components precipitate outward during the gentle centrifugation and are drawn off and added back to filtered plasma on the other side of the apparatus. Meanwhile, after a period of centrifugation, the plasma, containing mostly suspended nanorobots but few other solids, is drawn off through a simple filter, removing the devices as residue. Plasma filtrate is then recombined with centrifuged solid components and returned undamaged to the patient's body.

### 10.4 Cytocide and Virucide

The objective of nanomedical treatment is often cell repair or replacement (Chapter 21). However, it will frequently be necessary for medical nanorobots to destroy foreign viruses and pathogenic bacteria, protozoa, or metazoan parasites. It may also become necessary to eliminate native human tissue cells that have grown too numerous in a local region, cells that have adopted or developed major pathological, teratological or other undesirable patterns, or which have been too severely traumatized to permit efficient repair.

There are at least two classes of nanomedical cytocidal instrumentalities that may be employed to rid the human body of unwanted cells. The first may be termed the biochemical approach (Section 10.4.1), including methods which may take up to hours or days to proceed to completion with acceptable reliability. Biochemical methods can require relatively modest technological sophistication to implement—an important consideration in the early years of nanomedical device development. The second method may be termed the mechanical approach (Section 10.4.2), which may require only seconds or minutes to proceed to completion with very high reliability, but which involves more sophisticated nanomechanisms that might only be available in a mature nanomedical technology environment.

#### 10.4.1 Biochemical Cytocide and Virucide

Biochemical cytocide involves the killing of pathogenic cells using biochemical, cytochemical, or other means which cause the cell to die a "natural," though perhaps highly accelerated, death. Any agent or set of conditions that stresses the metabolism or normal response mechanisms of a cell may trigger the process of apoptosis, or programmed cell death, but the level of stress is crucial.

At high stress levels, including sharp pH changes or high toxin levels (Section 10.4.1.4), or high agitation rates, cells may die instantly by necrosis, largely because they have no time to respond to the stimulus. In necrotic cytocide, neighboring cells are showered with cellular debris as the necrotic cell swells and lyses, an event likely to cause wider injury and a process to be avoided in nanomedical situations wherever possible.

At intermediate levels of cell stress, the cell is injured but not killed, and so has time to activate its own suicide program. The contents of the dying cell are retained within sealed vesicles until removed through phagocytosis, thus the cell dies in a controlled and tissue-friendly way.

At low levels of environmental stress, cells can, for example, switch on the production of heat shock proteins which helps them to survive until the stress is removed largely by accelerating the refolding of proteins which have been damaged by heat and some other insults. Similarly, UV radiation activates a DNA damage response, cell starvation activates the proteasome protein-recycling response, and glucocorticoids might activate responses to eliminate all non-essential functions.

Of course, once a certain stress threshold is passed and survival is deemed impossible, death by apoptosis ensues. Nanomedical bio-cytocidal instrumentalities should seek to dispose of cells by apoptotic methods (Section 10.4.1.1), or by functionally equivalent techniques such as phagocytic flagging (Section 10.4.1.2) or cell division arrest (Section 10.4.1.3).

#### 10.4.1.1 Apoptosis

It is believed that all nucleated (eukaryotic) cells of all multicellular life on Earth incorporate an evolutionarily conserved self-destruct mechanism called programmed cell death, cell suicide, or simply apoptosis. Scattered reports on cell death have appeared in the literature (and were initially disbelieved) for more than a century, but advances in cytochemistry in the 1980s and early 1990s stimulated an explosion of interest in apoptosis, with ~2500 papers published during 1989-1994 and ~20,000 publications during 1994-1998.

Exactly what is apoptosis? Programmed cell death is the outcome of a programmed intracellular cascade of genetically determined steps. During apoptosis, the eukaryotic cell disassembles its DNA and breaks up its contents into membrane-wrapped packets which can then be cleared away without causing inflammation. Animals use apoptosis to eliminate extraneous, virus-infected, or otherwise dangerous cells. Apoptosis plays a central role both in development and in homeostasis of metazoans. Cells die by apoptosis in the developing embryo during morphogenesis or synaptogenesis, and in the adult animal during tissue turnover in the skin or gut or at the end of an immune response. Tight coupling of cell death and cell multiplication ensures in some tissues a constant, controlled flux of fresh cells which are crucial to the preservation and optimal functioning of the adult organism. Such tissues include those which are environmentally exposed, inside or outside, primarily epithelial tissue and severely insulted liver tissue, plus cells involved in reproductive function.

Apoptosis requires specialized machinery. The central component of this machinery is an irreversible proteolytic system involving a family of protein-chopping enzymes now called "caspases" from "cysteine-containing aspartate-specific proteases" (aka ICE or Interleukin-1β-Converting Enzyme). Caspases are expressed as 30-50 kD proenzymes with an NH₂-terminal domain, a large ~20 kD subunit, and a small ~10 kD subunit. These enzymes...
participate in a cascade that is triggered in response to proapoptotic signals and culminates in cleavage of a set of proteins, ultimately resulting in an orderly disassembly of the cell.\textsuperscript{2068}

Such complex proteolytic systems involve a combination of regulatory proteases, cofactors, feedbacks, and thresholds that converge to control the activity of an effector protease, which in turn carries out the function of the whole process.\textsuperscript{2071} These systems keep the effector protease inactive but are able to rapidly activate large amounts of it in response to minute quantities of an appropriate inducer.\textsuperscript{2068} Survival signals from the cell's environment and internal cellular integrity sensors normally hold a cell's apoptotic machinery in check.

Four kinds of events can trigger a suicidal cascade in a eukaryote.\textsuperscript{2069} First, if a cell loses its normal contact with its surroundings or sustains irreparable internal damage, then that cell initiates apoptosis. Second, a cell that simultaneously receives conflicting signals driving or attenuating its division cycle also undertakes apoptosis.\textsuperscript{2074} Third, the immune system can actively direct individual cells to self-destruct, an event called "instructive" apoptosis.\textsuperscript{2076-2078} Fourth, perhaps a majority of all apoptosis occurs during gestation and development, probably in response to external signal molecules or to a lack of sufficient levels of molecular signal gradients involved in biostuctural development regulation.

Death receptors—cell surface receptors that transmit apoptosis signals initiated by specific death ligands—play a central role in instructive apoptosis. Death receptors can activate death caspases within seconds of ligand binding, causing the apoptotic demise of the cell within hours.\textsuperscript{2069} Death receptors belong to the tumor necrosis factor (TNF) receptor gene superfamily. (TNF is produced mainly by activated macrophages and T cells in response to an infection.)\textsuperscript{2069,2079} The death receptors contain a homologous cytoplasmic sequence termed the "death domain."\textsuperscript{2080,2081} Death domains enable death receptors to engage the cell's apoptotic machinery, and mediate functions that are distinct from or even counteract apoptosis.\textsuperscript{2069} In 1998, the best characterized death receptors were CD95 (aka Fas)\textsuperscript{2066} and Apo1 and CD120a (aka TNFR1, Tumor Necrosis Factor Receptor 1, or p55), but others were known including avian CAR1, Death Receptor 3 (DR3; aka Apo3, WSL-1, TRAMP, LARD), DR4, and DR5 (aka Apo2, TRAIL-R2, TRICK-2, KILLER).\textsuperscript{2069}

When a caspase cascade is triggered, the first task is to inactivate proteins that protect living cells from apoptosis.\textsuperscript{2080} This accomplished, caspases next begin the direct disassembly of cell structure. One example is the destruction of the nuclear lamina (Section 8.5.4.3.), the rigid structure underlying the nuclear membrane that is involved in chromatin organization. Lamina is formed by head-to-tail polymers of intermediate filament proteins called lamins (Section 8.5.3.11.). During apoptosis, lamins are cleaved at a single site by caspases, causing the lamina to collapse, contributing to chromatin condensation.\textsuperscript{2082,2083}

Caspases also reorganize cell structures indirectly by cleaving several proteins involved in cytoskeletal regulation, including gelsolin,\textsuperscript{2084} focal adhesion kinase (FAK),\textsuperscript{2085} and p21-activated kinase 2 (PAK2).\textsuperscript{2086} Cleavage of these proteins results in deregulation of their activity.\textsuperscript{2086} Dissociation of regulatory and effector domains is another hallmark of caspase function. For example, they inactivate or deregulate proteins involved in DNA repair (such as DNA-PKcs), mRNA splicing (such as U1-70K), and DNA replication (such as replication factor C).\textsuperscript{2087,2088}

In 1998, 13 mammalian caspases—named caspase-1 to caspase-13—were known.\textsuperscript{2067,2068} Caspases cut their substrate proteins at tetrapeptide sites with high specificity and with characteristic motifs.\textsuperscript{2089} For example, caspase-3 recognizes DEVD sequences and, during apoptosis, cleaves and inactivates several significant cellular proteins in the cytosol, nucleus and cytoskeleton.\textsuperscript{2090-2092} Caspases participate in apoptosis "in a manner reminiscent of a well-planned and executed military operation."\textsuperscript{2068} They:

1. cut off contacts with surrounding cells, so the cell balls up;
2. import calcium,\textsuperscript{2075} strongly complexing phosphates;
3. shut down DNA replication and repair, interrupt splicing, disrupt the nuclear structure and condense the chromatin;
4. induce the loss of microvilli;\textsuperscript{2093}
5. destroy cellular DNA by mobilizing apoptotic-unique nuclease;\textsuperscript{2091,2092} cleave the double helix at regular intervals in an orderly fashion, first into large fragments of 300-750 kilobases, followed by fragments of ~50 kilobases, then finally oligonucleosomal length DNA fragments of ~200 base pairs\textsuperscript{2062-2064} (giving a characteristic pattern by gel electrophoresis called "DNA laddering")\textsuperscript{2094};
6. reorganize the cytoskeleton;
7. induce the cell to display signals in the outer membrane that mark the cell for phagocytosis by neighboring cells or macrophages;\textsuperscript{2095}
8. induce downstream blocking of mitochondrial respiratory chain components, resulting in mitochondrial malfunction (via transcriptional induction of redox related genes, the formation of reactive oxygen species, and finally the oxidative degradation of mitochondrial components\textsuperscript{2072,2073} including dissipation of mitochondrial inner transmembrane potential and the release of cytochrome c through the outer mitochondrial membrane\textsuperscript{2162}, with lysosomal involvement;\textsuperscript{2093}
9. initiate membrane blebbing (surface blisters) and cell shrinkage; and finally
10. disassemble the cell into multiple membrane-enclosed vesicles called apoptotic bodies without destroying organelle membranes.

In vivo, this process culminates with the engulfment and phagocytization of apoptotic bodies by other cells, preventing inflammation and other complications that would result from a release of intracellular contents. These changes occur in a predictable, reproducible sequence and can be completed within 30-60 minutes.\textsuperscript{2068}

Many different ways of activating the apoptotic process are known or suspected.\textsuperscript{2096} Simple detachment of tissue cells from all contacts with the ECM\textsuperscript{1553} or manipulation of cell shape\textsuperscript{218} have been shown to induce apoptosis experimentally. Apoptosis can also be indirectly initiated by a variety of cellular insults that can damage DNA, including ultraviolet- and x-irradiation, hypoxia, and chemotherapeutic neoplastic drugs (e.g., alkylating agents such as nitrogen mustard derivatives, antimetabolites and plant alkaloids) which inflict cell damage that is then translated via the Bcl-2 protein family\textsuperscript{2097} through several poorly understood steps into activation of caspase-9.\textsuperscript{2068} Proteins that sense DNA damage and help trigger apoptosis also affect the cell cycle—stopping cell division so the damage can be repaired or making the decision (with the activation of tumor suppressor p53) that the damage has gone too far and the cell must die.\textsuperscript{2074} B and T lymphocytes undergo apoptosis in response to anti-IgM antibodies and dexamethasone (a glucocorticoid), respectively.\textsuperscript{2098} The cell death program can also be activated by
inhibition of proteasome function,\textsuperscript{2099} or by infection with a wide variety of viruses.\textsuperscript{2118}

Caspases can be directly activated by two distinct intracellular mechanisms that interact with the death receptor complexes. First, because all caspases have similar cleavage specificity, the simplest way to activate a procaspase is to expose it to a previously activated caspase molecule. This caspase cascade is used extensively by cells for the activation of the downstream effector caspases: caspase-3, caspase-6, and caspase-7.\textsuperscript{2068}

A second method for activating caspases is the use of apoptotic chaperones, which help to bind inactive proenzymes to increase their local concentration and ease them into conformations that promote their activation. This “induced proximity” was first observed in caspase-8 (aka FLICE, MACH), an initiator caspase that acts downstream of the CD95 death receptor.\textsuperscript{2069} Upon trimeric ligand binding, CD95 receptor molecules are aggregated into a membrane-bound complex. This signaling complex recruits, via the receptor-bound adaptor protein FADD (Fas-Associated protein with Death Domain; aka Mort-1\textsuperscript{2100}), several procaspase-8 molecules, resulting in a high local concentration of procaspase-8. Under these conditions, the low protease activity inherent to procaspases is sufficient to drive intermolecular proteolytic activation of the receptor-associated procaspase-8 molecules.\textsuperscript{2068} Proximity-induced activation may also be used to activate caspase-9.\textsuperscript{2101} Procaspase-9 activation involves a complex with the cofactor Apaf-1 through the CARD (Caspase Recruitment Domain), but activation of caspase-9 also requires cytochrome c (released into the cytosol by mitochondria\textsuperscript{2073}) and deoxyadenosine triphosphate, indicating that caspase activation may require multiple cofactors.

Death receptors can be expressed on both normal and cancerous cells in the human body, so the challenge for conventional drug-based therapy is to find some way to activate death receptors selectively on cancer cells only.\textsuperscript{2068,2069} For medical nanorobots, such selectivity should be simple and routine using multiple chemosensors (Section 4.2), a benefit that is characteristic of most nanorobot-based therapeutics. If caspase cascade amplification is sufficient to permit single-site activation of the cascade, then in theory an extracellular nanorobot intending cytocide could press onto the outer surface of a target cell an appropriate ligand display tool. This tool might contain suitably exposed trimeric CD95L (aka FasL\textsuperscript{2069}) ligand (binds to the extracellular domains of three CD95 death receptors), TNF or lymphotoxin \(\alpha\) (binds to CD120a), Apo3L ligand aka TWEAK (binds to DR3), or Apo2L ligand aka TRAIL (binds to DR4 and DR5).\textsuperscript{2069,2070} The binding event would then activate a single death receptor complex, potentially triggering the entire irreversible cytocidal cascade. If necessary, multiple display tools could be employed. This technique avoids the storage requirement for bulky consumables onboard the medical nanorobot.

As another approach, molecular sorting rotors could be used to selectively extract from the cytoplasm specific crucial molecular species of IAPs (Inhibitors of Apoptosis\textsuperscript{2105}) that can hold the apoptotic process in check. Examples include survivin, commonly found in human cancer cells\textsuperscript{2058}, the transcription factor NF-\(kB\)\textsuperscript{2142}, and Akt, which delivers a survival signal that inhibits the apoptosis induced by growth factor withdrawal in neurons, fibroblasts, and lymphoid cells\textsuperscript{2101}. Conversely, decoy receptors (DcRs) that compete with DR4 and DR5 for binding to Apo2L\textsuperscript{2104-2106} could be saturated with intrinsically harmless but precisely engineered intracellular “chaff” ligands. With IAPs’ removed or DcRs blocked, apoptosis may be free to proceed.

Multicellular plants and animals,\textsuperscript{2074,2075} human fibroblasts,\textsuperscript{2107} and even slime molds show various forms of apoptotic cell death. Other processes of programmed cell death that may be distinct from either apoptosis or necrosis have been reported in lung fibroblasts,\textsuperscript{2108} the caspase-free yeast \textit{Saccharomyces cerevisiae},\textsuperscript{2109} and other cells.\textsuperscript{2109-2112} In 1998, apoptosis had not yet been reported in any bacterial species, nor was it expected because conventional evolutionary logic does not favor the emergence of cell suicide in the case of obligate single-celled organisms.\textsuperscript{2075,2076} Autolysis has been reported in bacteria in specialized circumstances,\textsuperscript{2113-2115,2106} including addiction modules\textsuperscript{2118-2120} and phage exclusion,\textsuperscript{2120-2122} though these cases, while sometimes termed “programmed cell death,”\textsuperscript{2111,2105-2106} all involve a (messy) necrotic cytocidal outcome or self-damage short of cell death, rather than a (clean) apoptotic cytocidal outcome. Apoptosis also is not found in viruses.

It is theoretically possible that an artificial plasmid could be designed that embodies an engineered self-contained apoptotic system capable of unleashing a non-necrotic cell death process inside a prokaryote. A single such proapoptotic artificial plasmid (similar to artificial chromosomes,\textsuperscript{2432} already in use by 1998) could then be injected into the target bacterium, giving the desired clean cytocidal outcome. R. Bradbury suggests the following useful components:

1. a restriction enzyme to fragment the bacterial DNA, that is effective against that specific bacteria;
2. DNase to digest the DNA entirely, if not already present;
3. RNase to digest the RNA entirely; if not already present;
4. one or more proteases such as trypsin, to digest internal proteins; and
5. one or more lipases to digest any lipids.

A lambda-phage container holding an engineered vector with only the first two capabilities could still effectively “neuter” the bacterium. Such a biorobot could be relatively harmless to eukaryotic cells, assuming its enzymes lacked nuclear localization targeting sequences; cleanup could be effectively handled by the immune system. Bradbury also suggests another approach—using a DNA replication blocker such as antisense DNA polymerase, in combination with a mitotic promoter, to induce the cell to attempt to divide forever while lacking any DNA in the progeny. Essential biomolecules would eventually be diluted to such a low level that growth would radically slow and perhaps even death would result. This second approach might yield a faster decline in bacterial activity than simple neutering, due to the diversion of bacterial resources in an attempt to regain intracellular molecular balances. In either approach (as with mechanical approaches; Section 10.4.2), it is necessary to avoid fragmenting the bacterium in a way which allows release of superantigen molecules such as LPS which could cause an immune system overreaction, leading to toxic shock.

10.4.1.2 Phagocytic Flagging

Target cells may be flagged with biochemical substances capable of triggering a reaction by the body’s natural defensive or scavenging systems. For example, novel recognition molecules are expressed on the surface of apoptotic cells. In the case of T lymphocytes, one such molecule is phosphatidylinerine, a lipid that is normally restricted to the inner side of the plasma membrane (Section 8.5.3.2) but, after the induction of apoptosis, appears on the outside.\textsuperscript{2117} Cells bearing this novel surface molecule are then capable of being recognized and removed by phagocytic cells. Seeding the outer wall of a target cell with phosphatidylinerine or other molecules with similar action could activate phagocytic behavior by macrophages which
had mistakenly identified the target cell as apoptotic. Loading the target cell membrane surface with B7 costimulator molecules permits T-cell recognition, allowing an immunologic response via what has been termed the immunological synapse. This tagging operation should work well against cells which have an apoptotic response which can be triggered by cytotoxic T cells—such as human cancer cells and cysts.

Aside from false apoptosis, removal or destruction of cell-surface MHC class I receptors might convince the immune system that the target cell was not native (e.g., “self”), making the target susceptible to attack by natural killer NK lymphocytes. (Macrophages should be able to engulf a eukaryotic cell, as they are large enough to absorb protozoa.) Alternatively, the target cell's surface would be salted with a non-self antigen-containing MHC Class I molecule, or with self MHC Class I molecules with foreign peptide attached (particularly one for which the patient has already been vaccinated), with the expectation that the cell would mistakenly be identified as foreign, and thus be phagocytosed. Other methods of engineering chemical reactivity by cell surface remodeling are readily envisioned, such as flagging with lymphokinones or oligohistidine-tagged fusion proteins. But the best approach may be to seed the target cell with antigens for which the patient already has a high titer of circulating antibodies, which then will bind to the antigens and attract the immune system. Antibodies occur in higher concentrations than B/T cells and therefore would presumably trigger a faster response. As a general rule, cells with bad MHC molecules are targeted for apoptosis, while cells with antibodies attached are targeted for phagocytosis.

### 10.4.1.3 Cell Division Arrest

In some circumstances it may be useful to eliminate a target cell's ability to replicate via cell division, thus rendering it incapable of further growth. For example, many prokaryotic microorganisms respond to starvation by entering growth-free stationary phases or by forming dormant spores; nanorobot manipulation of cellular biochemical states possibly could induce stasis in bacteria, though spore formation should be avoided. The Bcl-2 protein family can modulate eukaryotic cell cycle progression; under suboptimal growth conditions, Bcl-2 promotes exit into quiescence and retards reentry into cycle. A gene fragment on human chromosome 4, called mortality factor 4 (morf4), when added to cancerous (eukaryotic) cells in vitro converts proliferating cells into senescent cells that have stopped dividing. Ceramide, the breakdown product of sphingomyelins, can drive a eukaryotic cell either to cycle arrest, or to apoptosis, depending upon other conditions within the cell, and p16 and p27 are known eukaryotic cell cycle inhibitors. Cytochalasins, which are derived from molds, interfere with the division of cytoplasm, inhibit cell movement, and can cause extrusion of the interphase nucleus.

Eukaryotic mitosis has four distinct stages (prophase, metaphase, anaphase, telophase) and interphase (between divisions), each associated with distinct biochemical and cytoarchitectural states. Thus it is not surprising that specific biochemical agents have been found that are capable of selectively halting, accelerating, or altering mitosis at any of the many stages; a few examples are in Table 10.4. Cell cycle control is discussed further in Chapter 12.

Many chemotherapeutic agents function by eliminating the cell's ability to undergo mitosis. For example, piclumycin (mithramycin) is believed to form a complex with DNA which inhibits DNA-dependent or DNA-directed cellular RNA and enzymatic RNA synthesis—48-hour lethal dose in human HeLa cancer cells is ~1 part/million by weight or ~10^5 molecules/cell. Eukaryotic cell cycling may be interrupted by inhibitors of RNA polymerases I, II, and III (which produce RNAs specific to different functions in the cell), or by inhibitors of uracil synthesis or other steps in general RNA base production which would interfere with RNA synthesis overall. Vincristine applied at similar dosage (~10^-18 kg/cell) inhibits microtubule formation in the mitotic spindle, resulting in an arrest of dividing cells at the metaphase stage. Large doses lead to necrotic cytocide; a smaller dose may trigger apoptosis.

A theoretical enzyme capable of selectively lysing DNA telomeres in the nucleus would prevent cell replication, but would probably trigger apoptosis due to massive chromatin damage because TTAGGG sequences are scattered throughout the chromosome. Telomere length affects eukaryotic cell division only indirectly, either because short telomeres activate a DNA damage response that blocks cell division or because genes near the ends of chromosomes which are essential for cell division may be lost or misregulated. Telomerase enzymes would be useless against common bacteria (e.g., E. coli) and viruses (e.g., SV40) which have circular DNA chromosomes, or viruses such as the adenoviruses that cause bronchiitis and pneumonia which have linear DNA locked with a terminal protein—but no telomeres in either case. Alternatively, R. Bradbury suggests the introduction of DNA (or RNA) for antisense sequences, for genes essential for cell cycle progression or for proteases which cut or digest these same proteins, as for instance an antisense gene for cyclin or one of the many cdc genes—if phosphorylation or dephosphorylation of the cyclins or cdc genes is occurring, the introduction of a phosphatase or kinase cadid easily block the entire cell division process.

### 10.4.1.4 Chemical Poisoning

There are a great many (often inefficient) ways to poison a cell or virion simply by releasing chemical agents near, on, or within it, but if done incautiously this approach may lead to necrotic, rather than apoptotic, cell death. Here are a few examples of some well-known classes of biocidal agents, many of which may be inappropriate for use by in vivo medical nanorobots:

1. **Phagosomal Biochemicals** — Both types of phagocytic cells (e.g., neutrophils and macrophages) contain specialized organelles that fuse with newly formed phagocytic vesicles (phagosomes), exposing phagocytosed microorganisms to a barrage of enzymatically produced, highly reactive molecules of superoxide (O_2^-) and hypochlorite (HOCl, the active ingredient in bleach), called the “oxidative burst” that punctures cell walls, and to a concentrated mixture of lysosomal hydrolases. Human neutrophils also use antimicrobial peptides such as the serprocidins (e.g., proteinase 3, azurocidin, and cathepsin G, a metabolic inhibitor) against fungi and bacteria.

2. **Cytolytic Enzymes** — Lysosomes, present in tears, nasal mucus and sputum, destroy the cell walls of many airborne Gram-positive bacteria by catalyzing the hydrolysis of β-1,4 linkages between N-acetyl muramic acid and N-acetyl glucosamine (peptidoglycan degradation) in bacterial cell walls—causing the bacteria to burst open, spilling their contents. Lysozyme, zymolase, glucalase and lyticase are frequently used with bacteria and yeast cells to dissolve coats, capsules, or capsids. Granulysin released by cytolytic T lymphocytes directly kills intracellular Mycobacterium tuberculosis by altering the membrane integrity of the bacillus.
3. Oligomycin inhibits the FoF1-adenosine triphosphatase (ATPase) proton pump of the mitochondrial inner membrane, and cyanide poisons mitochondrial oxidative phosphorylation.\textsuperscript{2073} Nitric oxide (conc. \textasciitilde 60 nanomolar) inhibits mitochondrial respiration,\textsuperscript{2918,2919} specifically by inhibiting cytochrome oxidase by outcompeting oxygen at the oxygen binding site;\textsuperscript{2920} NO also prevents viral replication by inactivating a crucial protease.\textsuperscript{2969} Antipsychotic drugs such as chlorpromazine, thioridazine, and fluphenazine are potent peroxisome inhibitors.\textsuperscript{3067} The nucleus of an oocyte is ejected from a cell treated with etoposide and cycloheximide (chemical enucleation),\textsuperscript{3719,3720} microtubule poisons such as colchicine, colcemid and vincristine cause extrusion of nuclei,\textsuperscript{936} and EDDF is involved in erythroid

---

### Table 10.4. Organic Compounds Affecting Eukaryotic Cell Division at Specific Phases of Mitosis

<table>
<thead>
<tr>
<th>Organic Substance</th>
<th>Cell/Tissue</th>
<th>Specific Mitotic Cycle Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Propphase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aureomycin</td>
<td>Onion root</td>
<td>Membrane dissolution delayed</td>
</tr>
<tr>
<td>Glutathione</td>
<td>Amoeba proteus</td>
<td>Prophase accelerated</td>
</tr>
<tr>
<td>Nitrophenols</td>
<td>Sea urchin egg</td>
<td>Prophase blocked</td>
</tr>
<tr>
<td>Purines</td>
<td>Arboic egg</td>
<td>Reversion to interphase</td>
</tr>
<tr>
<td>Trypan blue</td>
<td>Rabbit fibroblast</td>
<td>Spindle formation slowed</td>
</tr>
<tr>
<td>Urethan</td>
<td>Rabbit fibroblast</td>
<td>Prophase accelerated</td>
</tr>
<tr>
<td><strong>Metaphase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol, DDT</td>
<td>Onion root</td>
<td>Nucleolus neformation</td>
</tr>
<tr>
<td>Benzene</td>
<td>Mammal marrow</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Colchicine (0.00001 M)</td>
<td>Chortophaga neuroblast</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Epinephrine (0.01%)</td>
<td>Wheat root</td>
<td>Spindle rotation</td>
</tr>
<tr>
<td>Methyl naphtholhydroquinone diacetate</td>
<td>Onion root</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Morphine (0.0001 M)</td>
<td>Chick fibroblast</td>
<td>Multipolar spindle induced</td>
</tr>
<tr>
<td>Nicotine (0.0001 M)</td>
<td>Rabbit fibroblast</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Nicotine</td>
<td>Sea urchin egg</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Phenylyurethan</td>
<td>Echinorachnius egg</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Podophyllotoxin (10^{-7} M)</td>
<td>Paracentrotus egg</td>
<td>Reversion to interphase</td>
</tr>
<tr>
<td>Sulfanilamide</td>
<td>Rabbit fibroblast</td>
<td>Mitotic poison—blocking agent</td>
</tr>
<tr>
<td>Testosteron, Estrone</td>
<td></td>
<td>Abnormal chromosome orientation</td>
</tr>
<tr>
<td><strong>Anaphase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caffeine</td>
<td>Onion root</td>
<td>Incomplete chromosome separation</td>
</tr>
<tr>
<td>Rianodine</td>
<td>Sand dollar egg</td>
<td>Incomplete chromosome separation</td>
</tr>
<tr>
<td>Trypanflavine</td>
<td>Rabbit fibroblast</td>
<td>Incomplete chromosome separation</td>
</tr>
<tr>
<td><strong>Telophase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aureomycin</td>
<td>Chicken fibroblast</td>
<td>Cytoplasmic division suppressed</td>
</tr>
<tr>
<td>Carbamate</td>
<td>Sea urchin egg</td>
<td>Cytoplasmic division suppressed</td>
</tr>
<tr>
<td>Chloroacetophenone</td>
<td>Pea seedling</td>
<td>Nuclear reconstruction retarded</td>
</tr>
<tr>
<td>Nicotine</td>
<td>Sea urchin egg</td>
<td>Spindle remnant persists</td>
</tr>
<tr>
<td>Rotenone</td>
<td>Chicken fibroblast</td>
<td>Cytoplasmic division suppressed</td>
</tr>
<tr>
<td>Thiourea</td>
<td></td>
<td>Nuclear reconstruction retarded</td>
</tr>
<tr>
<td><strong>Interphase:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acridines</td>
<td>Chicken fibroblast</td>
<td>Initiation of prophase inhibited</td>
</tr>
<tr>
<td>Azaguanine</td>
<td>Mouse tumors</td>
<td>Initiation of prophase inhibited</td>
</tr>
<tr>
<td>Dyes</td>
<td>Frog sperm</td>
<td>Initiation of prophase inhibited</td>
</tr>
<tr>
<td>Hypoxanthine</td>
<td>Chicken osteoblast</td>
<td>Interphase shortened</td>
</tr>
<tr>
<td>Glucose</td>
<td>Mouse epidermis</td>
<td>Division stimulated</td>
</tr>
<tr>
<td>Nitrogen mustard</td>
<td>Rat corneal epithelium</td>
<td>Initiation of prophase inhibited</td>
</tr>
<tr>
<td>Trypanflavine</td>
<td>Onion root</td>
<td>Destruction of interphase nucleus</td>
</tr>
<tr>
<td><strong>Multiple Phases:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aminobenzoate, Coumarin</td>
<td>Onion root</td>
<td>Chromosome breaks</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Protozoan</td>
<td>Chromosome reduction induced</td>
</tr>
<tr>
<td>Dyes, N-butyl gallate</td>
<td>Onion root</td>
<td>Chromosome adhesion</td>
</tr>
<tr>
<td>Ethoxycaffeine</td>
<td>Onion root</td>
<td>Chromosome rearrangements</td>
</tr>
<tr>
<td>Mustards</td>
<td>Spidervort mother cell</td>
<td>Centromere misdivision</td>
</tr>
<tr>
<td>Urea</td>
<td>Fruit fly salivary gland</td>
<td>Chromosome dispersion/despiralization</td>
</tr>
</tbody>
</table>

---

3. Organelle Poisons — Certain drugs and other substances, if injected into the cytoplasm, may interfere with the workings of specific organelles or cell subsystems. For example, the fungal metabolic drug brefeldin A disrupts the Golgi, causing it to collapse back into the ER.\textsuperscript{2347,3441,3442} Colchicine,\textsuperscript{359} vinblastine and vincristine\textsuperscript{936} bind to tubulin, causing microtubules to disassemble, griseofulvin (an antifungal agent) prevents microtubule assembly,\textsuperscript{996} and taxol inhibits microtubule depolymerization, while cytochalasin B inhibits the polymerization of cytoskeletal actin microfilaments.\textsuperscript{939} Adociasulfate-2 is a kinesin motor inhibitor.\textsuperscript{2390} High concentrations of vitamin A weaken the lysosomal membrane. Proteasome inhibitors such as vinyl sulphone\textsuperscript{2301} are usually lethal for the eukaryotic cell.
Table 10.5. Inhibitors of Protein or RNA Synthesis

<table>
<thead>
<tr>
<th>Biochemical Inhibitor</th>
<th>Specific Molecular Inhibitory Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acting Only on Prokaryotes:</strong></td>
<td></td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>Blocks the peptidyl transferase reaction on ribosomes (MIC ~8-14 µg/cm³)</td>
</tr>
<tr>
<td>Erythromycin</td>
<td>Blocks the translocation reaction on ribosomes (MIC ~1-2 µg/cm³)</td>
</tr>
<tr>
<td>Rifampicin</td>
<td>Blocks initiation of RNA chains by binding to RNA polymerase (prevents RNA synthesis) (MIC ~1 µg/cm³)</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>Prevents the transition from the initiation complex to chain-elongating ribosome and causes miscoding (MIC ~2-10 µg/cm³; 800-4000 molecules/cm³)</td>
</tr>
<tr>
<td>Tetracycline</td>
<td>Blocks binding of aminoacyl-tRNA to A-site of ribosome (MIC ~2-4 µg/cm³)</td>
</tr>
<tr>
<td><strong>Acting on Prokaryotes and Eukaryotes:</strong></td>
<td></td>
</tr>
<tr>
<td>Actinomycin D</td>
<td>Binds to DNA and blocks the movement of RNA polymerase (prevents RNA synthesis)</td>
</tr>
<tr>
<td>Pyromycin</td>
<td>Causes the premature release of nascent polypeptide chains by its addition to growing end chain</td>
</tr>
<tr>
<td><strong>Acting Only on Eukaryotes:</strong></td>
<td></td>
</tr>
<tr>
<td>α-Amanitin</td>
<td>Blocks mRNA synthesis by binding preferentially to RNA polymerase II</td>
</tr>
<tr>
<td>Anisomycin</td>
<td>Blocks the peptidyl transferase reaction on ribosomes</td>
</tr>
<tr>
<td>Cycloheximide</td>
<td>Blocks the translocation reaction on ribosomes</td>
</tr>
</tbody>
</table>

* The ribosomes of eukaryotic mitochondria (and chloroplasts) often resemble those of prokaryotes in their sensitivity to inhibitors MIC = Minimum Inhibitory Concentration

cell denucleation, is at least one report of nuclear extrusion in lymphocytes. Selected organelle populations (e.g., mitochondria) could also be ubiquitinated.

4. Antibiotics — Drugs that prevent bacteria from multiplying (bacteriostatic) or that kill bacteria outright (bactericidal) are called antibiotics; in 1998, ~100 such drugs were FDA approved for U.S. use. Major groups include the aminoglycosides, cephalosporins, macrolides, penicillins, polypeptides, quinolones, sulfonamides, and tetracyclines; many are naturally-derived products. All act by interfering with protein synthesis, cell wall construction, or DNA replication. For example, gonococci isolated in the preantibiotic (pre-resistance) era were inhibited by benzylpenicillin (C₁₀H₁₈N₂O₄S, MW = 334 daltons) in concentrations as low as ~7 µg/cm³; in the 1990s, some highly resistant bacterial isolates required ~1000-fold higher concentrations. Vancomycin derivatives are active against Gram-positive bacteria at 10-100 µg/cm³. Antibiotics are available with protein/RNA synthesis inhibition activity against either prokaryotic or eukaryotic cells (Table 10.5).

5. Bactericidal Phages — Bacteriophages (viruses that infect bacteria) are capable of penetrating bacterial membranes and delivering foreign DNA which can take control of all metabolic processes; in ~700 sec after penetration, the first complete virion bursts, necrotically liberating ~200 virus particles. These viruses are often very host-specific. Still, it may be possible to design artificial general-purpose bacteriophages that are capable of disabling or destroying all bacterial DNA, or are capable of replicating more bacteriophage particles to which only the targeted bacterial cells are susceptible (see also Section 10.4.1.1). Engineered macrophages or artificial neutrophils might also be deployed to achieve targeted bacterial digestion. Reaching intracellular parasites such as mycoplasmas and rickettsias might be one of the biggest challenges for such biorobots.

6. Bacteriocins — Probably 99% of all bacteria generate at least one bacteriocin, small proteins that function as narrow-spectrum antibiotics that may have developed as poisons to kill competing bacteria. In 1998, ~80 bacteriocins were known, most of them produced by fermentation microbes, including nisin which is used in 45 countries including the U.S. as a commercial food additive for pasteurized egg products. Bacteriocins seem to work by entering the outer membrane of a susceptible bacterium, congregating in groups, and forming pores that allow the unregulated outflow of essential ions. The target bacterium begins breaking down ATP in a vain attempt to produce enough new protons to recharge the membrane, a futile cycle that quickly results in exhaustion of bacterial ATP. Bacteriocins are most effective in acid and least effective in salty environments. Bacteriocins of Gram-positive bacteria such as *Listeria* and *Clostridium botulinum* are ineffective against Gram-negative bacteria such as *E. coli* and *Salmonella* which have a protective double-walled outer membrane. In 1998, several Gram-negative bacteriocins were known. One example is the colicins, water-soluble cytotoxins secreted by and active against *E. coli*, that form voltage-sensitive ion channels in the bacterial inner membrane that kill the cell by selectively siphoning out key cell nutrients and inhibit protein synthesis.

7. Porins and Superporins — C. Sublette notes that it should be possible to design and insert into a cell a piece of RNA or DNA designed to produce high levels of a “superporin” protein that migrates to the cell membrane, then self-assembles into a large open-ended channel in the membrane, allowing faster and more efficient transport of ions, sugars, and other molecules across the cell membrane.
channel in the membrane several orders of magnitude larger than an ion channel, allowing a large, lethal ion influx or outflux. For example, normal cytosolic Ca\(^{2+}\) concentration ranges from 60-3000 ions/micron\(^3\),\(^{3,531}\) Blood plasma concentration is ~10\(^6\) ions/micron\(^3\) (Appendix B), but even cytosolic concentrations as low as ~10\(^4\) ions/micron\(^3\) may be toxic (Section 7.4.5.3). (See also Section 10.4.2.1.) Possible templates for such subunits are the bacterial porins,\(^{2133}\) minus their charged amino acid residues lining the inner passage which produce nonspecific aqueous diffusion channels across the outermost LPS bacterial membrane, and granulysin,\(^{2165}\)\(^{2166}\) perforin and members of the amoebapore family\(^{2166,2167}\) which are thought to damage target cell membranes by inducing formation of microscopic pores.

8. Ionophores — Ionophores are small hydrophobic molecules that dissolve in lipid bilayers, thus increasing ion permeability of cell membranes.\(^{531}\) Most are synthesized by microorganisms, presumably as biological weapons to weaken their competitors. There are two classes of ionophores—(highly temperature-sensitive) mobile ion carriers, and channel formers—both operating by shielding the charge of the transported ion so that it can penetrate the hydrophobic interior of the lipid bilayer.\(^{531}\) Ionophores permit net movement only down their electrochemical gradients, since no energy sources are available.) Valinomycin, a ring-shaped polymer that increases K\(^+\) permeability of membranes, is an example of a mobile ion carrier. Another example is A23187, which acts as an ion-exchange shuttle, carrying two H\(^+\) out of the cell for every divalent cation (such as Ca\(^{2+}\) and Mg\(^{2+}\)) carried in. Transport rates through a mobile carrier are ~2 x 10\(^{11}\) ions/sec.\(^{531}\) A channel-forming ionophore is gramicidin (C\(_{16}\)H\(_{31}\)N\(_{3}\)O\(_{26}\), MW = 2822 daltons), a 15-residue linear polypeptide with all hydrophobic side chains. Two such molecules come together in the bilayer to form a transmembrane channel that selectively allows monovalent cations (most readily H\(^+\), K\(^+\) somewhat less readily, Na\(^+\) still less readily) to flow down their electrochemical gradients at a rate of ~2 x 10\(^5\) cations/sec.\(^{531}\) The dimers are unstable, constantly forming and dissociating, with average channel open time ~1 sec.\(^{531}\) Gramicidin is produced by Bacillus brevis and is active against Gram-positive cocci and bacilli. A 5 microgram dose kills, in vitro, 10\(^3\) pneumococci or group A streptococci in 2 hours at 30\(^\circ\)C,\(^{571}\) giving a fatal dose of ~1000 molecules/micron\(^3\) assuming a ~1 cm\(^3\) culture volume in this experiment.

9. Channel Blockers — Cells import and export nutrients, ions, water, and wastes through variety of gated channels (Section 3.3.3) and transporter molecular pumps (Section 3.4.1). If these channels or pumps are permanently blocked, cytocyte may ensue. Blocking the ~10\(^5\)/micron\(^2\) cellular transport systems embedded in a typical eukaryotic cellular membrane requires ~10\(^5\)/micron\(^2\) “steric plugs” that permanently jam in the throat of critical ion channels or pumps. For example, the acetylcholine receptor channel narrows to ~0.65 nm at its waist, so an efficiently designed steric plug could have a volume of ~1 nm\(^3\) or ~10\(^{-21}\) gm. (By comparison, a single molecule of 318 dalton tetrodotoxin, or puffer fish toxin, has mass ~0.53 x 10\(^{-21}\) gm; apparently, saxitoxin (shellfish toxin, human IV lethal dose ~68 micrograms) and palytoxin are slightly smaller, whereas botulin toxin is a fairly high-mass zinc metalloprotein that inhibits release of acetylcholine at the neuromuscular junction.) Plugging all channels in the surface of a (2 micron\(^3\) bacterial cell (~10 micro\(^3\)) or a (20 micron\(^3\)) tissue cell (~1000 micro\(^3\)) requires ~10\(^4\) or ~10\(^5\) plugs, respectively, a nanorobot-dispersed ~10\(^5\) micron\(^3\) or ~10\(^3\) micron\(^3\) dosage of plugs at the target cell, assuming ~100% efficiency. A mean migration distance per plug of ~1 micron gives a Brownian diffusion time of ~1 millisecc, so a micron-scale nanorobot with 0.1 micron\(^3\) of onboard storage carries sufficient plug-doses to incapacitate ~10\(^5\) bacterial cells in ~10 sec or ~10\(^2\) tissue cells in ~0.1 sec, ignoring cell-to-cell travel time. Assuming a mean interbacterial separation of ~100 micro\(^3\) (an implied pathogen number density of ~10\(^7\)/cm\(^3\) in the tissue) and a maximum nanodevice intercellular travel speed of ~100 micro\(^3\)/sec (Section 9.4.4.2), a nanorobot spends ~1 sec in transit between neighboring bacterial cells, ~0.002 sec in cell-type recognition (Section 8.5.2.2), and ~0.001 sec killing the bacterium once the nanorobot arrives.

10. Channel Destroyers — Some smaller number (perhaps as few as one) of an enzyme-like toxin molecule capable of destroying either the steric or the electrochemical specificity of each ion channel by permanently altering the physical structure of that channel. For example, in the case of the acetylcholine receptor channel, an artificial enzyme possibly could be designed to alter the ring of negatively charged residues at the entrance to the receptor, allowing ions to enter the pore and ultimately depolarizing the cell. Enzymes typically operate at ~1000 Hz and the diffusion time across ~20 nm between adjacent ion channels is ~0.5 microsec, so a single enzyme neutralizes all 10\(^3\) channels in a bacterial cell in ~10 sec and ~1000 artificial enzymatic molecules would incapacitate the cell in ~10 millisecc.

11. Complement — The complement system is a group of ~20 soluble serum-resident proteins acting in combination with specific cytolsin antibodies in amplification cascades along two separate pathways to initiate reactions to foreign antigens (Chapter 15). Complement proteins are enzymes that act on foreign cells by punching holes in their membranes by inserting lipid-soluble pores (the membrane attack complex, C5-C9), causing necrotic osmotic cytolsis. In addition, complement plus antibody designates cells to be engulfed by phagocytic cells.

12. Animal Venoms and Toxins — The bufagin from the Bufo arenarum toad of Argentina (arenobufagin, C\(_{31}\)H\(_{33}\)O\(_{6}\), MW = 430 daltons) has a mean lethal toxicity of 92 microgram/kg,\(^{585}\) the equivalent of ~10\(^5\) bufagin molecules per 20-micron tissue cell (~100 molecules/micron\(^3\)). Most reptile venoms have lethal toxicities of 200-1000 micrograms per kg of body weight.\(^{585}\) Sea anemone granulotoxin (4958 daltons) has a mouse LD50 of 400 micrograms/kg (~49 molecules/micron\(^3\));\(^{2458}\) another sea anemone protein extract has a mouse LD50 of 40 micrograms/kg.\(^{2459}\) Maurotoxin, a 34-residue scorpion venom (Scorpio mauro), has an intracebroventricular mouse LD50 of ~80 nanograms,\(^{2460}\) or ~0.4 molecules/micron\(^3\). The LD50 value to mice of Vespa luctuosa hornet venom is 1600 micrograms/kg, the most lethal known wasp venom.\(^{2457}\) Venoms and toxins may not be cytolicidal against all cell types—some act by interfering with nerve transmission or by paralyzing muscles.

13. Plant and Microbial Toxins — A variety of microbes manufacture neurotoxins or neurolysins that destroy the ability of human ganglion and cortical neural cells to function.\(^{1236}\) A single molecule of some neurotoxins can incapacitate a cell, though the rate of action tends to be very slow. For example, botulin toxin is the second deadliest poison known (e.g., median human lethality ~5-50 nanograms/kg).\(^{3462}\) The toxin acts at cholinergic nerve terminals to prevent acetylcholine release, a permanent physiologic denervation that causes muscle paralysis, often resulting in death by respiratory system failure in 2-8 days. Recovery requires the sprouting of new axon twigs and the formation of new myoneural junctions,\(^{2122}\) which is
why the recovery time for sub-lethal botulin poisoning is very long, several months to a year or more. Botulin is a -150,000 dalton protein comprised of 1285 amino acid,1716,3462 which implies a lethal concentration of 0.2-2 molecules per 20-micron tissue cell. Other toxins are capable of killing a cell outright with just a few molecules, such as ricin: mouse and rabbit IV LD50 ~400-4000 nanograms/kg or 30-300 molecules per 20-micron tissue cell,3463-3465 possibly due to lower specificity in binding to cells and other pharmacokinetic effects. Ricin is a phytotoxic enzymatic 65,750-dalton protein3466 taken from the seeds of the castor oil plant that cleaves the ribosome complexes, shutting down protein manufacturing. This damage is normally irreversible, killing the organism in ~4 days at minimum dosage and ~8 hours at the highest dosages, so a reasonable lower limit for killing an individual cell is probably ~10^4 sec.

14. Antifungals — Antifungal agents such as amphotericin B or fluconazole are fungicastic. For example, the growth of Candida and many other fungi is inhibited by amphotericin B (C_{47}H_{73}NO_{17}, MW = 924 daltons) concentrations of 30-1000 micrograms/cm^3 in vitro.1119 A number of natural antimicrobial peptides are induced in epithelial cells at sites of inflammation; for example, lingual antimicrobial peptide (LAP), a member of the b-defense class, has been isolated from bovine tongue and exhibits both antifungal (e.g., Candida) and antibacterial (e.g., E. coli) activity1229 at concentrations as low as -16 microgram/cm^3. In 1998, antifungals were known to target membrane function, metabolism, cell wall synthesis, protein and ergosterol synthesis, nuclear division, and nucleic acid synthesis and function.2133 In those circumstances where fungal cells have learned to communicate to effect mating or cooperative behaviors, such communication may be disrupted to their detriment. If the genetic pathways involved in these behaviors include primitive forms of apoptosis, these could be exploited as well.

15. Antivirals — Antiviral drugs may work by interfering with viral replication processes, including cell attachment, cell uptake, viral coat removal, and viral DNA or RNA replication by the cell, and are primarily virostatic. For example, acyclovir (MW = 225 daltons) in concentrations of 0.01-13.5 micrograms/cm^3 inhibits by 50% the growth of herpes simplex virus in vitro.2119 In 1998, 20 antiviral were FDA approved for U.S. use. Antibodies may also serve a virostatic function. For example, the rabies-like vesicular stomatitis virus (VSV) rhabdovirus has a surface envelope with ~1200 identical glycoprotein molecules that form a regular and densely ordered pattern of spike tips.2139 Rhabdoviruses are neutralized if they cannot dock with their cellular receptors; this requires a minimum of 200-500 IgG antibody molecules bound per virion.2140 Another challenge for medical nanorobots will be the retroviruses, which insert their genome into the host DNA, thus will require chromosomal editing or replacement (Chapter 20) to remove.

16. Iodine-Based Microbicides (iodophors) — In tincture of iodine, all of the iodine is in the free form at ~10% by weight and is readily available for instantaneous reaction and killing of bacteria in a few seconds.360 The minimum lethal iodine dose for a 2-micron bacterium (assuming complete exhaustion of a 0.1-micron perimicrobial iodophor layer) may be of order ~0.001 picomoles or ~0.1 picogram iodine (~10^7 molecules), which is ~3% of the mass of a bacterium—not terribly efficient, as expected for a broad-spectrum agent. Iodoacetate inactivates most cytoplasmic enzymes and blocks anaerobic metabolism, or glycolysis (fluoride is also a glycolytic poison).758 Yodoxin (64% organically bound iodine) is amoebicidal.2119 The broadest-known spectrum iodine-based microbicide is betadine (povidone-iodine) 10% solution,2119 employed in surgical scrub and many other general antiseptic uses. Betadine kills most bacteria in 15-30 sec, and also kills viruses, fungi, yeasts and protozoa; no bacterial resistance has been reported.

17. Silver — Silver foil has been used in bacteriocidal wound dressings,2159 and silver metal has been employed for centuries to purify water. In 1998, there were more than 600 silver-based antibacterial products available in Japan including silver-impregnated pens, floppy disks, calculators, ATM machines, floor tiles, plastic food wrap, socks, shirts, public park sand, and toilet seats.2126 At least 20 companies in the U.S. offer silver/copper-based systems for swimming pool sanitation and commercial air conditioning cooling towers. The passive dissociation of silver from the metallic phase into a wound with antimicrobial results2123 and the antiseptic action of silver compounds2124 are well known. A quantitative study of the metal’s antimicrobial properties in vitro found that a silver ion concentration of ~25 micrograms/cm^3 (~10^5 ions/micron^3) reduced the number density of Staphylococcus and Pseudomonas bacterial cells and Candida fungal cells by a factor of 1-10 million relative to microbial control aliquots.2125

18. Platinum, Bismuth, and Other Elements — Platinum-based cisplatin (PtCl_2H_6N_2, MW = 300 daltons) is an antitumor and bacteriostatic drug that interacts with tumor-cell DNA by forming a Pt-GG intrastrand crosslink as the critical lesion leading to cytotoxicity.2136 A typical ~100 mg whole-body dose2119 produces a mean concentration of ~2000 molecules/micron^3. Colloidal bismuth subcitrate exerts a direct antimicrobial effect against H. pylori (gastritis) at concentrations of 4-25 micrograms/cm^3 (6000-38,000 molecules/micron^3), with treated bacteria showing deposits of bismuth on their surface and internally.2137 Mercury (chronic toxicity in blood ~600 atoms/micron^3), lead (toxic in blood at ~2900 atoms/micron^3), and arsenic (toxic in blood at ~4800 atoms/micron^3) are general poisons (Appendix B) that have been used for medicinal purposes. Indirect poisons such as iron and copper catalyze free-radical formation. Interestingly, bacterial growth is constrained by essential nutrients, and human immune cells scavenge and sequester iron when faced with a bacterial threat; nanorobotic “deferritization” of a bacterium could kill the organism.

19. Sugar — Medical lore has it that sugar, and especially honey2158,2334 (Section 1.2.1.2), makes an excellent antibacterial disinfectant. Blood glucose levels of 0.01-0.07 M in untreated diabetics cause tissue and cellular damage, at least partly via osmotic cellular dehydration (honey and NaCl salt have a similar effect), and protein glycosylation and Maillard reaction (Section 6.3.4.4) which diminish enzyme function; ~15% cytosolic glucose concentrations are toxic.

20. Hydrogen-Bond and Disulfide-Bond Breaking Agents — At high concentrations, guanidine hydrochloride disintegrates the bacterial S-layer coat which is held together only by noncovalent forces.525 Local heating (hyperthermia; Section 10.4.2.3) also disrupts H-bonds. Other agents are available to break disulfide bonds, disrupting many classes of proteins.

21. Nuclear Alkalination and Other Nucleic Disruption — RNA breaks down in slightly alkaline environments (e.g., pH ~8.0,1593 to increase the pH of a typical cellular nucleus from 7.2 to 8.0 would require the injection of ~10^5 OH- ions. RNA is also broken down by RNases; DNA is broken down by DNases, or by using restriction enzymes.
22. **Tissue Liquefaction** — A 2-hour exposure to 0.25% trypsin enzyme solution digests and breaks up the extracellular matrix, turning an organized tissue into a jumble of individual cells.

23. **Other Poisons** — Cyanide causes cells to swell and lyse at concentrations of −0.002 M (−1 × 10⁶ ions/micron³). Assuming serum reference levels, ethanol should be cytotoxic at −0.09 M (−50 × 10⁶ molecules/micron³). Ethylene oxide (C₂H₄O) gas is a microbial sterilant and fungicid; chlorine dioxide (chlorite ion) is an antimicrobial sometimes found in swimming pools and hot tubs. Most chemotherapy agents also kill cells. Common lab detergents such as Triton X-100 and tributyl phosphate readily destroy bacteria and the majority of viruses sheathed in protein envelopes.

Nanorobot biocidal delivery vehicles can be used to present the chemical agent on a cell-by-cell basis. For example, a major whole-body infection involving −10¹² pathogens (e.g., −10⁶ bacteria/cm³) with each microbe capable of being killed by −10⁵ precisely-delivered biocidal molecules each of molecular weight −10⁴ daltons requires a minimum treatment dose of −10¹⁷ biocidal molecules or −0.1 mm³ of material. This dose could be carried and dispensed in −1000 sec by one billion −1 micron³ “pharmacyte” nanorobots (total volume of therapeutic nanorobots −0.001 cm³) assuming each nanorobot has −0.1 micron³ of internal storage for the biocidal agent and assuming as before a mean −1 sec cell-to-cell transit time.

### 10.4.2 Mechanical Cytocide and Virucide

Mechanical cytocide involves the killing of biological cells or the destruction of virion particles partially or primarily by mechanical, rather than by strictly biochemical, means. Such means may include disassembly, partial physical destruction of critical components, or even complete disassembly of the target cell, bacterium, or viral particle.

#### 10.4.2.1 Transmembrane Siphonation and Ionic Disequilibrium

A number of rigid diamondoid tubes with membrane-locking lipophilic external coatings and hydrophilic end rings may be inserted through the target cell membrane, preventing self-sealing of the breach (Section 9.4.5.6) and resulting in drainage of internal fluids and disruption of ionic balances, probably triggering apoptosis in eukaryotic cells after −10⁴ sec.

What magnitude of siphonation may be required to achieve cell lethality? A cytotoxic concentration of ≅10⁴ Ca++ ions/micron³ is believed to be toxic (Section 7.4.5.3) and there are −10⁶ Ca++ ions/micron³ present in blood plasma (Appendix B) or interstitial fluid, so replacing −1% of cell volume (e.g., −0.1 micron³ for a bacterium, −80 micron³ for a tissue cell) with raw extracellular fluid should be lethal. Macrophages can ingest up to −25% of their volume per hour,⁵⁹⁰ or −10% of their volume in −1000 sec, but the paramicium, a large protozoan, excretes through its contractive vacuole a quantity of water equal to 100% of its volume in 15-20 minutes,⁷₅₈ or −10% in −100 sec.

Conservatively assuming that a −10% target cell volume influx of extracellular fluid in 1 sec will cause serious cell damage or induce apoptosis, the size and number of passive transmembrane siphons may be crudely estimated. Mean interstitial fluid pressure is −0.0002 atm (Section 8.4.2). Thus from Eqn. 9.25, the total water flow rate through water-bearing transmembrane tubes of length l_olute = 300 nm with pressure differential Δp = 0.0002 atm is V_{HP} = 4 × 10¹⁰ N tube r_tube⁻⁴ (m³/sec) for N tube tubes each of inside radius r_tube. To exchange 10% of the volume of a −2 micron diameter bacterium by siphonation with extracellular fluids in −1 sec (V_{HP} = 1 micron³/sec) requires one tube of radius r_tube = 70 nm. For a −20 micron tissue cell, V_{HP} = 1000 micron³/sec which requires one tube of radius r_tube = 400 nm or 15 tubes of radius r_tube = 200 nm.

It is important that passive siphon tubes should be employed in minimum numbers and never be left unattended. They should be tethered to the nanorobot during use and retrieved after each use. Left unattended in the plasma membrane of a dying cell, diamondoid siphon tubes could become serious systemic poisons. For example, macrophages phagocytosing apoptotic siphon-studded cells would be unable to digest the tubes and might themselves become accidentally siphoned, or might pass the lethal tubes to other innocent cells in the liver, spleen, or elsewhere.

Death by siphonation may be hastened if the nanorobot actively pumps extracellular fluids (or cations directly) into the target cytosol. A pressure differential of −1 atm applied along a simple bulk fluid injector with r_tube = 50 nm and l_tube = 300 nm gives V_{HP} = 1000 micron³/sec and draws −76 pW of power during the transfer.

Artificial cytocidal cation inflows must surpass the net natural cell-outpumping rates of those cations. For instance, there are at least five known Ca++ concentration-maintenance mechanisms in eukaryotic cells:

1. a cell membrane ATP-driven Ca++ exporting pump,
2. a cell membrane 3Na+/Ca++ exchange transporter,
3. a Ca++ transporter into the mitochondria,
4. a 2Na+/Ca++ exchange transporter into the mitochondria, and
5. an ATP-driven Ca++ pump into the endoplasmic reticulum.

What is the maximum natural outflow rate? Normal cytosolic Ca++ concentration ranges from 60 ions/micron³ for a resting cell up to 3000 ions/micron³ for an activated cell (Section 7.4.5.3). From Eqn. 3.4, the diffusion-limited ion current through the surface of a 20-micron tissue cell is −2 × 10⁷ ions/sec for a resting cell and −10⁹ ions/sec for an activated cell. Taking the extracellular concentration of −10⁶ Ca++ ions/micron³ gives a minimum lethal extracellular bulk fluid inflow rate of −1000 micron³/sec to defeat the maximum possible pumping rate of an activated 20-micron cell, consistent with the previous estimate. (The diffusion limit for a nanorobot pumping Ca++ from the extracellular fluid into the cytosol is −3 x 10¹⁰ ions/sec, far higher than the maximum rate at which the cell can outpump.) Similarly, Na’/K’ transporters present in the eukaryotic membrane at 10⁰⁰/micron⁴, each device transporting 500 ions/sec (Section 3.4.1), move at most −10⁹ ions/sec through the −2400 micron² (Table 8-17) plasma membrane. The Na’/K’ pumping cost is −16 zJ/ion (Section 3.4.1); devoting the entire 30 pW power budget of the typical cell (Section 6.5.1) to such pumping would allow the transport of −2 x 10⁹ cations. Thus, >1000 molecular sorting rotors each operating at −10⁶ ions/rotor-sec (Section 3.4.2)—the entire array transporting >10⁹ ion/sec—should allow an attacking nanorobot to defeat a cell’s attempts to resist ionic disequilibrium.

#### 10.4.2.2 Mechanical Cytoskeletolysis and Monkeywrenching

It may have occurred to the alert reader that a simple way to kill a eukaryotic cell might be to cytopenetrate and then motor around inside the cytoplasm at higher-than-recommended velocities, thus indiscriminately bursting lysosomes, shredding the thin membranes of the endoplasmic reticulum (ER) and the Golgi, and trashing other...
delicate cellular structures, depending upon the motility mechanism employed (Section 9.4). Such a course would almost certainly result in an unwanted necrotic cellular lysis. Rupture of lysosomes (Section 8.5.3.8) should especially be avoided because lysosomal lipases and other corrosive enzymes would be released into the cytosol, possibly provoking some autolysis, local disintegration of the plasma membrane, and inflammation. Peroxisome (Section 8.5.3.9) rupture may produce similar effects, since these organelles catalyze some of the fatty acid breakdown in the cell. Selective destruction of ribosomes (Section 8.5.3.4), the ER (Section 8.5.3.5), the Golgi complex (Section 8.5.3.6), the mitochondria (Section 8.5.3.10), or the nucleus (Section 8.5.4) will more slowly kill the cell but again this may be a nonapoptotic death if the transcription, protein synthesis, or energizing mechanisms necessary to sustain the apoptotic cascade are disabled. Loss of power due to mitochondrial destruction eliminates the cell’s ability to maintain long-term membrane recycling; the cell will eventually come apart due to the accumulated damage in the membrane. However, elements of the cytoskeleton probably can be safely destroyed without disabling apoptosis, while simultaneously severely reducing the chances of post-attack cell survival, particularly in cells that are incapable of further division. Chemical cytoskeletolysis by caspases (Section 10.4.1.1) and various mitotic inhibitors such as the microtubule inhibitor vincristine (Sections 9.4.7.4 and 10.4.1.3) have been described earlier. We can speculate that mechanical cytoskeletolysis may trigger apoptosis in eukaryotic cells because the entire cytoskeletal structure per se does not appear to be a crucial operational component of the apoptotic cascade.

How best to perform the cytoskeletolysis? Microtubules in cyto preferentially absorb light at ~350 nm, so in theory a sufficiently intense laser source at this frequency could selectively disrupt the microtubule network. However, proteolytic or mechanical chopping should be more energy efficient (especially if used to activate a few of the molecules at the upstream end of the cascade). For example, p56 severs microtubules slowly in an ATP-independent manner. Katanin is a heterodimeric protein that severs and disassembles microtubules in an ATP-dependent manner. Human elongation factor 1 (EF-1, ~48 kilodaltons) rapidly severs taxol-stabilized fluorescent microtubules in vitro, and induces rapid fragmentation of cytoplasmic microtubule arrays when microinjected into fibroblasts, at concentrations of ~15 microgram/cm² (~200 molecules/micron²). In theory, a pair of fiber cleavage tools modeled after EF-1α operating at ~500 Hz (Section 9.4.6) would lyse ~1000 fibers/sec. If the cytoskeleton of a typical tissue cell has a total of ~10⁶ intermediate filament segments (Table 8.17), then 10% of all such filaments can be mechanochemically cleaved by a single tool pair in ~100 sec, ignoring travel time.

The cytoskeleton can also be mechanically cleaved. Individual microfilaments have a tensile tearing strength of ~0.1 nN, microtubules ~1 nN, and intermediate filaments (IFs) ~100 nN. Consider a cutting blade of length 50 nm rotating at ~1 KHz (blade tip speed ~ 0.3 mm/sec). The energy required to cut a single 10-nm thick IF is ~0.001 pJ (Section 9.3.5.1); the power required to cut 10% of all cellular IFs in ~100 sec is ~1 pW. The continuous drag power of the freely rotating blade in plasma during the 100-sec cytoskeletalysis program is ~0.1 pW (Eg., 9.5). A 1-micron² cross-section nanorobot traversing the entire 8000 micron³ volume of a 20-micron tissue cell traces an 8000-micron path, a 100-sec travel time at a mean ~80 microns/sec velocity; continuous nanorobot drag power at this velocity even in a cytoskeleton-rich cytoplasm of viscosity ~10 kg/m-sec (Section 9.4.6) is ~1 pW (Eqn. 9.73).

Another simple technique for mechanical disruption of the cell may informally be termed “monkeywrenching.” As examples, active DNA cutters (e.g., restriction enzymes), or in situ manufactured H₂O₂ along with some iron borrowed from local ferritin molecules, could be injected into the eukaryotic nucleus, rapidly generating enough double-strand breaks to activate the fatal DNA damage response. R. Bradbury notes that a nanorobot could seize a single mitochondrion (~0.4% of nuclear volume; Table 8.17), transport it to the nuclear envelope and then forcibly insert it into the nucleus—if there are sufficient oxygen and other small reactant molecules present, the mitochondrion would probably produce enough local free radicals to initiate the DNA-damage (apoptotic) cascade. Methods for synthesizing superoxide, hydrogen peroxide, and hypochlorite from locally available materials are described in Chapter 19. In the case of prokaryotes, medical nanorobots could introduce restriction enzyme molecules that are alien to the target strain of bacterium, followed by some DNase (in case the strain has none). R. Bradbury believes that ribonucleases and peptidases may be unnecessary because some endogenous ribonucleases should always be present and the remaining lipid bag of proteins would not represent much of a threat.

10.4.2.3 Gross Cellular Disruption

A number of crude cytotic techniques have been proposed which for the most part are inadvisable. For instance, simple lancing or “harpooning” of cells (e.g., Feynman “stab the paramecium”) would likely be followed by self-sealing or a messy necrosis, depending on the duration and severity of the penetration. Acoustic shock waves would cause random mechanical damage, ending in necrosis. Bulk pressurization (see Section 10.3.3) can be lethal to biology—high pressure treatment to kill bacteria was first described in 1895 by Royer, and by Hite and colleagues in 1899 in connection with the high pressure preservation of milk. Cells are generally inactivated between 1000-5000 atm, bacteria and viruses above 5000 atm, antibodies and enzymes over 10,000-20,000 atm (Table 10.3). A pressure-initiated apoptotic response might be difficult to control because the apoptosis-inducing and necrosis-inducing pressure thresholds are not widely separated and may vary among cell types, physiological states, external conditions, and even among cells of apparently similar characteristics. Bulk pressurization followed by release would almost certainly permanently inactivate viruses (Table 10.3), but could lead to inflammation if large populations of deactivated viruses are set adrift in the tissues. Some fragmentation of DNA molecules in aqueous solution by ~1 MHz ultrasound irradiation has been reported at intensities as low as ~2000 watts/m² for >15-min exposures, but 10-min exposures to 50,000 watts/m² completely fragment the DNA, due to shearing forces acting on the large molecules rather than cavitation (which requires higher intensities). Plant cell chromosomes damage occurs at 10⁴-10⁵ watts/m², and red blood cell suspensions suffer increased membrane permeability but no membrane disruption up to 30,000 watts/m², although platelets in aqueous suspension may be disrupted by >2000 watts/m² at 1 MHz for >5 min.

Local hyperthermia can selectively destroy cancer cells in vivo, presumably with an apoptotic outcome. Temperatures of 42.5˚C or higher for 20-30 minutes appear necessary for tumoricidal effects, presumably with an apoptotic outcome. Temperatures of 42.5˚C or higher for 20-30 minutes appear necessary for tumoricidal effects, and several sessions are needed for significant tumor regression.
Focused 1 MHz ultrasound beams produce brain lesions, with in situ spatial-peak intensities at 2 megawatts/m² for 10 sec causing purely thermal damage and higher intensities up to 0.2 gigawatts/m² for 300 microsec adding a contribution from direct mechanical effect, possibly cavitation. Large numbers of nanorobots in a small tissue volume could produce significant localized or cell-wide heating for extended durations that might be difficult or impossible for individual nanorobots to achieve.

Very high intensity laser light (>10¹¹ watts/m²) can necrotically "optocute" biological cells; bioparticle optical trapping pioneer A. Ashkin notes that during his experiments, "if the power got too high, the bacteria would just explode." Threshold laser ablation rates are ~1 nJ/micron² for a 248-nm KrF excimer laser photodissociating organic material and 0.5 nJ/micron² for ablative etching of corneal tissue at 193 nm (probably as a result of photodecomposition of the peptide bonds). At 193 nm, a single 20-nanosec, 20 nJ/micron² (~10¹² watt/m²) pulse ablates 2.4 microns of bire duct tissue; hard biomaterials like tooth dentine and dental enamel ablate 0.5-1.9 microns at 248 nm and a fluence of ~10¹³ watts/m². Microlasers are an established technology, but such methods produce necrosis and require far too much power to be practical for in vivo cytoide.

However, ultraviolet UVC band photons (190-290 nm) induce cell destruction much more selectively and energy-efficiently. For example, nucleic acids have an optical absorption maximum near ~260 nm that is characteristic for each base; the absorption of DNA itself is ~40% less than would be displayed by a mixture of free nucleotides of the same composition, known as the hypochromic effect. The primary mode of UV damage is the formation of thymine dimers in the DNA, which block both transcription and replication until repaired. In sufficient numbers, these defects probably activate the DNA damage response (repair or apoptotic) pathways. At higher-energy shorter wavelengths (e.g., 193 nm), the UV photons don't penetrate the cell far enough to reach the nucleus. Lower-energy longer-wavelength UV photons can still damage DNA by producing oxidizing chromophores and decreasing enzyme synthesis, which has the effect of reducing the repair and regrowth properties of DNA; protein containing the aromatic (chromophoric) amino acids tryptophan and tyrosine have maximum absorption near 275 nm and phenylalanine more weakly near 260 nm and peptide bonds absorb strongly at wavelengths under 240 nm. Thus photons of a precise wavelength can cause very selective damage. Suspicions of Streptococcus, Lactobacilli and Actinomyces bacteria stained with toluidene blue experienced millionfold kill rates after 60-sec exposure to helium-neon laser photons at an intensity of only 5600 watts/m². Viral genetic material is also strongly susceptible to UV damage.

Particulate radiation is another poor choice for inflicting selective damage to individual cells in a controlled manner. It is true that Co⁶⁰ irradiation is used to sterilize sutures and antitumor radiation treatments have been commonplace in 20th century medicine. However, the range of natural-emission α-rays in protoplasm is >20 microns, raising the likelihood of some collateral damage, and an accelerator capable of generating artificial lower-energy alpha beams would be energetically and geometrically prohibitive in most in vivo nanomedical contexts.

10.4.2.4 Mechanical Virucide

Viruses are acellular bioactive parasites that attack virtually every form of cellular life. Viruses have diameters ranging from 16-30 nm—for example, poliomyelitis ~18 nm, yellow fever ~25 nm, influenza ~100 nm, herpes simplex and rabies ~125 nm, and psittacosis ~275 nm. Their shape is either pseudospherical with icosahedral symmetry, as in the poliomyelitis virus, or rodlike, as in the tobacco mosaic virus (TMV). Viruses consist of a core of RNA (most plant viruses and animal viruses such as the rhinoviruses, polio and flu viruses, and all retroviruses) or DNA (most bacterial and some animal viruses), but never both. This nuclear material is surrounded by a protein coat called a capsid, a quasi-symmetrical structure assembled from one or only a few protein subunits called capsomeres. A virus surrounded only by capsid is a naked virus; some viruses acquire a lipid membrane envelope from their host cell upon release, and are called enveloped viruses. Attached to the capsid are other protein structures necessary for host infection, especially attachment or docking proteins. The interior capsid volume is usually only slightly larger, and never more than twice as large, as the volume of the enclosed nuclear material. For instance, the adenovirus (one family of viruses that causes the common cold in humans) is an icosahedral particle ~70 nm in diameter (~180,000 nm³) containing one double-stranded DNA molecule (~11 microns (~35,000 bases or ~10⁷ daltons) in length (~15,000 nm), giving a composition of approximately 92% protein and 8% DNA by volume. By comparison, the TMV has 5% nucleic acid, the bushy stunt virus 17%, and the tobacco ringspot virus 40%.

There are at least two methods by which such viruses may be eliminated from the human body, as follows.

10.4.2.4.1 Sequester and Transport (ST)

Nanorobots may extend chemosensory pads (Section 4.2.8) which will selectively and reversibly adhere to the capsid coat of a nonenveloped viral target. For enveloped viruses, the nanorobot must carefully pick over the camouflage lipid membrane to find viral-specific antigens. R. Bradbury suggests that such lipid-coated viruses might also be distinguishable from host cells by measuring their membrane curvature (viruses are much smaller than eukaryotic cells) or local ion concentrations (viruses emit no metabolic effluents). Once the virus is bound to the pad, the nanorobot packs the whole virus particle into an internal storage cannister. When the cannister is full, the nanorobot ceases operations and either delivers its cargo to a dedicated biodisposal organ or other in vivo facility, or eliminates itself from the body (Chapter 16). A nanorobot with a 1.8 micron³ cannister can warehouse up to ~10,000 compacted adenovirus particles. With viral infections reaching ~10¹⁰ particles/cm³ localized in the upper respiratory mucosal tissues, mean distance between free-floating particles is of order ~5 microns, or ~100 millisecond travel time at ~100 micron/sec, probably dominating the exclusive antigenic identification time of ~2 millisecond (Section 8.5.2.2). As part of a nanomedical treatment for viral infection (Chapter 23), a minimum therapeutic dosage of ~10⁸ nanorobots/cm³ can extract and sequester the entire adenovirus population from infected human tissue in ~1000 sec (~17 minutes); larger doses can clear the tissues even faster, and may also assist in the cytoplasmic and nucleoplastic clearance of viral DNA.

10.4.2.4.2 Digest and Discharge (DD)

After virion acquisition as previously described, the particle is placed in a leakproof transfer chamber. The virus is then pistonned into a morcellation chamber (Section 9.3.5.1) where it is chopped by orthogonally traveling diamondoid blades into ~10 nm pieces, greatly increasing the reactive surface area of the virus material. The morcellate is then pistonned into a reaction chamber. Capsid-specific proteases and peptidases are introduced, reducing all virus proteins to amino acids, which are removed from solution by molecular sorting.
rotors (Section 3.4.2) and discharged. Lipases are required for digesting enveloped viruses. Deoxyribonuclease and ribonuclease enzymes are also introduced, reducing viral DNA or RNA to nucleotides which are themselves removed from the solution by a second set of molecular sorting rotors and discharged. The enzymes remaining in solution are pumped back into storage vessels via sorting rotors, and the reaction chamber is ready for the next cycle.

For example, a 70-nm adenovirus particle containing 8% DNA by weight is reduced to 1.7 x 10^{-16} gm (~10^6 molecules) of amino sorting rotors and discharged. The enzymes remaining in solution species; the largest known bacterium is 30-400 microns in diameter, and from 1-10 microns in length for the nonspherical curves. Bacteria range in size from 0.2-2 microns in width or curvature, spiral or comma-like (spirilla). Bacilli may remain metabolism, growth, and replication. Their shapes are generally mantoplas is surrounded by a usually thin lipid bilayer plasma membrane. Most Gram-positive bacteria are surrounded by an unsaturated inner lipid layer but a more rigid outer leaflet additional two-layer coat atop the peptidoglycan layer. This coat approximately linked in a single LPS molecule.2134

Its volume halved, the bacterium is packed into a nanorobot storage cannister of volume ~2 micron^3 and is delivered to an unwelcome microbe from the human body, described below. An additional serious problem in bacterial infections is that many microbes release toxins into the bloodstream that can wreak havoc even if all the bacteria are killed. Toxin cleanup (Chapter 19) is a necessary component of any comprehensive antibacterial treatment strategy.

10.4.2.5 Mechanical Bacteriocide

Bacteria are unicellular microorganisms capable of independent metabolism, growth, and replication. Their shapes are generally spherical or ovoid (cocci), cylindrical or rodlike (bacilli), and curved-rod, spiral or comma-like (spirilla). Bacilli may remain associated after cell division and form colonies configured like strings of sausages. Bacteria range in size from 0.2-2 microns in width or diameter, and from 1-10 microns in length for the nonspherical species; the largest known bacterium is *Thiomargarita namibiensis*, with spheroidal diameters from 100-750 microns.325 Spherical bacterial as small as 50-500 nm in diameter have been reported, but it has been theorized that the smallest possible cell size into which the minimum essential molecular machinery can be contained within a membrane is a diameter of ~40-50 nm.2149 Many spherical bacteria are ~1 micron in diameter; an average rod or short spiral cell is ~1 micron wide and 3-5 microns long. Each bacterial cell consists of a mass of protoplasm enclosed within the usual thin lipid bilayer plasma membrane. Most Gram-positive bacteria are surrounded by a thick, mechanically strong but porous peptidoglycan cell wall. Gram-negative bacteria like *E. coli* surround themselves with an additional two-layer coat atop the peptidoglycan layer. This coat has an unsaturated inner lipid layer but a more rigid outer leaflet composed of an unusual lipid, called lipopolysaccharide (LPS), in which fatty acid chains are all saturated and 6-7 chains are covalently linked in a single LPS molecule.2134

Bacterial cells have no internal membranous surfaces and no internal organelles, although some functional compartmentalization does exist.3616 Bacteria possess small internal vacuoles, ribosomes, and granules of stored food, and usually one or more externally-attached flagella (Section 9.4.2.5.2). Some rod-shaped bacteria can form tiny spherical or oval endospores that survive when conditions become inhospitable for metabolism, such as extreme heat (>100˚C) or desiccation (for up to 60 years).1225 As prokaryotes, bacteria have no distinct nucleus. However, a single circular chromosome is organized into one or more compact aggregates, called nucleoids, that may occupy about one-third of cell volume.397 *E. coli*, a well-studied cylindrical bacterium measuring 0.65 microns wide and 1.7 microns long (cell volume ~ 0.6 micron^3), has one double-stranded DNA chromosome ~1.3 mm (~4.2 megabases or ~10^7 daltons) in length (strand volume ~ 0.002 micron^3), organized in ~40 kilobase loops.397 Chromosomes spilled from lysed bacteria show ~80% DNA content, suggesting loop stabilization by proteins and at least a primitive membrane-associated cytoskeleton.

Table 10.6 gives the approximate mean composition of a 4 micron^3 bacterium. There are at least three methods for eliminating an unwelcome microbe from the human body, described below. An additional serious problem in bacterial infections is that many microbes release toxins into the bloodstream that can wreak havoc even if all the bacteria are killed. Toxin cleanup (Chapter 19) is a necessary component of any comprehensive antibacterial treatment strategy.

10.4.2.5.1 Desiccate, Sequester and Transport (DST)

After the target bacterium has been identified by chemosensory pads binding, the nanorobot secures itself to the cell exterior. The first objective is to desiccate the bacterium. An efficiently designed molecular sorting rotor for water may require ~30 nm^2/receptor of exposed surface with 600 nm^3/rotor which includes 50% volume overhead for housings and other mechanical elements, and transports ~10^-12 molecules/sec at a cost of ~0.01 pW/rotor (Section 3.4.2). Of total bacterial water, ~2/3 is freely diffusible as bulk water and ~1/3 is loosely bound as hydration water (Section 8.5.3.3). Thus to accomplish the desiccation, the nanorobot inserts a 0.3-micron wide, 2-micron long cylindrical water extraction probe into the cell interior. Taking the experimentally measured19140 failure strength of the wet peptidoglycan wall as ~3 x 10^6 N/m^2 (Table 9.3), a probe tip configured as a core sampler tool (Section 9.3.2) with an annular cutting edge thickness ~20 nm can be compacted into a 2-micron wide storage cylinder by a 2-micron piston by applying an Euler buckling force

The extraction probe is tiled with ~66,000 sorting rotors of volume ~0.04 micron^3. Extraction probe volume is ~0.14 micron^3, which leaves ~0.10 micron^3 for probe structure, probe plumbing manifold, pipes and pumps, and probe control mechanisms. All bulk water is extracted from the bacterium in ~1 sec, shrinking the cell by half from ~4 micron^3 to a volume of ~2 micron^3. (Sorting rotors covering ~2 micron^3 remove ~3 x 10^11 molecules/sec-micron^3, far outpacing possible backflows which may be crudely estimated from maximum macrophage water-ingestion rates (Section 10.4.2.1), suggesting at most ~10^-12 molecules/sec-micron^3 across a ~13 micron^-3 bacterial surface.) Engulf formations of metamorphic nanorobots (Section 5.3.4) could be especially useful in this application.

Its volume halved, the bacterium is picked into a nanorobot storage canister of volume ~2 micron^3 and is delivered to an appropriate dedicated biodisposal organ or other biodisposal facility (Section 10.4.2.4)—one dead bacterium per nanorobot. (A rigid wet hollow peptidoglycan sphere of diameter ~2 microns and thickness ~20 nm can be compacted into a 2-micron wide storage cylinder by a 2-micron piston by applying an Euler buckling force
(Eqn. 9.44) of ~1540 nN, or ~5 atm; a dry peptidoglycan sphere of this size would require ~10,000 atm to crush.) Serious infections of ~10^7 bacteria/cm^3 (mean separation ~50 microns) thus require a minimum therapeutic dose of 10^7 nanorobots/cm^3. Bacterial clearance time for infected tissue at minimum dosage is dominated by entry and withdrawal times (Chapter 16) for single-cell-payload nanorobots, which may require several blood circulation times, but is certainly less than 1000 sec. An alternative suggestion is that a desiccated bacterium could be "shrink-wrapped," possibly using a tightly wound surface webbing consisting of polymerized glucose (e.g., cellulose or "string"), then flagged for phagocytic removal (Section 10.4.1.2) and released, with the objective of improving the microbe-processing speed of nanorobots. Unfortunately this incautious approach allows bacterial DNA to remain intact. If the wrapping is easily digestible, the prokaryote might escape confinement prior to disposal, and reinflate; if not easily digestible, blockade of the macrophage system could result.

Note that cellular RNA synthesis and protein synthesis stop below 70%-80% hydration. All metabolism of small molecules, lipid synthesis, amino acid synthesis, and CO_2 fixation into organic molecules ceases below ~35% hydration. 941

10.4.2.5.2 Neuter and Release (NR)

After identification and docking, the nanorobot extends an adhesion antenna tool (Section 9.3.2) into the bacterial interior and sweeps it around inside. The antenna recognizes DNA and RNA material including chromosomes, plasmids, and ribosomes. Any such material found during antenna sweeps through the protoplasm adheres to reversible binding sites on the antenna. The antenna of diameter d_{ant} is then rotated N_{rot} = \frac{L_{DNA}}{d_{ant} \pi} turns which is sufficient to randomly enspool a chromosome strand of maximum length L_{DNA}; taking d_{ant} = 0.1 micron and L_{DNA} = 1.3 mm for ~4.2 Mb E. coli, N_{rot} = 4100 turns which probably may be executed in ~1 sec without strand breakage because the bonding between chromosome and cellular matrix is likely noncovalent. The mean thickness of the bolus of DNA that is tightly spooled around the antenna is ∆X = \left((V_{DNA} / \pi L_{bolus}) + (d_{ant}^2 / 4)\right)^{1/2} - (d_{ant}/2) = 12 nm, where DNA volume V_{DNA} = 4.2 x 10^8 nm^3 at ~1 nm^3/bp and L_{bolus} = 1 micron is the bolus length.

The adhesion antenna * is retracted, with patching lipids ejected from the tip as it clears the hole, thus auto-sealing the hole. The adhered genetic material is pushed into a reaction chamber (Section 10.4.2.4.2), whereupon deoxyribonuclease enzymes are introduced to reduce bacterial DNA to nucleotides which are then removed from the solution by molecular sorting rotors and discharged. Ribonucleases and peptidases are also present to reduce RNA and ribosomes to dischargeable effluent, leaving the cell with no genetic or transcription capability, effectively neutering the bacterium. After minor surface modifications to enhance immune system recognition (Section 10.4.1.2), the cell is abandoned to allow natural phagocytic processes to run their course.

* R. Bradbury suggests that a DNA-binding grappling hook mounted on a cable sliding inside a nanotube sheath could be inserted less disruptively through a bacterial cell wall. The interior DNA strands are snagged and then pulled into the nanorobot through the sheath by retracting the cable.

Table 10.6. Summary of Biochemical Composition and Net Digestion Breakdown Products of a Representative 4 µm^2 Bacterium (modified from Becker313 and Lewin997)

<table>
<thead>
<tr>
<th>Cell Component</th>
<th>Avg. Molecular Weight (daltons)</th>
<th>Cell Mass %</th>
<th>Cell Mass (gm)</th>
<th>Number of Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inorganic:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water</td>
<td>18</td>
<td>70.0%</td>
<td>3.0 x 10^-12</td>
<td>1.0 x 10^11</td>
</tr>
<tr>
<td>Salts</td>
<td>~55</td>
<td>1.0%</td>
<td>4.3 x 10^-14</td>
<td>4.7 x 10^8</td>
</tr>
<tr>
<td>Organic:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carbohydrates</td>
<td>~180</td>
<td>3.0%</td>
<td>1.3 x 10^-13</td>
<td>4.3 x 10^8</td>
</tr>
<tr>
<td>Amino Acids</td>
<td>~100</td>
<td>0.5%</td>
<td>2.1 x 10^-14</td>
<td>1.3 x 10^8</td>
</tr>
<tr>
<td>Nucleotides</td>
<td>~308</td>
<td>0.5%</td>
<td>2.1 x 10^-14</td>
<td>4.2 x 10^7</td>
</tr>
<tr>
<td>Large Molecules:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proteins</td>
<td>~30,000</td>
<td>14.5%</td>
<td>6.2 x 10^-13</td>
<td>1.2 x 10^7</td>
</tr>
<tr>
<td>Lipids</td>
<td>~700</td>
<td>2.0%</td>
<td>8.5 x 10^-14</td>
<td>7.3 x 10^7</td>
</tr>
<tr>
<td>Polysaccharides</td>
<td>~1000</td>
<td>2.0%</td>
<td>8.5 x 10^-14</td>
<td>5.1 x 10^7</td>
</tr>
<tr>
<td>RNA</td>
<td>~10^8</td>
<td>6.0%</td>
<td>2.6 x 10^-13</td>
<td>1.5 x 10^5</td>
</tr>
<tr>
<td>DNA</td>
<td>~10^9</td>
<td>0.5%</td>
<td>2.1 x 10^-14</td>
<td>~1</td>
</tr>
<tr>
<td>TOTALS:</td>
<td></td>
<td>100.0%</td>
<td>4.3 x 10^-12</td>
<td>1.01 x 10^11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Net Digestion Products</th>
<th>Digestion Product # of Molecules</th>
<th>Digestion Product Mass (gm)</th>
<th>Natural Blood Conc. (gm/cm^3)</th>
<th>Blood/Bacterium Equivalents (#/cm^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Water)</td>
<td>1.0 x 10^11</td>
<td>3.0 x 10^-12</td>
<td>~0.94</td>
<td>3.1 x 10^15</td>
</tr>
<tr>
<td>Inorganic Salts</td>
<td>4.7 x 10^8</td>
<td>4.3 x 10^-14</td>
<td>~7 x 10^-5</td>
<td>1.6 x 10^15</td>
</tr>
<tr>
<td>Fatty Acids</td>
<td>2.0 x 10^8</td>
<td>8.5 x 10^-14</td>
<td>~4 x 10^-3</td>
<td>6.7 x 10^10</td>
</tr>
<tr>
<td>Simple Sugars</td>
<td>7.1 x 10^8</td>
<td>2.1 x 10^-13</td>
<td>~2 x 10^-3</td>
<td>9.4 x 10^7</td>
</tr>
<tr>
<td>Amino Acids</td>
<td>3.9 x 10^9</td>
<td>6.4 x 10^-13</td>
<td>~5 x 10^-5</td>
<td>7.8 x 10^7</td>
</tr>
<tr>
<td>Nucleotides</td>
<td>5.8 x 10^8</td>
<td>3.0 x 10^-13</td>
<td>~1 x 10^-6</td>
<td>3.3 x 10^6</td>
</tr>
<tr>
<td>TOTALS</td>
<td>1.1 x 10^11</td>
<td>4.3 x 10^-12</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
10.4.2.5.3 Liquefy, Digest and Discharge (LDD)

After identification and docking, the nanorobot extends a retractible, self-cleaning, flexible diamondoid filament rotating “eggbeater” tool (Section 9.3.2) into the bacterial interior and sweeps the tool around inside the cell. A mechanical sensor that serves as a registration bumper detects the presence of cell wall, allowing the rapidly rotating mechanism to avoid damaging the bacterial outer perimeter. A beater with four tines of length 500 nm, tine width 50 nm, and forward cutting edge 10 nm, bowed to a 200 nm diameter and rotating at ~100 Hz has an equatorial velocity of ~50 micron/sec and can cut material having tearing strength ~10^6 N/m^2 (~maximum for soft biological materials; Table 9.3) while consuming ~30 pW (Eqn. 9.75) in continuous operation in highly viscous E. coli cytoplasm (absolute viscosity η ~ 1000 kg/m-sec; Table 9.4). As liquefaction proceeds, the protoplasm may become less viscous, allowing rotation rates to be increased and input power to be reduced.

After 1-10 sec, the appropriately liquefied material is pumped from the exemplar 4 micron^3 bacterium interior into the 1 micron^3 enzyme reaction chamber in ~0.01 micron^3 aliquots and is processed as described earlier in ~100 millisecond cycles consuming ~30 pW (Section 10.4.2.4.2). However, the LDD chamber also includes lipases to digest lipids, amylases and related enzymes to digest carbohydrates, and molecular sorting rotors able to transport and discharge fatty acids, sugars, and inorganic ions. A few specialized additional enzymes may be required to fully digest unusual or rare metabolites that might already be present or might appear during processing. Processing time for ~400 cycles is ~40 sec. Finally, the bacterial coat, plasma membrane, and flagella (another ~0.3 micron^3 plus ~0.001 micron^3 flagellum) are sectioned, drawn into the reaction chamber, processed in like fashion, and then discharged in an additional 30 chamber cycles or ~3 sec. Care should be taken not to fragment the coat until nearly all of the cell contents are evacuated, thus minimizing contamination of the extracellular environment. A few additional sorting rotors and special processing may be required for endotoxins, native cellular enzymes, digestible bacterium-resident poisons and heavy metals, and so forth. Operating simultaneously and in parallel, all 10^11 water molecules are discharged in ~50 sec by 2000 sorting rotors drawing 20 pW. Thus, total processing time is ~50 sec, consuming ~50 pW.

Table 10.6 also compares the composition of bacterial digestion products to natural bloodstream concentrations of those same products (Appendix B), and shows that the (extremely conservative) maximum “safe” bacterium material discharge rate is nucleotide-limited to ~10’ molecules/cm^3, which is near the typical microbial number density in serious infections. Digested-bacterium material discharges should not significantly augment natural serum concentrations of most of these substances.

Processing 20-micron tissue cells may proceed by similar, but more cautious, means, and will require on the order of one nanorobot-day to digest and discharge each such cell. A cooperative group of ~100 LDD nanorobots occupying ~10% of a tissue cell surface can perform a complete disassembly and discharge in ~1000 sec, employing a group power draw of ~5000 pW for the duration.

10.4.2.6 Cytocarriage Disposal

Unwanted bacterial or other motile cells may be piloted to natural disposal sites in the human body via cytocarriage (Section 9.4.7), a potentially efficient pathogen-clearing technique that is elaborated in greater detail in Volume II.

10.5 Temperature Effects on Medical Nanorobots

Will medical nanorobots retain functionality at unusually high or low operating temperatures? This is an important question for nanodevices at work in human limbs that may be subjected to hot scalding, combustion, explosion, or other burn traumas, or at the cold temperature extreme, to severe frostbite or exposure of the extremities whether in space or in arctic conditions, accidental ice burial after an avalanche, or in situations requiring the repair and resuscitation of cold-vitrified or cryogenically-preserved tissues, organs, or whole organisms.

A complete review of the many effects of environmental temperature on the operations of medical nanodevices is beyond the scope of this text. This Section can only briefly mention a few of the many design issues that may arise if nanorobot operations are contemplated significantly above or below normal human body temperatures.

10.5.1 Dimensional Stability and Strength

At high applied stress, covalent bonds cleave much more readily at higher temperatures. For example, bond a C=C double bond cleaves in ~10^-4 sec at 0 K, ~10^-3 sec at 300 K, and ~10^-2 sec at 500 K. The positional uncertainty is somewhat less problematic in design—the mean classical thermal longitudinal displacement of a diamondoid logic rod varies only as ~T/1210. The endpoint of such a rod 1 nm wide and 100 nm long displaces ~0.05 nm at 77 K (liquid N2), ~0.10 nm at 300 K, and ~0.14 nm at 600 K (liquid lead). Mechanical elements thus become more reliable at lower temperature—e.g., the probability of error in a force sensor (Section 4.4.1) scales as ~exp(1/T), so a 1% error in a sensor at 310 K increases to a 10% error at 600 K, but falls to just 10^-6 % error at 77 K.

Most materials contract when cold and expand when hot (although zirconium tungstate is a well-known exception2938). Thus a 1000-nm long diamondoid rod at 310 K contracts to ~999 nm at 77 K and expands to 1001 nm at 600 K. However, the coefficients of thermal expansion and various other thermophysical parameters are themselves temperature-dependent. The coefficients of volumetric thermal expansion at -298 K (-25°C) are 3.5 x 10^-6 K^-1 for diamond, 15.6 x 10^-6 K^-1 for sapphire, 1.2 x 10^-6 K^-1 for vitreous silica and 36 x 10^-6 K^-1 for crystalline silica or quartz.567

Young’s modulus (modulus of elasticity) and the other moduli are also temperature-sensitive, especially and most obviously near phase changes—e.g., sapphire melts at 2310 K but “softens” at 2070 K1602. At ~1773 K, diamond plastically deforms at just -1 atm pressure, but >60,000 atm are required to deform diamond at ~1300 K.2041 At the cold extreme, anyone who has seen the hammering of a nail using a piece of banana cooled in liquid N2 has witnessed the temperature sensitivity of materials strength.

10.5.2 Viscosity and Locomotion in Ice

The viscosity of liquids generally declines with rising temperature following Andrade’s formula (Section 9.4.1.1); water is ~6 times more viscous at 273 K (near freezing) than at 373 K (near boiling) (Table 9.4). Viscosity affects internal fluid transfers but also locomotion. Natrium in liquid nitrogen should require relatively less power than in water, since liquid air (at 81 K) is only one-quarter as viscous as liquid water at 310 K (Table 9.4).

An equally relevant but far more serious challenge is locomotion through solid water ice. Just below freezing, crystalline ice viscosity is
-10^{10} \text{ kg/m-s-sec}, requiring a 1-micron nanorobot to expend on the order of -200,000 pW to creep forward at 1 micron/sec (Eqn. 9.73) by viscoplastic flow in which ice crystals are deformed without breaking. Just halfway from freezing to liquid nitrogen temperature, at 164 K, viscosity has already risen to -10^{21} \text{ kg/m-s-sec}, roughly equivalent to solid mantle rock, and the power requirement has increased 100-billionfold, clearly prohibitive. (Microdroplets of pure water may be supercooled in the liquid state to 235 K at 1 atm, or 181 K at -2000 atm.\textsuperscript{2965}).

One solution that avoids this problem at temperatures near the melting point is baronatation, which depends upon the fact that water, almost uniquely, is less dense as a solid than a liquid (i.e., ice floats), suggesting a freezing point depression effect with increasing pressure that is visible as the short downleg from 0°C to -16°C in the phase diagram for Ice Ih (Figure 10.11). This is confirmed experimentally by suspending two heavy weights from a wire stretched across a block of ice. The wire passes slowly through the block, the wire exerting a pressure that melts a thin layer of water ahead of it, allowing the wire to progress; as the water passes behind the wire to the lower pressure region, it refreezes, a process known as regelation.\textsuperscript{1697} The melting ice ahead of the wire absorbs the heat of fusion while the refreezing water gives up the heat of fusion, with the heat steadily transferred by the wire, hence a good conductor cuts better than a poor one.\textsuperscript{1697} The barostatic freezing point depression constant for ice is 134 atm/°C\textsuperscript{390,2050} up to -2100 atm. By exerting a higher pressure (force per unit area) ahead of it than behind it, a baronatating nanorobot of roughly conical geometry can progress slowly through ice that is no colder than -16°C.

Taking the heat of fusion for water ice as \( \Delta H_{\text{fus}} = 306 \text{ pJ/micron}^3 \) (334 J/gm at 0°C) and assuming at least -3 micron\(^3\) of ice must be melted to allow 1 micron of forward progress, then baronatation power requirement is very conservatively estimated as \( P_{\text{baro}} = 3 \Delta H_{\text{fus}} v_{\text{nano}} \approx 900 \text{ pW for } v_{\text{nano}} = 1 \text{ micron/sec}. \) (This is energy flow, which is not necessarily energy dissipation, since most of the heat loss will come from water refreezing at the rear of the nanorobot and only the losses from finite thermal conductivity must be made up.) Below -16°C the ice would have to be heated to that temperature before melting can occur, requiring a prohibitively prohibitive additional power dissipation \( P_{\text{heat}} = 10^9 \text{ K} \text{ T}^2 \approx 22,000,000 \text{ pW} \) to produce a \( T = 10^\circ C \) warming of size \( L_{\text{nano}} = 1 \text{ micron} \) in ice assuming thermal conductivity of \( K = -2,200,000 \text{ pW/micron-°C} \) at -24°C. Bejan and Tyvand\textsuperscript{2861} have analyzed gravity-induced pressure-melting of ice due to the passage of solid bodies having square, disklike, or cylindrical contact surfaces. The constants may be estimated as:\textsuperscript{390}

\[
k_f = \frac{R T_{\text{fus}}^2 \text{MW}}{1000 \Delta H_{\text{fus}}} \quad (\circ/C/molal) \quad [\text{Eqn. 10.22}]
\]

\[
k_b = \frac{R T_{\text{fus}}^2 \text{MW}}{1000 \Delta H_{\text{vap}}} \quad (\circ/C/molal) \quad [\text{Eqn. 10.23}]
\]

where \( R = 8.31 \text{ J/mole-K} \) (universal gas constant), MW is molecular weight in gm/mole or daltons, \( T_{\text{fus}} \) and \( T_{\text{boil}} \) in K, and \( \Delta H_{\text{fus}} \) and \( \Delta H_{\text{vap}} \) are the heats of fusion and vaporization, respectively (Table 10.8).

In the case of electrolyte solutes (e.g., KCl) in dilute solutions (e.g., \( M_1 = 0.1 \text{ molal} \)), \( n \) is approximately the moles of ions per mole of solute (e.g., \( n \sim 2 \) for KCl). In more concentrated solutions, experimental values of \( n \) are slightly lower, as explained by Debye-Huckel theory due to ion-ion and ion-solvent interactions. For example, with KCl in water, \( n = 1.94 \) at \( M_1 = 0.01 \text{ molal} \) but \( n = 1.80 \) at \( M_1 = 0.50 \text{ molal} \). The most concentrated salt solution (e.g., 6.2 molal NaCl) depresses the freezing point of water by -20°C.

At still lower temperatures, solvents with much lower melting points may be employed to avoid solvent freezing. In general, substances dissolve in liquids that are chemically similar. Water (itself a polar solvent) and many organic compounds including

**An additional complication is that the uppermost molecular layers of ice may not be fully frozen. Experiments by Van Hove and Somorjai\textsuperscript{2090} show that even as cold as 90 K, the amplitude of vibrational motions in the topmost surface water monolayer is several times that of the water molecules buried deeper in the bulk ice. The second molecular layer also has enhanced vibrational motion, but far less than the top monolayer. This excess motion is attributed to a lack of water molecules above the monolayer, hence the monolayer molecules have fewer motion-restricting hydrogen bonds to other water molecules than those of the layers beneath the surface. At 200 K, the amorphous film becomes thicker; above 230 K, the film becomes a quasi-liquid layer measured as ~12 nm thick at 249 K and ~30 nm thick at 268 K, rising to ~70 nm thick at 272.5 K.\textsuperscript{2701}**
in liquid ammonia which melts at -78˚C (195 K). Sodium chloride is slightly soluble in ethanol, and also glucose are soluble in ethanol, which remains a liquid to -117˚C (156 K). Sodium chloride is slightly soluble in ethanol, and also in liquid ammonia which melts at -78˚C (195 K). Some natural enzymes are known to retain their function in liquid ammonia, and others in supercritical carbon dioxide; artificial enzyme systems based on natural peptides, or other polymers with protein-like conformational properties, could in principle operate in cryogenic solvents such as tetrafluoromethane, carbon monoxide (a polar solvent, liquid from -191.5˚C to -199˚C) or liquid nitrogen (a nonpolar solvent, liquid -195.8˚C to -209.9˚C). Lactate dehydrogenase enzyme found in cold water Antarctic fishes operates as fast as the related enzyme in animals with higher body temperatures, even though enzyme action normally halves for every 10˚C temperature drop (Section 4.6.4). The Antarctic fish enzyme compensates for the cold with modifications near the enzyme’s active site that increase flexibility and mobility, in effect “greasing the hinges so that the enzyme can move more quickly.”

| Table 10.7. Molal Freezing and Boiling Point Constants |
| Solvent | Freezing Point (˚C) | k_f (˚C/molal) | Boiling Point (˚C) | k_b (˚C/molal) |
| Acetic acid | 16.6 | 3.90 | 118.5 | 2.93 |
| Acetone | -95.4 | --- | 56.2 | 1.71 |
| Benzene | 5.5 | 5.12 | 80.1 | 2.53 |
| Camphor | 176.0 | 40.0 | 209.0 | 5.95 |
| Chloroform | -63.5 | 4.73 | 61.2 | 3.63 |
| Cyclohexane | 6.5 | 20.2 | 81.0 | 2.79 |
| Diethyl ether | -116.2 | --- | 34.6 | 2.11 |
| Ethanol | -117.3 | 1.85 | 75.8 | 1.19 |
| n-Hexane | -95.0 | 1.75 | 68.0 | 2.75 |
| Methanol | -97.8 | 2.58 | 65.0 | 0.83 |
| Naphthalene | 80.2 | 6.85 | 217.9 | --- |
| n-Octane | -56.5 | 2.14 | 125.5 | 4.02 |
| Phenol | 43.0 | 7.27 | 182.0 | 5.56 |
| Water | 0.0 | 1.86 | 100.0 | 0.512 |

| Table 10.8. Molar Heats of Fusion and Vaporization at 1 atm |
| Substance | \( \Delta H_{\text{fus}} \) (J/mole) | \( \Delta H_{\text{vap}} \) (J/mole) |
| Acetic acid | 11,700 | 24,400 |
| Aluminum | 10,500 | 230,000 |
| Ammonia | 5,660 | 23,400 |
| Argon | 1,110 | 6,540 |
| Benzene | 9,850 | 30,800 |
| n-Butane | 4,670 | 22,400 |
| Carbon monoxide | --- | 6,080 |
| Chlorine (Cl₂) | 6,410 | 20,100 |
| Chloroform | 9,220 | 29,500 |
| Diethyl ether | --- | 26,000 |
| Ethane | 2,860 | 14,700 |
| Ethanol | 5,030 | 38,600 |
| Fluorine (F₂) | 1,590 | 6,320 |
| Helium | 21 | 84 |
| n-Hexane | 13,000 | 28,900 |
| Hydrogen (H₂) | 117 | 905 |
| Iodine | 16,800 | 44,000 |
| Iron | 1,290 | 380,000 |
| Lead | 5,080 | 170,000 |
| Mercury | 2,300 | 56,600 |
| Methane | 943 | 8,190 |
| Methanol | 3,170 | 35,300 |
| Nitrogen (N₂) | 721 | 5,580 |
| n-Octane | 20,800 | 35,000 |
| Oxygen | 444 | 6,830 |
| Platinum | 22,200 | 520,000 |
| Sapphire/Ruby | 109,000 | (decomp.) |
| Silver | 9,520 | 250,000 |
| Sodium Chloride | 28,900 | 180,000 |
| Water | 6,017 | 40,690 |

\[ t_{\text{EQ}} = \frac{L^2}{C_v K_t} \] (Eqn. 10.24)

For water at 310 K, \( C_v = 4.19 \times 10^6 \text{ J/m}^3\text{K} \) and \( K_t = 0.623 \text{ W/m} \cdot \text{K} \), hence \( t_{\text{EQ}} = (6.7 \times 10^6) L^2 \) (sec) for in vivo medical nanorobot refrigerators. Thus a cold box 1 mm wide requires \( t_{\text{EQ}} \approx 7 \text{ sec} \), but a 1 micron³ box equilibrates in only ~7 microsec. Diamond is far more conductive; a 1 micron³ cold box embedded in a surplus of diamondoid structure (\( C_v = 1.8 \times 10^6 \text{ J/m}^3\text{K} \)) inside a nanorobot has a \( t_{\text{EQ}} = (900) L^2 = 0.9 \) nanosec equilibration time (giving a ~10⁻¹¹ Kelvin/sec cool rate, assuming \( \Delta T = 100 \) K). Thus in order to avoid thermal re-equilibration with the surroundings, a ~1 micron³ refrigeration mechanism must circulate working coolant fluid at a velocity \( v_{\text{fluid}} = L / t_{\text{EQ}} \approx 1000 \text{ m/sec} \), very near the speed of sound in most fluids, hence sub-ambient refrigerators smaller than ~1 micron³ are not feasible using this method.
cycles as the contained ice alternately freezes and thaws, combined with a low-viscosity, low-melting point carrier such as a light hydrocarbon. The heat absorbed by a substance that melts at constant temperature (the melting point) is the heat (or enthalpy) of fusion; the much larger heat required to boil a liquid is the heat (or enthalpy) of vaporization (Table 10.8); the heat of sublimation is simply the sum of the two.\textsuperscript{390,2036} Phase changes provide the most efficient cooling—ice absorbs \(306 \text{ pJ/micron}^3\) of heat when it melts at 0°C, and water absorbs \(2170 \text{ pJ/micron}^3\) when it boils at 100°C, but water at 37°C absorbs only \(42 \text{ pJ/micron}^3\) when it warms by 10°C. The exact temperature of a reversible phase transition in a refrigerant working fluid can be precisely controlled by judicious selection of operating pressures and fluid materials. Thermal-driven phase-change microactuators have been tested.\textsuperscript{545}

Many other possible refrigeration technologies are known but have yet to be investigated in the context of nanorobot refrigeration, including “magnetic cooling” by adiabatic demagnetization,\textsuperscript{1031} or magnetocaloric refrigerators,\textsuperscript{2159} Seebeck effect or Peltier effect electronic cooling,\textsuperscript{1034,1035,2160,2933} thermoacoustic refrigeration (Section 6.3.3), optical refrigerators,\textsuperscript{549} chemomechanical turbines operated in reverse,\textsuperscript{597} heat of solvation mediated by molecular sorting rotors (e.g., KMnO\textsubscript{4} cools solvent water by 44,000 J/mole, which is ~750 pJ/micron\textsuperscript{3} or ~73 zJ/molecule, as it dissolves,\textsuperscript{763}) heat of allotropic-transition cooling (e.g., transition from red \(\alpha\)-HgI\textsubscript{2} to yellow \(\beta\)-HgI\textsubscript{2} absorbs ~13,000 J/mole or ~180 pJ/micron\textsuperscript{3},\textsuperscript{2036}) and acoustic, polymeric, or mechanical prevention of ice crystallization during supercooling of working fluid.

10.5.5 Other Temperature-Dependent Properties

There are a great many temperature-dependent materials properties of nanomedical relevance, but there is only space here to mention just a few:

1. Denaturation and Combustion — Protein denaturation and reduced receptor-ligand fidelity may occur at temperatures as low as 50-100°C. It is true that not much happens to a block of diamond dropped into boiling water,\textsuperscript{280} but the maximum combustion point for diamond in air is ~800°C,\textsuperscript{601} and carbon nanotubes start to burn in air at ~700°C.\textsuperscript{1855} At low temperatures, receptor-ligand binding may occur with greater fidelity but more slowly, and there are various unusual biological effects that occur at low temperature—for example, the lens of the human eye becomes opaque when cooled to below freezing [G. Fahy, personal communication, 1997].

2. Speed of Sound — Acoustic waves travel at different speeds in cold and hot media, potentially affecting medical nanorobot sensing, energy transmission, communication and navigation. In general, the speed of sound \(v_{\text{sound}}\) in liquids depends upon the adiabatic bulk modulus \(B\) and density \(\rho\), both of which are temperature-dependent, as \(v_{\text{sound}} = \frac{B}{\rho}\frac{1}{2}\) (Eqn. 4.30). The temperature dependence of the speed of sound in pure water at 1 atm has been carefully studied and is approximated fairly well by:\textsuperscript{2155}

\[
v_{\text{sound}} = 1557 - (0.0245)(347 - T)^2\quad \text{[Eqn. 10.25]}
\]

where \(T\) is temperature in K. Thus the speed of sound increases with rising temperature up to a peak at 347 K (74°C), then decreases thereafter. For practically all other liquids, \(v_{\text{sound}}\) decreases with rising temperature over the entire range in which the material remains a liquid.\textsuperscript{2156} (In most liquids, \(v_{\text{sound}}\) increases linearly, though only very slightly, with pressure. For example, \(v_{\text{sound}}\) in benzene rises ~17% when pressure is increased from 1 atm to 500 atm;\textsuperscript{2156} see also Section 10.3.3.) The speed of sound in water-ice just below the freezing point, estimated from Eqn. 4.30, is ~1000 m/sec. The speed of sound in dry air at 1 atm increases with rising temperature and is approximated by:\textsuperscript{1164}

\[
v_{\text{sound}} = 332 \times \left[1 + (0.003366)(T - 273)\right]^{1/2}\quad \text{[Eqn. 10.26]}
\]

3. Energy Absorption — Acoustic absorption and attenuation coefficients change with temperature, affecting the efficiency of acoustic power transfer. Absorption per unit of radio frequency (rf) energy in tissue during diathermic heating also varies with temperature.

4. Surface Tension — The surface tension of liquids at the air-liquid interface tends to decline as temperature rises, falling to zero at the boiling point. For instance, the air-liquid surface tension for a 48% volumetric ethanol-water mixture (96 proof, U.S. spirit) is 30.10 x 10\textsuperscript{-3} N/m at 20°C but 28.93 x 10\textsuperscript{-3} N/m at 40°C,\textsuperscript{763} values for pure water are given elsewhere (Section 9.2.3).

5. Dielectric Constant — The electrical properties of materials may be temperature-dependent. For example, the dielectric constant of water declines with rising temperature and is crudely approximated by:\textsuperscript{2157}

\[
\kappa_{\text{water}} = 80 - 0.4 (T - 293)\quad \text{[Eqn. 10.27]}
\]

for temperatures in kelvins from \(T = 273-373\) K (0-100°C), for rf frequencies up to 100 MHz, and at 1 atm pressure. The dielectric constant of ice at 0°C is virtually the same as that of water (88.0), but decreases rapidly with decreasing temperatures below 0°C, and with increasing frequency; by 0.1 MHz, \(\kappa_{\text{ice}} = 2-4\) with little influence of temperature.\textsuperscript{5157} Relative permittivity decreases by a large factor for many other substances as they change state at the freezing point; for example, \(\kappa\) falls from 35 for liquid nitrobenzene at 279 K to 3 for solid nitrobenzene at 279 K.\textsuperscript{727} In general, nonpolar liquids have a small dielectric constant (e.g., 1.5-2.5) that is nearly independent of the temperature, whereas polar liquids have a larger value that declines rapidly as temperature rises.\textsuperscript{2036}
Introduction

One of life’s pleasures is writing an afterword for a classic-in-the-making. Not only will some of the glory inevitably rub off, there’s also the illicit pleasure of having peeked at the future, like peeking at the Christmas presents before Christmas.

A few of the possibilities of this new field of nanomedicine have been hinted at, a few more have been sketched in some research papers, but only with the publication of *Nanomedicine* have we started to see the full richness of it.

Like cresting the top of a hill and beholding, for the first time and in one sweep, the whole of a new land, our minds are both captivated by the prospects and at another level churning with plans and ideas and tasks. For at the same time we see what is possible, we are also aware of the work that remains to be done to convert this vision into reality. *Nanomedicine* is more than just a description of what might be; it is a call to action. While it will take decades to convert the possible to the actual, that is what we are called upon to do—not only for the good of all, not only to advance our knowledge, not only to help future generations, but to help ourselves as well.

The Long View

While planning beyond a decade is rare in this society, our lives can span over a century. We should not be shortsighted or timid about this. When I was a child, my sister was wise and very old: all of twenty years in age! My parents, in their 40’s, were old beyond concepts of antiquity. Like the sky above and the ground beneath, they had existed since the beginnings of time—at least, of my time. Yet somehow I am now 47, and when I protest my age I am laughed at by my grandmother-in-law who views me, from her 90’s, as a mere youth.

The field of nanomedicine will take decades to develop, but those decades will pass and that future will arrive. Most of us will find we are still here: a bit older, a bit slower, perhaps a bit wiser, yet still filled with the excitement of life and the joy of living. Think of yourself on that future day, looking back on what was and looking forward to what will be. Will the future still be bright, still be open, filled with the excitement of life and the joy of living. Think of yourself on that future day, looking back on what was and looking forward to what will be. Will the future still be bright, still be open, filled with uncharted possibilities?

That depends on what we do today. If we ignore the future, if we dismiss the decades ahead and focus narrowly on the next few weeks or months, then the future will catch us by surprise, unprepared. But if we start now, if we raise up our eyes from the distractions of the moment and prepare for the future that we know will come, then when that future arrives we will look back and be pleased with what we did, and will look forward and be pleased with the even greater possibilities of what we can do.

The Tasks Ahead

Perhaps the first task is to decide whether the capabilities described so well in *Nanomedicine* are indeed possible. If they are, then developing this new technology is a matter, quite literally, of life and death for many of us, our children, and future generations. Making this decision is harder than it might seem. As a society, we deal with new ideas poorly if we deal with them at all. Most people do not have the intellectual resources to directly evaluate new proposals, and so must rely on the statements of others. But those who in principle might be able to evaluate a new idea and so help our collective understanding often get it wrong. Looking back at a few historical examples, we can begin to see the magnitude of the difficulty.

John Aubrey, a contemporary of William Harvey, wrote this account of the response to the publication in 1628 of Harvey’s book *De Motu Cordis* in which Harvey described his discovery of the blood’s circulation:

“...I heard Harvey say that after his book came out, he fell mightily in his practice. ’Twas believed by the vulgar that he was crack-brained, and all the physicians were against him. I knew several doctors in London that would not have given threepence for one of his medicines.”

In 1873 Sir John Erichsen offered this grim assessment of the future of surgery:

“There cannot always be fresh fields of conquest by the knife; there must be portions of the human frame that will ever remain sacred from its intrusions, at least in the surgeon’s hands. That we have already, if not quite, reached these final limits, there can be little question. The abdomen, the chest, and the brain will be forever shut from the intrusion of the wise and humane surgeon.”

*Nanomedicine*, as *Nanosystems* before it, is based on the laws of physics which describe our world with phenomenal accuracy. Both books advance arguments grounded on those laws, and both can therefore be evaluated with respect to the accuracy of their conclusions with respect to those laws. *Nanosystems* was published in 1992, and no significant flaws have been found. Given the volume of public debate and the number of people who have read the book, the simplest explanation for this absence of reported errors is that its logic is basically correct and its conclusions are basically sound. Today, these conclusions are working their way into our collective decision making processes and guiding our next steps. Research is being focused on how best to develop this new technology, companies are being formed to achieve the goals that we now accept as possible, and people are beginning to grapple with the potential consequences.
An Immediate Concern

Actually, there is one thing we must do even earlier: ensure the completion of this exceptional series of books. What you are reading is only Volume I. Volumes II and III, and the popular book to follow, do not yet (as of 1999 when this is being written) exist, except in Freitas’ head.

We need to support him, in order to move this first and most critical series to completion. This is always the hardest time for a new idea—before it has been codified and laid out, before it has been clothed in words, when it exists only as thoughts. The work of making it solid and substantial is great, and yet this work is given the least support.

What funding committee will agree to fund a book describing an entire new field that has never before been dreamt of? Committees base their conclusions on a shared understanding of a common body of knowledge. Their members are drawn from an existing society of experts to evaluate the next incremental improvement. What do you do when there are no experts? Who lays claim to expertise in nanomedicine? Who has spent their life in this field which is just being conceived? No one. The committee process breaks down when we move into truly new terrain. It fails us just when failure is most expensive: at the beginnings of new things. Here we must fall back on individuals—individuals who are bold enough to believe in themselves when there are no experts to turn to for help and support. Individuals who are willing to back up their own beliefs with action, who will nurture the truly new and the truly groundbreaking objectives often requires taking many steps, and all of those steps except the first one are (pretty much by definition) not experimentally accessible. While experimental work is focused on taking the next step, the theoretical and computational work should be focused on clarifying the whole pathway from today’s technology to the future applications. This feeds back into the experimental work in two ways. First, it provides information about which approaches are more likely or less likely to succeed. Second, it provides a reason for supporting the experimental work. The value and feasibility of the long-term objectives makes experimental progress more valuable, and as this understanding spreads it becomes easier for experimentalists to get funding for work that moves us closer to those long-range objectives.

Consider one example: Freitas’ respirocyte\(^3\) is based on the observation that a red blood cell stores very little oxygen when compared with a tank of similar size which holds oxygen compressed to ~1,000 atmospheres. The design calls for strong materials (to hold oxygen at high pressure) and very finely detailed structural components (to control the release and storage of the oxygen). On a theoretical and computational front, sub-components that are composed of not-too-many atoms can be modeled in great detail. The position of each atom and the forces between the atoms can be modeled using techniques that provide remarkable accuracy and as this understanding spreads it becomes easier for experimentalists to sidestep awkward issues that might cause the design to fail.

The requirement for highly detailed structures made of stiff hydrocarbons in turn implies we must analyze chemical reactions able to synthesize such materials. The chemical reactions involved
in the growth of diamond are reasonably well understood, and many reaction pathways have been proposed by which such growth can occur. We can adopt reaction pathways similar to those seen in the chemical vapor deposition (CVD) growth of diamond, but provide finer control over where they occur by positioning the reacting compounds using positional devices. Better computational methods for analyzing individual reactions are possible using ab initio methods, which can also provide accurate descriptions of the interactions of small numbers of atoms which then feed into the design of better PEFs. Better understanding of reactions relevant to the growth of diamond can also be pursued experimentally, and particular reactions of interest can be looked at in the laboratory as well as on a computer.

The need to position molecular components in its turn implies we must consider positional devices—both improvements to today’s SPMs (Scanning Probe Microscopes) and future molecular scale versions that are faster, more accurate, and have a greater range of tip configurations. This implies a strong interest in experimental and theoretical work on positioning devices, as well as work aimed at improving SPM tips. Experimental work that shows greater flexibility in arranging individual atoms and molecules should be supported, as the potential consequences of this work are very great.

This process of working backwards from our desired goal to near-term research objectives was called backward chaining by Drexler. As can be seen, it is a method of analyzing a long-term objective (e.g., using respirocytes to treat medical conditions) and breaking down the steps needed to achieve that objective into nearer-term objectives (e.g., improving PEFs, experimental work in SPMs). While the outline of the process given here is necessarily very short, it should give the reader a feeling for the basic idea.

The procedure of targeting near-term research goals based on their utility in achieving long-term objectives not only provides a focus for research, it also produces a wealth of results which further bolster the underlying arguments supporting the feasibility of the objectives and the desirability of such research. This creates a recursive spiral of knowledge. A little research shows there are no fundamental barriers that prevent us from achieving the objectives of molecular nanotechnology. Further research gives a better understanding of which molecular machine systems should be feasible and provides initial targets for additional research. Ongoing work is providing a clearer picture of the routes that can move us from our present technology base to the proposed molecular machines of the future, and produces yet more targets for near-term efforts.

The Recursive Spiral of Knowledge
Every time we pursue further research in nanotechnology we find that our original assessment of its basic feasibility is strengthened, our understanding of the specific near-term research targets that we must pursue is broadened, our conviction that further research can speed the development of this fundamentally new and revolutionary technology grows stronger, and our awareness of the astonishingly pervasive benefits this technology can bring is widened. In this recursive spiral of knowledge, research emboldens our interest and increasing interest produces yet more research. The rate-limiting process is the speed with which people take the first step, for having taken the first step the second step comes a little faster, and the third step faster still.

This should come as no real surprise, for either the ability to arrange and rearrange molecular structures in most of the ways permitted by physical law is feasible, or, alternatively, it is not. But since Feynman’s famous 1959 talk There’s Plenty of Room at the Bottom, every informed observer who has studied the issue has drawn the same conclusion: it’s feasible. The only way to break the recursive spiral would be to discover a fundamental objection that makes molecular machine systems impossible. As we are surrounded by biological molecular machines, this possibility seems remote. If thermal noise was a fundamental obstacle to molecular machine design, then biological systems could not copy DNA and molecular rotary motors could not rotate. If quantum uncertainty was a fundamental obstacle, then ribosomes could not synthesize proteins and sodium channels could not distinguish between sodium and potassium.

A New Medical Technology And A New Era Of Medicine
We are left, then, with a fairly clear set of conclusions. Living systems exist. Living systems can usually heal and cure their own injuries, unless those injuries are severe enough to prevent the living system from functioning. Too often, we suffer injuries that are indeed this severe. Molecular nanotechnology is feasible. As we master the ability to design molecular machines that can continue to function when the living system around them has failed, those molecular machines can restore the function of the living system. They can support and sustain the processes of the living system until that living system can once again function on its own. Whether this is done by a temporary assist from respirocytes or by any of the myriad other techniques discussed in Nanomedicine, the underlying message is clear: life and health can be restored and sustained in the face of greater injury, greater damage, greater trauma, and greater dysfunction than has ever before been realized. This will usher in a new era of medicine—an era in which health and long life will be the usual state of affairs while sickness, debility and death will be the mercifully rare exceptions.

The future capabilities of nanomedicine give hope and inspiration to those of us who still have decades of life to look forward to, but some are not so fortunate. Many others who rightfully should live several decades more might find that chance cuts short their expected time. Heart attacks and cancer can strike us down even in the prime of our lives. They do not always wait their turn and politely arrive only when expected. How can today’s dying patient take advantage of a future medical technology that is as yet only described in a handful of theoretical publications? How can we preserve the physical structure of our bodies well enough to permit that future medical technology to restore our health?

The extraordinary medical prospects ahead of us have renewed interest in a proposal made long ago: that the dying patient could be frozen, then stored at the temperature of liquid nitrogen for decades or even centuries until the necessary medical technology to restore health is developed. Called cryonics, this service is now available from several companies. Because final proof that this will work must wait until after we have developed a medical technology based on the foundation of a mature nanotechnology, the procedure is experimental. We cannot prove today that medical technology will (or will not) be able to reverse freezing injury 100 years from now. But the patient dying today must choose whether to join the experimental group or the control group. The luxury of waiting for a definitive answer before choosing is simply not available. So the decision must be made today, on the basis of incomplete information. We already know what happens to the control group. The outcome for the experimental group has not yet been confirmed. But given the wonderful advances that we see coming, it seems likely that we should be able to reverse freezing injury—especially when that injury is minimized by the rapid introduction through the vascular system of cryoprotectants and other chemicals to cushion the tissues against further injury.
Conclusion

The development of nanomedicine depends on us: what we do and how rapidly we do it. Research is not done by a faceless "them," nor is it something that happens spontaneously and without any human intervention. It is done by and supported by people. Unless we decide to support and pursue this research, it won’t happen. How long it takes to develop depends on us. We are not idle bystanders watching the world go by. We are a part of it. If we sit and wait for someone else to develop this technology, it will happen much more slowly. If we jump in and work to make it happen, it will happen sooner. And developing a life saving medical technology within our lifetimes seems like a very good idea—certainly better than the alternative.

References

1. For more examples of this kind and references for the above quotations, see http://www.foresight.org/News/negativeComments.html#loc026.
Appendix A. Useful Data in Medical Nanodevice Design

Compiled from various sources, including (but not limited to) the following References: 10, 460, 536-537, 567, 763, 1164, 1597, 1662, 2036, 2153, 2154, 2223-2224, 3229-3232, 3429-3431, 3469.

I. Systeme International (SI) Metric Prefixes

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Symbol</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yotta-</td>
<td>Y</td>
<td>10^24</td>
</tr>
<tr>
<td>Zeetta-</td>
<td>Z</td>
<td>10^21</td>
</tr>
<tr>
<td>Exa-</td>
<td>E</td>
<td>10^18</td>
</tr>
<tr>
<td>Peta-</td>
<td>P</td>
<td>10^15</td>
</tr>
<tr>
<td>Tera-</td>
<td>T</td>
<td>10^12</td>
</tr>
<tr>
<td>Giga-</td>
<td>G</td>
<td>10^9</td>
</tr>
<tr>
<td>Mega-</td>
<td>M</td>
<td>10^6</td>
</tr>
<tr>
<td>Kilo-</td>
<td>k, K</td>
<td>10^3</td>
</tr>
<tr>
<td>Hecto-</td>
<td>h</td>
<td>10^2</td>
</tr>
<tr>
<td>Deka-</td>
<td>da, dk</td>
<td>10^1</td>
</tr>
<tr>
<td>Deci-</td>
<td>d</td>
<td>10^-1</td>
</tr>
<tr>
<td>Centi-</td>
<td>c</td>
<td>10^-2</td>
</tr>
<tr>
<td>Milli-</td>
<td>m</td>
<td>10^-3</td>
</tr>
<tr>
<td>Micro-</td>
<td>µ, μ</td>
<td>10^-6</td>
</tr>
<tr>
<td>Nano-</td>
<td>n</td>
<td>10^-9</td>
</tr>
<tr>
<td>Pico-</td>
<td>p</td>
<td>10^-12</td>
</tr>
<tr>
<td>Femto-</td>
<td>f</td>
<td>10^-15</td>
</tr>
<tr>
<td>Atto-</td>
<td>a</td>
<td>10^-18</td>
</tr>
<tr>
<td>Zepto-</td>
<td>z</td>
<td>10^-21</td>
</tr>
<tr>
<td>Yocto-</td>
<td>y</td>
<td>10^-24</td>
</tr>
</tbody>
</table>

II. Units of Measurement and Useful Conversions

Angles and Rotations:
180 degrees = π radians
1 radian = 57.29578 deg = 0.1591549 cycle (revolution)
1 cycle (revolution) = 6.2831853 radians = 360 deg
1 deg = 0.01745329 radian = 0.0027777 cycle
= 60 arcmin = 3600 arcsec
1 arcsec = 4.848136 microradians
1 hertz (Hz) = 1 cycle/sec = 6.2831853 radians/sec
1 radian/sec = 9.5492966 revolutions per minute (rpm)

Areas and Surfaces:
1 m^2 = 10^4 cm^2 = 10^6 mm^2 = 10^{12} microns^2 = 10^{-18} nm^2
= 1.1959900 yd^2 = 10.763910 ft^2 = 1550.031 inch^2
1 acre = 43,560 ft^2 = 4046.8564 m^2 = 0.0015625 mile^2
1 stadian = 0.079577472 spheres = 3282.8063 deg^2
1 sphere = 12.566371 stadians

Concentration:
1 molar (M) = 1 mole/liter = 0.602257 molecules/nm^3
1 molecule/nm^3 = 1.66042 moles/liter
1 molal = 1 mole/kg
milliequivalent:
mEq/liter = (mg/liter) x valence / molecular weight
mEq/kg = (mg/kg) x valence / molecular weight

Density:
1 gm/cm^3 = 1000 kg/m^3 = 8.345444 lbs/gallon (U.S. liq.)
= 1 gm/ml = 1000 gm/liter = 62.42601 lbs/ft^3
1 gallon water at 4˚C (U.S. liq.) = 8.34517 lbs = 3.785305 kg

Electrical:
1 coulomb (C) = 1 ampere-sec = 6.24196 x 10^{18} electron charges
1 ampere (A) = 1 coulomb/sec
1 volt (V) (electric potential, emf) = 1 joule/coulomb
= 1 watt/ampere
1 ohm (Ω) = 1 volt/ampere
1 siemen (S) (conductance) = 1 ampere/volt = 1 mho
1 farad (F) (capacitance) = 1 coulomb/volt

Energy:
1 joule (J) (work, heat) = 1 N-m = 1 W-sec = 1 kg-m^2/sec^2
= 0.000239006 Kcal
= 10^7 ergs = 0.000984851 Btu
= 2.7777 x 10^{-7} kilowatt-hours
= 0.00986895 liter-atm
= 6.241506 x 10^{-18} eV
1 nanojoule (nJ) = 1000 picojoules (pJ) = 10^{-9} joules
1 zeptojoule (zJ) = 10^{-21} J = 6.241506 x 10^{-3} eV
= 0.1439325 Kcal
1 electron-volt (eV) = 160.217733 zJ
1 MeV = 10^6 eV = 1000 KeV = 0.001 GeV = 10^{-6} TeV
1 kilocalorie (Kcal) (dietary Calories) = 1000 calories = 4184 J
1 kT at 310 K = 4.28004 zJ = 0.026714 eV
1 Hartree = 627.5095 Kcal/mole = 27.2116 eV = 4359.7482 zJ
1 foot-pound = 1.355818 J
1 megaton (of TNT explosive) ~ 5 x 10^{15} J

Energy (molar):
1 Kcal/mole = 6.947700141 zJ/molecule
1 KJ/mole = 29.069177 zJ/molecule
1 zJ/molecule = 0.1439325 Kcal/mole = 0.0344007 KJ/mole
1 eV/atom = 23.0622 Kcal/mole = 160.217733 zJ/atom
**Force:**
1 newton (N) (force) = 1 kg-m/sec²
= 10⁵ dynes = 0.22480894 pounds
1 nanonewton (nN) = 1000 piconewtons (pN) = 10⁻⁹ newtons

**Illumination:**
1 candela (cd) (luminous intensity) = 1 lumen/m²
1 lumen (luminous flux) (at 550 nm wavelength) = 0.0014705882 watt
1 candle = 1 lumen/steradian
1 candle/m² (at 550 nm wavelength) = 0.00462 watt/m²

**Length and Distance:**
1 Angstrom (Å) = 0.1 nm = 10⁻¹⁰ meters
1 nanometer (nm) = 1000 picometers = 10⁻⁹ meters
1 micron = 1000 nm = 0.001 mm = 10⁻⁶ m
1 mile = 5280 ft = 63,360 inches = 1609.344 m
1 light-year = 9.46055 x 10¹⁵ m

**Magnetism:**
1 weber (Wb) (magnetic flux) = 1 volt-sec
1 henry (H) (inductance) = 1 weber/ampere
1 tesla (T) (magnetic flux density) = 1 weber/m² = 1 N/ampere-m

**Mass:**
1 kilogram (kg) = 1000 grams (gm) = 2.2046226 pounds (lbs)
= 771.61792 scruples (apoth) = 15,432.358 grains
1 nanogram (ng) = 1000 picograms (pg) = 10⁻⁹ grams
1 pound = 16 ounces (avdp oz) = 453.59237 gm
1 troy ounce = 31.103486 gm = 1.0971429 (avdp) oz
1 gram (gm) = 0.25720597 drams (apoth/troy)
= 0.5643839 drams (avdp)
= 15.432358 grains = 0.77161792 scruples (apoth)
1 ton (metric) = 1000 kg
1 ton (long) = 2240 lbs = 1016.0469 kg
1 ton (short) = 2000 lbs = 907.1847 kg
1 carat (metric) = 0.200 grams = 3.08647 grains
1 dalton = 1 amu = 931.752 MeV (mass energy)
= 1.661 x 10⁻²⁷ kg/molecule = 1 gm/mole

**Power:**
1 watt (W) = 1 joule/sec = 0.00134102 horsepower
= 10⁷ erg/sec
= 3.41443 Btu/hour = 20.650 Kcal/day
1 nanowatt (nW) = 1000 piconewtons (pW) = 10⁻⁹ watts
1 horsepower (hp) = 746 W
2000 Kcal/day = 96.85 W

**Pressure or Stress:**
1 Pascal (Pa) = 1 N/m²
1 dyne/cm² = 0.1 N/m² (Pa) = 10⁻⁶ bars = 0.986923 x 10⁻⁶ atm
1 atm = 101,325 N/m² (Pa) = 1.01325 bars
= 760 mmHg = 760 torr = 14.6960 lbs/in² (PSI)
1 mmHg = 0.00131579 atm = 13.332 N/m² (Pa)
= 0.00133322 bar
= 0.0193368 lbs/in² = 1.35953 cmH₂O
1 kg/mm² = 96.784 atm = 9.80665 x 10⁶ N/m² (Pa)
= 1422.33 lbs/in² = 98.0665 bars
= 10⁶ kg/m² = 100 kg/cm²

**Radioactivity:**
1 curie (Ci) = 3.70 x 10¹⁰ nuclear disintegrations/sec
1 bequerel (Bq) = 1 nuclear disintegration/sec
= 3.70 x 10⁻⁹ curies

**Temperature:**
Kelvin (K) degrees = Celsius (C) degrees + 273.15
1 deg K = 1 deg C (Celsius or Centigrade)
Fahrenheit (F) degrees = 32 + (9 x Celsius degrees / 5)
Celsius (C) degrees = 5 x (Fahrenheit degrees - 32) / 9
310 K = 37°C = 98.6°F

**Time:**
1 second = (1 / 31,556,925.9747) of a tropical year
1 day = 24 hours = 1440 minutes = 86,400 sec
1 year (tropical) = 365.24219 mean solar days = π x 10⁷ sec

**Velocity:**
1 m/sec = 3.2808399 ft/sec = 3.6 km/hr = 20.650 Kcal/day = 1.94385 knots (nautical miles per hour)
1 mph = 1.609344 km/hr

**Viscosity (absolute or dynamic):**
1 Poise = 0.1 kg/m-sec = 100 centipoise (cp) = 0.1 Pascal-sec

**Volume:**
1 cm³ (cc) = 1 milliliter (ml)
1 m³ = 10⁹ cm³ = 10¹⁵ microm³ = 10²⁷ nm³
1 cm³ = 10⁻³ liters = 10⁻⁶ m³ = 10⁻¹⁵ liters = 10⁻¹² cm³
1 micron³ = 10⁻¹⁸ m³ = 10⁻¹⁵ liters = 10⁻¹² cm³
1 nm³ = 10⁻³⁰ m³ = 10⁻²⁷ liters = 10⁻²⁴ cm³
1 liter = 10⁶ cm³ = 10⁸ micron³ = 10²⁴ nm³

**Albedo**
Earth, mean: 31%
Angular momentum:
proton (quantized spin): \(5.28 \times 10^{-35} \text{ J-sec}\)

Areal number density:
C atoms, monoatomic graphite sheet: \(\approx 38.2 / \text{nm}^2\)

Atmosphere, Earth:
\(\begin{align*}
\text{N}_2 & \quad 78.084\% \text{ (by volume)} \\
\text{O}_2 & \quad 20.946\% \\
\text{H}_2\text{O} & \quad 0.01\% \text{ (arctic)} - 3\% \text{ (tropical)} \\
\text{Ar} & \quad 0.934\% \\
\text{CO}_2 & \quad 0.031\%
\end{align*}\)

Atom and ion diameters (approx.):
\(\begin{align*}
\text{H} & : \quad 0.060 \text{ nm} \\
\text{F} & : \quad 0.128 \text{ nm} \\
\text{O} & : \quad 0.132 \text{ nm} \\
\text{N} & : \quad 0.140 \text{ nm} \\
\text{C} & : \quad 0.154 \text{ nm} \\
\text{He} & : \quad 0.186 \text{ nm} \\
\text{Na}^+ & : \quad 0.190 \text{ nm} \\
\text{S} & : \quad 0.208 \text{ nm} \\
\text{P} & : \quad 0.220 \text{ nm} \\
\text{Ne} & : \quad 0.224 \text{ nm} \\
\text{F}^- & : \quad 0.272 \text{ nm} \\
\text{O}^- & : \quad 0.280 \text{ nm} \\
\text{S}^- & : \quad 0.368 \text{ nm} \\
\text{Na} & : \quad 0.372 \text{ nm} \\
\text{K} & : \quad 0.462 \text{ nm} \\
\text{Fr} & : \quad 0.544 \text{ nm}
\end{align*}\)

Atomic mass unit (amu):
\(1.661 \times 10^{-27} \text{ kg (1/12th of C}^{12}\text{ nuclide mass)\)}

Avogadro's number:
\(N_A = 6.02257 \times 10^{23} \text{ molecules/mole}\)

Bandgap:
\(\begin{align*}
\text{Silicon:} & \quad 1.12 \text{ eV} \\
\text{Diamond:} & \quad 5.47 \text{ eV}
\end{align*}\)

Bohr radius:
\(5.29177249 \times 10^{-11} \text{ m}\)

Boltzmann's constant:
\(k = 1.38058 \times 10^{-23} \text{ J/molecule-K}\)

Bond energy (covalent, dissociation): \(556 \text{ zJ (C-C)}\)

Bond length:
\(\begin{align*}
\text{C-C in graphite:} & \quad 0.1415 \text{ nm} \\
\text{C-C in diamond:} & \quad 0.154484 \text{ nm}
\end{align*}\)

Breakdown voltage, electrical:
\(\begin{align*}
\text{Air:} & \quad 1.3 \times 10^6 \text{ volts/m} \\
\text{Porcelain:} & \quad 4 \times 10^6 \text{ volts/m} \\
\text{Titanium dioxide:} & \quad 3.9-8.3 \times 10^6 \text{ volts/m} \\
\text{Paraffin, 25° C:} & \quad 9.8 \times 10^6 \text{ volts/m} \\
\text{Pyrex glass:} & \quad 1.3 \times 10^7 \text{ volts/m} \\
\text{Paper:} & \quad 1.4 \times 10^7 \text{ volts/m}
\end{align*}\)

Bulk modulus, diamond:
\(K = (4.4-5.9) \times 10^{11} \text{ N/m}^2\)

Cleavage energy:
\(\begin{align*}
\text{Diamond [111] crystal plane:} & \quad 10.6 \text{ J/m}^2 \\
\text{Diamond [110] crystal plane:} & \quad 13.0 \text{ J/m}^2 \\
\text{Diamond [100] crystal plane:} & \quad 18.4 \text{ J/m}^2
\end{align*}\)

Compressibility, diamond:
\(1.7 \times 10^{-11} \text{ m}^2/\text{kg}\)

Crack velocity:
\(\begin{align*}
\text{Soda-lime glass (maximum):} & \quad 1580 \text{ m/sec} \\
\text{Sapphire (maximum):} & \quad 4500 \text{ m/sec} \\
\text{Diamond (maximum):} & \quad 7200 \text{ m/sec}
\end{align*}\)

Density:
\(\begin{align*}
\text{H}_2, \text{STP:} & \quad 0.0899 \text{ kg/m}^3 \\
\text{He, STP:} & \quad 0.1785 \text{ kg/m}^3 \\
\text{Air, 1 atm, 20° C:} & \quad 1.205 \text{ kg/m}^3 \\
\text{Air, STP:} & \quad 1.2929 \text{ kg/m}^3 \\
\text{N}_2 \text{ at 1000 atm, 37° C:} & \quad 493 \text{ kg/m}^3 \\
\text{O}_2 \text{ at 1000 atm, 37° C:} & \quad 672 \text{ kg/m}^3 \\
\text{CO}_2 \text{ at 1000 atm, 37° C:} & \quad 816 \text{ kg/m}^3 \\
\text{Crude oil:} & \quad 862 \text{ kg/m}^3 \\
\text{(137 kg/barrel)} \\
\text{Water-ice, STP:} & \quad 0.917 \text{ kg/m}^3 \\
\text{Water, 100° C:} & \quad 958.384 \text{ kg/m}^3 \\
\text{Water, 37° C:} & \quad 993.360 \text{ kg/m}^3 \\
\text{Water, 0° C:} & \quad 999.868 \text{ kg/m}^3 \\
\text{Water, 4° C:} & \quad 1000.000 \text{ kg/m}^3 \\
\text{Blood plasma, 37° C:} & \quad 1025 \text{ kg/m}^3 \\
\text{Glucose:} & \quad 1562 \text{ kg/m}^3 \\
\text{Silicon, 25° C:} & \quad 2330 \text{ kg/m}^3 \\
\text{Boron:} & \quad 2340 \text{ kg/m}^3 \\
\text{Quartz:} & \quad 2650 \text{ kg/m}^3 \\
\text{Diamond:} & \quad 3510 \text{ kg/m}^3 \\
\text{Sapphire:} & \quad 3970 \text{ kg/m}^3 \\
\text{Germanium, 25° C:} & \quad 5323 \text{ kg/m}^3 \\
\text{Earth (mean):} & \quad 5522 \text{ kg/m}^3 \\
\text{Gadolinium:} & \quad 7895 \text{ kg/m}^3 \\
\text{Polonium:} & \quad 9320 \text{ kg/m}^3 \\
\text{Silver, 20° C:} & \quad 10,500 \text{ kg/m}^3 \\
\text{Lead, 20° C:} & \quad 11,350 \text{ kg/m}^3 \\
\text{Gold, 20° C:} & \quad 19,320 \text{ kg/m}^3 \\
\text{Platinum, 20° C:} & \quad 21,450 \text{ kg/m}^3 \\
\text{Nylon, 20° C:} & \quad 1.9 \times 10^7 \text{ volts/m} \\
\text{Rubber, hard, 25° C:} & \quad 1.9 \times 10^7 \text{ volts/m} \\
\text{Polystyrene:} & \quad 2.0-2.8 \times 10^7 \text{ volts/m} \\
\text{Teflon:} & \quad 6 \times 10^7 \text{ volts/m} \\
\text{Ruby mica:} & \quad 1.6 \times 10^8 \text{ volts/m} \\
\text{SiC:} & \quad 4 \times 10^8 \text{ volts/m} \\
\text{Diamond:} & \quad 2 \times 10^9 \text{ volts/m}
\end{align*}\)

Dielectric constant:
\(\begin{align*}
\text{Vacuum:} & \quad 1.00000 \\
\text{Air:} & \quad 1.00059 \\
\text{Paraffin, 25° C:} & \quad 2.0-2.5 \\
\text{Teflon:} & \quad 2.1 \\
\text{Polystyrene:} & \quad 2.4-2.65 \\
\text{Rubber, hard, 25° C:} & \quad 2.8 \\
\text{Nylon, 25° C:} & \quad 3.5
\end{align*}\)
Paper: 3.5
Pyrex glass: 4.5
Diamond, 300 K: 5.7
Porcelain: 6.5
Sapphire: 9.0
SiC: 9.0
Ruby, parallel to optical axis: 11.28
Silicon: 12.0
Ruby, normal to optical axis: 13.27
Ethanol, 25˚C: 24.3
Glycerol, 25˚C: 42.5
Water, pure, 37˚C: 74.31
Water-ice, 0˚C: 88.0
Titanium dioxide: 100.0

Distance between adjacent graphite sheets: 0.335 nm
e: 2.718 281 828 459

Electron mobility:
Diamond (holes): 0.12 m²/volt-sec
Silicon: 0.15 m²/volt-sec
Diamond: 0.18-0.19 m²/volt-sec

Electron-volt: 1.602 177 33 x 10⁻¹⁹ J
Elementary charge: 1.602 177 33 x 10⁻¹⁹ coulomb

Energy content:
Glucose (burned in pure oxygen at 100% efficiency): 4.765 x 10⁻¹⁸ J/molecule
Sugar: -26 Kcal (-10⁷ J) per “lump”

Faraday constant: 96,487.0 coulomb/mole
Gravitational constant: \( G = 6.67259 \times 10^{-11}\) N·m²/kg²
Gyromagnetic ratio, protons: 4.26 x 10⁷ Hz/tesla

Heat capacity (\( C_V \)):
Air, room temp.: 1.19 x 10¹ J/m³·K
Water-ice, 77 K: 0.63 x 10⁶ J/m³·K
Water-ice, 249 K: 1.76 x 10⁶ J/m³·K
Diamond: 1.82 x 10⁶ J/m³·K
Water-ice, 271 K: 1.93 x 10⁶ J/m³·K
Platinum, 298 K: 2.84 x 10⁶ J/m³·K
Sapphire: 2.89 x 10⁶ J/m³·K
Water, 310 K: 4.19 x 10⁶ J/m³·K

Heat of fusion:
Water, 1 atm, 0˚C: 3.34 x 10⁸ J/m³
(79.7 cal/gm)

Heat of vaporization:
Water, 1 atm, 100˚C: 2.17 x 10⁹ J/m³
(539 cal/gm)

Hemoglobin:
Atom count, per molecule: -10,000 atoms/molecule
Free Hb, phagocytic half-life: 10-30 minutes
Molecular weight: -67,000 daltons
Molecule count, per red cell: -2.7 x 10⁶ molecules/RBC

Lattice spacing at 300 K:
Diamond: 0.356 683 nm
Silicon: 0.543 095 nm
Germanium: 0.564 613 nm
Tin, grey: 0.648 920 nm

Magnetic moment:
Electron: 9.2732 x 10⁻²⁴ J/tesla (Bohr magneton)
Proton: 1.41049 x 10⁻²⁶ J/tesla

Mass, particle:
Electron rest mass: \( m_e = 9.109 389 7 \times 10^{-31}\) kg
Proton rest mass: \( m_p = 1.672 623 1 \times 10^{-27}\) kg
Neutron rest mass: \( m_n = 1.674 928 6 \times 10^{-27}\) kg

Mass, planetary:
Atmosphere (Earth): 5.27 x 10¹⁸ kg
Oceans (Earth): 1.36 x 10²¹ kg
Moon: 7.34 x 10²² kg
Earth: 5.98 x 10²⁴ kg
Jupiter: 1.90 x 10²⁷ kg
Sun: 1.97 x 10³⁰ kg

Molar volume of ideal gas at STP:
0.022 414 10 m³/mole
37.22 nm³/molecule
0.02687 molecule/nm³

Molecular mass:
Diamond: 1.99 x 10⁻²⁶ kg/atom
Water: 2.99 x 10⁻²⁶ kg/molecule
\( \text{N}_2 \): 4.65 x 10⁻²⁶ kg/molecule
\( \text{O}_2 \): 5.32 x 10⁻²⁶ kg/molecule
\( \text{CO}_2 \): 7.31 x 10⁻²⁶ kg/molecule
Glucose: 2.99 x 10⁻²⁵ kg/molecule

Molecular number density:
Glucose, solid: 5.24 molecules/nm³
\( \text{N}_2 \) at 1000 atm: 10.6 molecules/nm³
\( \text{CO}_2 \) at 1000 atm: 11.1 molecules/nm³
\( \text{O}_2 \) at 1000 atm: 12.6 molecules/nm³
Water-ice at 273 K: ~30 molecules/nm³
Water, liquid: 33.4 molecules/nm³
Diamond: ~176 carbons/nm³

Molecular volume:
Diamond: 0.00567 nm³/atom
Water, liquid: 0.0299 nm³/molecule
\( \text{O}_2 \) at 1000 atm: 0.0791 nm³/molecule
\( \text{CO}_2 \) at 1000 atm: 0.0897 nm³/molecule
\( \text{N}_2 \) at 1000 atm: 0.0943 nm³/molecule
Glucose, solid: 0.1910 nm³/molecule

Permeability (magnetic) constant: \( \mu_0 = 1.256 637 061 4 \times 10^{-6}\) henry/m
Permittivity (electric) constant: 
\[ \varepsilon_0 = 8.854 \, 187 \, 817 \times 10^{-12} \text{ farad/m} \]
\[ \pi = 3.141 \, 592 \, 653 \, 589 \]

Planck's constant:
\[ h = 6.626 \, 075 \, 5 \times 10^{-34} \text{ J-sec} \]
\[ \hbar = h / 2 \pi = 1.054 \, 572 \, 7 \times 10^{-34} \text{ J-sec} \]

Power dissipation, global:
- Human physiological worldwide: 5.3 x 10^{11} W (in 1998)
- Human technological worldwide: 1.2 x 10^{13} W (in 1998)
- Earth vegetation worldwide: 1.4 x 10^{14} W

Radius, particle:
- Proton (Compton wavelength): 1.3214 \times 10^{-15} \text{ m}
- Electron (classical): 2.81777 \times 10^{-15} \text{ m}

Radius, mean planetary:
- Moon: 1.738 \times 10^{6} \text{ m}
- Earth: 6.371 \times 10^{6} \text{ m}
- Jupiter: 6.988 \times 10^{7} \text{ m}
- Sun: 6.938 \times 10^{8} \text{ m}
- Earth-Moon orbit: 3.80 \times 10^{8} \text{ m}
- Earth-Sun orbit: 1.496 \times 10^{11} \text{ m}

Refractive index at \( \lambda = 589.29 \text{ nm (Na)} \):
- Air, STP, rel. to vacuum: 1.0002926
- Water-Ice: 1.3104
- Water, pure, 37°C: 1.33093
- Water, pure, 20°C: 1.33299
- Ethanol, 99.8%, 37°C: 1.35348
- Sucrose, 25% aq. solution, 20°C: 1.3723
- Sucrose, 50% aq. solution, 20°C: 1.4200
- Quartz, fused, 18°C: 1.45845
- Sucrose, 75% aq. solution, 20°C: 1.4774
- Sucrose, 75% aq. solution, 30°C: 1.4793
- Rock salt: 1.544
- Glass: 1.52-1.89
- Diamond: 2.41726

Resistivity:
- Diamond, synthesized: 10^7-10^8 \text{ ohm-m}
- Diamondoid materials: up to 10^{10} \text{ ohm-m}
- Diamond, natural, dark: up to 10^{18} \text{ ohm-m}

Rest mass energy:
\[ (mc^2/m) = 8.99 \times 10^{16} \text{ J/kg} \]

Saturation velocity (electronic), diamond: 2.7 \times 10^5 \text{ m/sec}

Shear modulus, diamond: \[ G = (4.8-5.5) \times 10^{11} \text{ N/m}^2 \]

Solar constant (at Earth): 1370 W/m^2

Solar luminosity (mean, total): 3.92 \times 10^{16} \text{ W}

Speed of light in vacuum:
\[ c = 2.99 \, 792 \, 458 \times 10^8 \text{ m/sec} \]

Speed of sound:
- Air, dry, STP: 331.36 m/sec
- Water, 15°C: 1450 m/sec
- Water, 37°C: 1500 m/sec
- Sapphire: 10,000 m/sec
- Diamond: 17,300 m/sec

Standard atmosphere:
\[ 1.01325 \times 10^5 \text{ N/m}^2 (760 \text{ mmHg}) \]

Standard molar gas volume at STP: 22.4141 liter-atm/mole

Standard temperature and pressure (STP): 273.15 K and 1 atm

Stefan-Boltzmann constant:
\[ \sigma = 5.6697 \times 10^{-8} \text{ W/m}^2\text{-K}^4 \]

Surface area, Earth: 5.10 \times 10^{14} \text{ m}^2

Surface energy of graphite: 0.234 J/m^2

Thermal conductivity (K_t):
- Air, 1 atm, 290 K: 0.02524 W/m-K
- Glass wool, 293 K: 0.038 W/m-K
- Cork, 293 K: 0.042 W/m-K
- Wood, 293 K: 0.10 W/m-K
- Water, 273 K: 0.561 W/m-K
- Water, 310 K: 0.623 W/m-K
- Water, 373 K: 0.681 W/m-K
- Brick, 293 K: 0.71 W/m-K
- Water-ice, 270 K: 1.56 W/m-K
- Water-ice, 249 K: 2.21 W/m-K
- Sapphire: normal to c-axis, 310 K: 2-20 W/m-K parallel to c-axis, 310 K: 20-40 W/m-K
- Lead, 293 K: 35 W/m-K
- Platinum, 300 K: 73 W/m-K
- Iron, 300 K: 80.3 W/m-K
- Aluminum, 300 K: 237 W/m-K
- Gold, 300 K: 315 W/m-K
- Silver, 300 K: 427 W/m-K
- Diamond, natural, 293 K: 2000 W/m-K
- Diamond, pure C_12, 293 K: 3500 W/m-K
- Diamond, natural, 65 K: 17,500 W/m-K

Thermal emissivity (er):
- Silver, polished: 0.02
- Aluminum: 0.022-0.08
- Platinum: 0.037-0.05
- Porcelain: 0.25-0.50
- Brass: 0.60
- Water: 0.67
- Earth's surface, mean: 0.69
- Glass: 0.76-0.94
- Cotton cloth: 0.77
- Wood: 0.80-0.90
- Corundum, emory, 80°C: 0.86
- Snow: 0.89
- Paper: 0.93
- Asphalt pavement: 0.93
- Concrete, rough: 0.94
- Water-ice: 0.97
- Human skin: 0.97-0.99
- Carbon black: 0.97-0.99
**Thermal expansion coefficient, linear:**

<table>
<thead>
<tr>
<th>Material</th>
<th>Coefficient (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diamond, 293 K</td>
<td>0.8 x 10^{-6}</td>
</tr>
<tr>
<td>Glass, pyrex, 273-373 K</td>
<td>3.6 x 10^{-6}</td>
</tr>
<tr>
<td>Sapphire, normal</td>
<td></td>
</tr>
<tr>
<td>to c-axis, 293 K</td>
<td>4.5 x 10^{-6}</td>
</tr>
<tr>
<td>Sapphire, parallel</td>
<td></td>
</tr>
<tr>
<td>to c-axis, 293 K</td>
<td>5.5 x 10^{-6}</td>
</tr>
<tr>
<td>Glass, ordinary, 273-373 K</td>
<td>8.9 x 10^{-6}</td>
</tr>
<tr>
<td>Platinum, 273-373 K</td>
<td>9.0 x 10^{-6}</td>
</tr>
<tr>
<td>Steel, 273-373 K</td>
<td>12.0 x 10^{-6}</td>
</tr>
<tr>
<td>Marble, 273-373 K</td>
<td>11.7 x 10^{-6}</td>
</tr>
<tr>
<td>Copper, 273-373 K</td>
<td>14.1 x 10^{-6}</td>
</tr>
<tr>
<td>Brass, 273-373 K</td>
<td>19.3 x 10^{-6}</td>
</tr>
<tr>
<td>Aluminum, 273-373 K</td>
<td>23.8 x 10^{-6}</td>
</tr>
<tr>
<td>Lead, 273-373 K</td>
<td>29 x 10^{-6}</td>
</tr>
<tr>
<td>Water-ice, 363-373 K</td>
<td>51 x 10^{-6}</td>
</tr>
<tr>
<td>Hard rubber, 273-373 K</td>
<td>80 x 10^{-6}</td>
</tr>
</tbody>
</table>

**Thermal expansion coefficient, volumetric:**

<table>
<thead>
<tr>
<th>Material</th>
<th>Coefficient (K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vitreous silica, 298 K</td>
<td>1.2 x 10^{-6}</td>
</tr>
</tbody>
</table>

**Universal gas constant:**

\[ R_{\text{gas}} = 8.206 \times 10^{-5} \text{ m}^3\text{-atm/mole-K} \]
\[ = 8.314510 \text{ J/mole-K} \]
\[ = 1.99 \text{ cal/mole-K} \]

**Year, tropical, in ephemeris (mean solar) seconds:**

3.1556926 x 10^7 sec

**Young's modulus:**

<table>
<thead>
<tr>
<th>Material</th>
<th>Modulus (N/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sapphire</td>
<td>E = 4.0 x 10^{11}</td>
</tr>
<tr>
<td>Diamond</td>
<td>E = 1.05 x 10^{12}</td>
</tr>
</tbody>
</table>
## Appendix B. Concentrations of Human Blood Components
(compiled and adapted from References 749, 753–757, 825, 1019, 1604, 1712, 2223).

<table>
<thead>
<tr>
<th>Component</th>
<th>In Whole Blood (gm/cm³)</th>
<th>In Plasma or Serum (gm/cm³)</th>
<th>Component</th>
<th>In Whole Blood (gm/cm³)</th>
<th>In Plasma or Serum (gm/cm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>0.81-0.86</td>
<td>0.93-0.95</td>
<td>Ascorbic acid (Vitamin C)</td>
<td>1-15 x 10⁻⁶</td>
<td>6-20 x 10⁻⁶</td>
</tr>
<tr>
<td>Acetocacetate</td>
<td>8-40 x 10⁻⁷</td>
<td>4.43 x 10⁻⁷</td>
<td>Aspartic acid</td>
<td>0-3 x 10⁻⁶</td>
<td>9-12 x 10⁻⁹</td>
</tr>
<tr>
<td>Acetone</td>
<td>3.20 x 10⁻⁶</td>
<td></td>
<td>Aspartic acid (in WBCs)</td>
<td>2.5-4.0 x 10⁻⁴</td>
<td>5-5.7 x 10⁻⁴</td>
</tr>
<tr>
<td>Acetylcholine</td>
<td>6.6-8.2 x 10⁻⁸</td>
<td></td>
<td>Bicarbonate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenosine triphosphate</td>
<td>- total</td>
<td>3.1-5.7 x 10⁻⁴</td>
<td>Bile acids</td>
<td>2-30 x 10⁻⁶</td>
<td>3-30 x 10⁻⁶</td>
</tr>
<tr>
<td></td>
<td>- phosphorus</td>
<td>5-10 x 10⁻⁵</td>
<td>Bilirubin</td>
<td>2-14 x 10⁻⁶</td>
<td>1-10 x 10⁻⁶</td>
</tr>
<tr>
<td>Adrenocorticotrophic hormone</td>
<td>- @ 6AM, mean</td>
<td>2.5-12 x 10⁻¹¹</td>
<td>Blood Urea Nitrogen (BUN)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- @ 6AM, maximum</td>
<td>5.5 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- @ 6PM, mean</td>
<td>&lt;12 x 10⁻¹¹</td>
<td>Bradykinin</td>
<td>8-23 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- @ 6PM, maximum</td>
<td>3.5 x 10⁻¹¹</td>
<td>Bromide</td>
<td>7 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>Alanine</td>
<td>2.7-5.5 x 10⁻⁵</td>
<td>2.4-7.6 x 10⁻⁵</td>
<td>Cadmium</td>
<td>7-10 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Albumin</td>
<td>3.5-5.2 x 10⁻²</td>
<td>1-88 x 10⁻⁸</td>
<td>- normal</td>
<td>1-5 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Aluminum</td>
<td>1-40 x 10⁻⁸</td>
<td></td>
<td>- toxic</td>
<td>0.1-3 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Aldosterone</td>
<td>- supine</td>
<td>3-10 x 10⁻¹¹</td>
<td>Calciferol (Vitamin D₂)</td>
<td>1.7-4.1 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- standing, male</td>
<td>6-22 x 10⁻¹¹</td>
<td>Calcitonin (CT)</td>
<td>&lt;1.0 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- standing, female</td>
<td>5-30 x 10⁻¹¹</td>
<td>Calcium</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino acids</td>
<td>- total</td>
<td>3.8-5.3 x 10⁻⁴</td>
<td>- ionized</td>
<td>4.48-4.92 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- nitrogen</td>
<td>4.6-6.8 x 10⁻⁵</td>
<td>- total</td>
<td>4.25-5.25 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>α-Aminobutyric acid</td>
<td>1-2 x 10⁻⁶</td>
<td>1-2 x 10⁻⁶</td>
<td>Carbon dioxide (respiratory gas)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>δ-Aminolevulinic acid</td>
<td>1.5-2.3 x 10⁻⁷</td>
<td></td>
<td>- arterial</td>
<td>8.8-10.8 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>Ammonia nitrogen</td>
<td>1-2 x 10⁻⁶</td>
<td>1.0-4.9 x 10⁻⁷</td>
<td>- venous</td>
<td>3.0-7.9 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>AMP, cyclic</td>
<td>- male</td>
<td>5.6-10.9 x 10⁻⁹</td>
<td>Carbon monoxide (as HbCO)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>- female</td>
<td>3.6-8.9 x 10⁻⁹</td>
<td>- nonsmokers</td>
<td>0.5-1.5% total Hb</td>
<td></td>
</tr>
<tr>
<td>Androstenedione</td>
<td>- male &gt;18 yrs</td>
<td>2-30 x 10⁻¹⁰</td>
<td>- smokers,</td>
<td>0.5-1.5% total Hb</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- female &gt;18 yrs</td>
<td>8-30 x 10⁻¹⁰</td>
<td>1-2 packs/day</td>
<td>4-5% total Hb</td>
<td></td>
</tr>
<tr>
<td>Androstosterone</td>
<td></td>
<td>1.5 x 10⁻⁷</td>
<td>- smokers,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Angiotensin I</td>
<td>1.1-8.8 x 10⁻¹¹</td>
<td></td>
<td>2 packs/day</td>
<td>8-9% total Hb</td>
<td></td>
</tr>
<tr>
<td>Angiotensin II (Renin)</td>
<td>1.2-3.6 x 10⁻¹¹</td>
<td></td>
<td>- toxic</td>
<td>&gt;20% total Hb</td>
<td></td>
</tr>
<tr>
<td>α-Antitrypsin</td>
<td>7.8-20 x 10⁻⁴</td>
<td></td>
<td>- lethal</td>
<td>&gt;50% total Hb</td>
<td></td>
</tr>
<tr>
<td>Arginine</td>
<td>6-17 x 10⁻⁶</td>
<td>1.3-3.6 x 10⁻⁵</td>
<td>Carcinoembryonic antigen</td>
<td>&lt;2.5 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Arsenic</td>
<td>- normal range</td>
<td>2-62 x 10⁻⁹</td>
<td>β-Carotene</td>
<td>3-25 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- chronic poisoning</td>
<td>100-500 x 10⁻⁹</td>
<td>Carotenoids</td>
<td>2.4-23.1 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- acute poisoning</td>
<td>600-9300 x 10⁻⁹</td>
<td>Cephalin</td>
<td>3-11.5 x 10⁻⁴</td>
<td>0-1 x 10⁻⁴</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ceruloplasmin</td>
<td>1.5-6 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Chloride, as NaCl</td>
<td>4.5-5 x 10⁻³</td>
<td>3.5-3.8 x 10⁻³</td>
</tr>
<tr>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------</td>
<td>----------------------------</td>
<td>---------------------------------</td>
<td>-------------------------</td>
<td>----------------------------</td>
</tr>
<tr>
<td>Cholecalciferol (Vitamin D3)</td>
<td></td>
<td></td>
<td>Copper</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1, 25-dihydroxy</td>
<td>2.5-4.5 x 10⁻¹¹</td>
<td></td>
<td>- male</td>
<td>9-15 x 10⁻⁷</td>
<td>7-14 x 10⁻⁷</td>
</tr>
<tr>
<td>24,25-dihydroxy</td>
<td>1.5 x 10⁻⁹</td>
<td></td>
<td>female</td>
<td>8-15.5 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>25-hydroxy</td>
<td>1.4-8 x 10⁻⁹</td>
<td></td>
<td>Corticosteroids (male &amp; female)</td>
<td>1-4 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Cholecystokinin (pancreozymin)</td>
<td>6.04 x 10⁻¹¹</td>
<td></td>
<td>Corticosterone</td>
<td>4-20 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Cholesterol</td>
<td></td>
<td>0.5-2 x 10⁻³</td>
<td>Cortisol</td>
<td>3-23 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>-LDLC</td>
<td>2.9-9 x 10⁻⁴</td>
<td>- 8 AM</td>
<td>- 4 PM</td>
<td>6-23 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>-total</td>
<td>1.15-2.25 x 10⁻³</td>
<td>- 10 PM</td>
<td>- 50% of 8 AM value</td>
<td>3-15 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>Choline, total</td>
<td>1.1-3.1 x 10⁻⁴</td>
<td>3.6-3.5 x 10⁻⁴</td>
<td>C-peptide</td>
<td>0.5-2 x 10⁻⁹</td>
<td>4 x 10⁻⁹</td>
</tr>
<tr>
<td>Chorionic gonadotropin</td>
<td></td>
<td></td>
<td>C-reactive protein</td>
<td>6.8-820 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>- Menstrual</td>
<td>0-3 x 10⁻¹¹</td>
<td></td>
<td>Creatine</td>
<td>1.7-5.0 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- Pregnancy, 1st trimester</td>
<td>5-3300 x 10⁻¹⁰</td>
<td></td>
<td>- male</td>
<td>3.5-9 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- Pregnancy, 2nd trimester</td>
<td>20-1000 x 10⁻¹⁰</td>
<td></td>
<td>- female</td>
<td>1.0-2 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>- Pregnancy, 3rd trimester</td>
<td>20-50 x 10⁻¹⁰</td>
<td></td>
<td>Creatinine</td>
<td>5-15 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- Menopausal</td>
<td>3-30 x 10⁻¹¹</td>
<td></td>
<td>- male</td>
<td>0.8-1.5 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>Citric acid</td>
<td>1.3-2.5 x 10⁻⁵</td>
<td>1.6-3.2 x 10⁻⁵</td>
<td>Cystine</td>
<td>0.7-1.2 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>Citrulline</td>
<td>2-10 x 10⁻⁶</td>
<td></td>
<td>Cyanide</td>
<td>4 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Clotting Factors</td>
<td></td>
<td></td>
<td>- nonsmokers</td>
<td>6 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- I Fibrinogen</td>
<td>1.2-1.6 x 10⁻³</td>
<td>2.4 x 10⁻³</td>
<td>- smokers</td>
<td>10-60 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- II Prothrombin</td>
<td>1 x 10⁻⁴</td>
<td></td>
<td>- nitroprusside therapy</td>
<td>&gt;100 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- III Tissue thromboplastin</td>
<td>1 x 10⁻⁴</td>
<td></td>
<td>- toxic</td>
<td>&gt;1000 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- V Proaccelerin</td>
<td>5-12 x 10⁻⁶</td>
<td></td>
<td>- lethal</td>
<td>&gt;1000 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- VII Proconvertin</td>
<td>1 x 10⁻⁶</td>
<td></td>
<td>Dihydroepiandrosterone</td>
<td>6-12 x 10⁻⁶</td>
<td>1.8-5 x 10⁻³</td>
</tr>
<tr>
<td>- VIII Antihemophilic factor</td>
<td>1 x 10⁻⁷</td>
<td></td>
<td>DHEA</td>
<td>1.99-3.34 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- IX Christmas factor</td>
<td>4 x 10⁻⁶</td>
<td></td>
<td>DHEA sulfate (male)</td>
<td>1.67-3.6 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- X Stuart factor</td>
<td>5 x 10⁻⁶</td>
<td></td>
<td>DHEA sulfate (female)</td>
<td>1.0-6.0 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- XI Plasma thrmb. anteced.</td>
<td>4 x 10⁻⁶</td>
<td></td>
<td>- newborn</td>
<td>2-3.3 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>- XII Hageman factor</td>
<td>2.9 x 10⁻⁵</td>
<td></td>
<td>- pre-pubertal children</td>
<td>1.0-6.0 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- XIII Fibrin stabiliz. factor</td>
<td>1 x 10⁻⁵</td>
<td></td>
<td>- premenopausal</td>
<td>8.2-33.8 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- Fibrin split products</td>
<td>&lt;1 x 10⁻⁵</td>
<td></td>
<td>- pregnancy</td>
<td>2.3-11.7 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- Fletcher factor</td>
<td>5 x 10⁻⁵</td>
<td></td>
<td>- postmenopausal</td>
<td>1.1-6.1 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- Fitzgerald factor</td>
<td>7 x 10⁻⁵</td>
<td></td>
<td>11-Deoxycorticisol</td>
<td>1-7 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>- von Willebrand factor</td>
<td>7 x 10⁻⁶</td>
<td></td>
<td>Dihydrotestosterone (DHT)</td>
<td>3-8 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Cobalamin (Vitamin B₁₂)</td>
<td>6-14 x 10⁻¹⁰</td>
<td>1-10 x 10⁻¹⁰</td>
<td>- male</td>
<td>1-10 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Corcarboxylase</td>
<td>7.9 x 10⁻⁸</td>
<td></td>
<td>- female</td>
<td>1-10 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>Complement</td>
<td></td>
<td></td>
<td>Diphosphoglycerate (phosphate)</td>
<td>8-16 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>- Clq</td>
<td>5.8-7.2 x 10⁻⁵</td>
<td></td>
<td>DNA</td>
<td>0-1.6 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>- C1r</td>
<td>2.5-3.8 x 10⁻⁵</td>
<td></td>
<td>Dopamine</td>
<td>&lt;1.36 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>- C1s (C1 esterase)</td>
<td>2.5-3.8 x 10⁻⁵</td>
<td></td>
<td>Enzymes, total</td>
<td>&lt;6 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>- C2</td>
<td>2.2-3.4 x 10⁻⁵</td>
<td></td>
<td>Epidermal growth factor (EGF)</td>
<td>&lt;1 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>- C3 (β₂C-globulin)</td>
<td>8-15.5 x 10⁻⁴</td>
<td></td>
<td>Epinephrine</td>
<td>3.1-9.5 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>- Factor B</td>
<td>2.4-3.2 x 10⁻⁵</td>
<td></td>
<td>- after 15 min rest</td>
<td>3.8 x 10⁻⁹</td>
<td>2-2.5 x 10⁻⁹</td>
</tr>
<tr>
<td>(C3 proactivator)</td>
<td>1-3 x 10⁻⁹</td>
<td></td>
<td>when emitted</td>
<td>Ergothioneine</td>
<td>1-20 x 10⁻⁵</td>
</tr>
<tr>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>----------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Erythrocytes (#/cm³)</td>
<td></td>
<td></td>
<td>α₁-Fetoprotein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult male, avg.</td>
<td>5.2 (4.6-6.2) x 10⁹</td>
<td></td>
<td>Flavine adenine dinucleotide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(range)</td>
<td></td>
<td></td>
<td>Fluoride</td>
<td>1-4.5 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- adult female, avg.</td>
<td>4.6 (4.2-5.4) x 10⁹</td>
<td></td>
<td>Folate</td>
<td>2.2-17.3 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>(range)</td>
<td></td>
<td></td>
<td>- in erythrocyte</td>
<td>1.67-7.07 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>- children,</td>
<td>4.5-5.1 x 10⁹</td>
<td></td>
<td>Folic acid (Vitamin M)</td>
<td>2.3-5.2 x 10⁻⁸</td>
<td>1.6-2 x 10⁻⁸</td>
</tr>
<tr>
<td>varies with age</td>
<td></td>
<td></td>
<td>Fructose</td>
<td>0.5 x 10⁻⁵</td>
<td>7-8 x 10⁻⁵</td>
</tr>
<tr>
<td>- reticulocytes</td>
<td>25-75 x 10⁹</td>
<td></td>
<td>Furosemide glucuronide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Erythropoietin</td>
<td></td>
<td></td>
<td>Galactose (children)</td>
<td>&lt;2 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>- adult, normal</td>
<td>0.5-2.5 x 10¹⁰</td>
<td></td>
<td>Gastric inhibitory peptide (GIP)</td>
<td>&lt;1.25-4.0 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>- pregnant</td>
<td>2.7-6.2 x 10¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- hypoxia or anemia</td>
<td>0.8-8.0 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estradiol (E₂) (male)</td>
<td>8-36 x 10⁻¹²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estradiol (E₂) (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- follicular (days 1-10)</td>
<td>1-9 x 10⁻¹¹</td>
<td>2.2-4 x 10⁻²</td>
<td>Globulin, total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- mean</td>
<td>5 x 10⁻¹¹</td>
<td>1-4 x 10⁻³</td>
<td>α₁-Globulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pre-fertile (days 10-12)</td>
<td>10-15 x 10⁻¹¹</td>
<td>4-10 x 10⁻³</td>
<td>α₂-Globulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- fertile (days 12-14)</td>
<td>35-60 x 10⁻¹¹</td>
<td>5-12 x 10⁻³</td>
<td>β-Globulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- luteal (days 15-28)</td>
<td>20-40 x 10⁻¹¹</td>
<td>6-17 x 10⁻³</td>
<td>γ-Globulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy</td>
<td>3-70 x 10⁻⁷</td>
<td>7.1-7.9 x 10⁻³</td>
<td>Glucagon</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- postmenopausal</td>
<td>1-3 x 10⁻¹¹</td>
<td>5-15 x 10⁻¹⁵</td>
<td>Glucose</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estradiol (E₃) (male)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estradiol (E₃) (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- nonpregnant</td>
<td>&lt;2 x 10⁻⁹</td>
<td>4.2-5.5 x 10⁻⁴</td>
<td>Glucosamine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 22-30</td>
<td>3.5 x 10⁻⁹</td>
<td>5.2-6.9 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 32-37</td>
<td>6-11 x 10⁻⁹</td>
<td>6.1-8.2 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 38-41</td>
<td>25-170 x 10⁻⁹</td>
<td>7.0-8.9 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estrogen (male)</td>
<td>4-11.5 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Estrogen (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- prepubertal</td>
<td>&lt;4 x 10⁻¹¹</td>
<td>2-4 x 10⁻³</td>
<td>Glutathione, reduced</td>
<td>2.5-4.1 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>- 1-10 days</td>
<td>6.1-39.4 x 10⁻¹¹</td>
<td>7-10.5 x 10⁻⁴</td>
<td>Glycerol, free</td>
<td>2.9-17.2 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- 11-20 days</td>
<td>12.2-43.7 x 10⁻¹¹</td>
<td>8-54 x 10⁻⁴</td>
<td>Glycine</td>
<td>1.7-2.3 x 10⁻⁵</td>
<td>0</td>
</tr>
<tr>
<td>- 21-30 days</td>
<td>15.6-35 x 10⁻¹¹</td>
<td>8-11 x 10⁻⁶</td>
<td>Glycogen</td>
<td>1.2-16.2 x 10⁻⁵</td>
<td>0</td>
</tr>
<tr>
<td>- postmenopausal</td>
<td>&lt;4 x 10⁻¹¹</td>
<td>2-28 x 10⁻⁶</td>
<td>Glycoprotein, acid</td>
<td>4-15 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>Estrone (E₁)</td>
<td></td>
<td></td>
<td>GMP, cyclic</td>
<td>0.6-4.4 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Estrone (female)</td>
<td></td>
<td></td>
<td>Gonadotropic releasing hormone</td>
<td>1-80 x 10⁻¹²</td>
<td>0</td>
</tr>
<tr>
<td>- male</td>
<td>2.9-17 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female, follicular</td>
<td>2-15 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female, 1-10 days of cycle</td>
<td>4.3-18 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female, 11-20 days of cycle</td>
<td>7.5-19.6 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female, 20-29 days of cycle</td>
<td>13.1-20.1 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 22-30</td>
<td>3.5 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 32-37</td>
<td>5-6 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 38-41</td>
<td>7-10 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- social high</td>
<td>0.5 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- reduced coordination</td>
<td>0.8 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- depression of CNS</td>
<td>&gt;1 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- confusion, falling down</td>
<td>2.0 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- loss of consciousness</td>
<td>3.0 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- coma, death</td>
<td>&gt;4 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fat, neutral (see Triglycerides)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatty acids, nonesterified</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(free)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- male</td>
<td>5.3-9 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female</td>
<td>7.2-10 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatty acids, total</td>
<td>1.9-4.5 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ferritin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- male</td>
<td>1.5-30 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- female</td>
<td>0.9-18 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>17-Hydroxyprogesterone, male</td>
<td>20-250 x 10⁻¹¹</td>
<td></td>
<td>Phagocytes: birth, range</td>
<td>6-26 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- female: follicular</td>
<td>20-80 x 10⁻¹¹</td>
<td></td>
<td>- birth, median</td>
<td>11 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- luteal</td>
<td>80-300 x 10⁻¹¹</td>
<td></td>
<td>- pediatric, range</td>
<td>1.5-8.5 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- postmenopausal</td>
<td>4-50 x 10⁻¹¹</td>
<td></td>
<td>- pediatric, median</td>
<td>4.1 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>17-Hydroxyprogesterone, child</td>
<td>20-140 x 10⁻¹¹</td>
<td></td>
<td>- adult, range</td>
<td>3.5-9.2 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>Immunoglobulin A (IgA)</td>
<td>5-39 x 10⁻⁴</td>
<td></td>
<td>- CD4 cell count</td>
<td>0.5-1.5 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>Immunoglobulin D (IgD)</td>
<td>0.5-8.0 x 10⁻⁵</td>
<td></td>
<td>Lipase P</td>
<td>1.2-1.4 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>Immunoglobulin G (IgG)</td>
<td>5.0-19 x 10⁻⁵</td>
<td></td>
<td>Lipids, total</td>
<td>4.45-6.1 x 10⁻³</td>
<td>4.8-5 x 10⁻⁴</td>
</tr>
<tr>
<td>Immunoglobulin M (IgM)</td>
<td>3.0-30 x 10⁻⁴</td>
<td></td>
<td>Lithium</td>
<td>1.5-2.5 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>Immunoglobulin E (IgE)</td>
<td>&lt;5 x 10⁻⁷</td>
<td></td>
<td>Lysine</td>
<td>1.3-3 x 10⁻⁵</td>
<td></td>
</tr>
<tr>
<td>Indican</td>
<td>8-50 x 10⁻⁶</td>
<td></td>
<td>Lysozyme (muramidase)</td>
<td>1-15 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>Inositol</td>
<td>3-7 x 10⁻⁶</td>
<td></td>
<td>α₂-Macroglobulin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insulin</td>
<td>2.0-8.4 x 10⁻¹⁰</td>
<td></td>
<td>- pediatric</td>
<td>2.7 x 10⁻³</td>
<td></td>
</tr>
<tr>
<td>Insulin-like growth factor I</td>
<td>9.9-50 x 10⁻⁸</td>
<td></td>
<td>- male, adult</td>
<td>0.9-4 x 10⁻³</td>
<td></td>
</tr>
<tr>
<td>Iodine, total</td>
<td>2.4-3.2 x 10⁻⁸</td>
<td>4.5-14.5 x 10⁻⁸</td>
<td>- female, adult</td>
<td>1.2-5.4 x 10⁻³</td>
<td></td>
</tr>
<tr>
<td>Iron, adult</td>
<td>4.6-10⁻⁴</td>
<td>6.18 x 10⁻⁷</td>
<td>Melatonin</td>
<td>1.35-1.45 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>Isocitric acid</td>
<td>9-15 x 10⁻⁶</td>
<td>1.2-4.2 x 10⁻⁵</td>
<td>- Day</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ketone bodies</td>
<td>2.3-10 x 10⁻⁶</td>
<td>1.5-30 x 10⁻⁶</td>
<td>- Night</td>
<td></td>
<td></td>
</tr>
<tr>
<td>α₂-Ketonic acids, adult</td>
<td>1-30 x 10⁻⁶</td>
<td></td>
<td>Mercury</td>
<td>6.07-7.13 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>L-Lactate</td>
<td>&lt;1.13 x 10⁻⁵</td>
<td>4.5-14.4 x 10⁻⁵</td>
<td>- normal</td>
<td>&lt;1 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>- arterial</td>
<td>8.1-15.3 x 10⁻⁵</td>
<td>4.5-19.8 x 10⁻⁵</td>
<td>- chronic</td>
<td>&gt;20 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>Lead</td>
<td>1-5 x 10⁻⁷</td>
<td>1.78 x 10⁻⁸</td>
<td>Methemoglobin</td>
<td>4.6 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>- normal</td>
<td>&gt;6-10 x 10⁻⁷</td>
<td></td>
<td>Methionine</td>
<td>4.6 x 10⁻⁶</td>
<td>1.15 x 10⁻⁶</td>
</tr>
<tr>
<td>Lecithin</td>
<td>1.1-1.2 x 10⁻³</td>
<td>1.2-25 x 10⁻³</td>
<td>β₂-microglobulin</td>
<td>8-24 x 10⁻⁷</td>
<td></td>
</tr>
<tr>
<td>Leptin</td>
<td>1.2 x 10⁻⁸</td>
<td></td>
<td>MIP-1α</td>
<td>2.3 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>Leucine</td>
<td>1.4-2 x 10⁻⁵</td>
<td>1.2-5.2 x 10⁻⁵</td>
<td>MIP-1β</td>
<td>9 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>Leukocytes (#/cm³)</td>
<td></td>
<td></td>
<td>Mucopolysaccharides</td>
<td>1.75-2.25 x 10⁻³</td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td></td>
<td>Mucoproteins</td>
<td>8.65-9.6 x 10⁻⁴</td>
<td></td>
</tr>
<tr>
<td>Neutrophils: birth</td>
<td>6-26.0 x 10⁻⁵</td>
<td></td>
<td>Nerve growth factor (NGF)</td>
<td>6-10 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>- pediatric</td>
<td>1.5-8.5 x 10⁻⁵</td>
<td></td>
<td>Nicotinic acid</td>
<td>5-8 x 10⁻⁶</td>
<td>2-15 x 10⁻⁷</td>
</tr>
<tr>
<td>- adult, range</td>
<td>4.3-11.0 x 10⁻⁵</td>
<td></td>
<td>Nitrogen</td>
<td>8.2 x 10⁻⁶</td>
<td>9.7 x 10⁻⁶</td>
</tr>
<tr>
<td>- median</td>
<td>7.0 x 10⁻⁶</td>
<td></td>
<td>- respiratory gas</td>
<td>3-3.7 x 10⁻²</td>
<td></td>
</tr>
<tr>
<td>Neutrophils: adult</td>
<td>3.65 x 10⁻⁶</td>
<td></td>
<td>Norepinephrine</td>
<td>2.15-4.75 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>- pediatric</td>
<td>0.4 x 10⁻⁶</td>
<td></td>
<td>- after 15 min rest</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult, range</td>
<td>0.2-0.3 x 10⁻⁶</td>
<td></td>
<td>- when emitted</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- median</td>
<td>0.05-0.7 x 10⁻⁶</td>
<td></td>
<td>8.1 x 10⁻⁹</td>
<td>8.5 x 10⁻⁹</td>
<td></td>
</tr>
<tr>
<td>Eosinophils: birth</td>
<td>0.4 x 10⁻⁶</td>
<td></td>
<td>Nucleotide, total</td>
<td>3.1-5.2 x 10⁻⁴</td>
<td>4-14 x 10⁻⁶</td>
</tr>
<tr>
<td>- pediatric</td>
<td>0.2-0.3 x 10⁻⁶</td>
<td></td>
<td>Ornithine</td>
<td>1-2.4 x 10⁻⁶</td>
<td></td>
</tr>
<tr>
<td>- adult, range</td>
<td>0.05-0.7 x 10⁻⁶</td>
<td></td>
<td>Oxalate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- median</td>
<td>0.15 x 10⁻⁶</td>
<td></td>
<td>Oxygen (respiratory gas)</td>
<td>2.4-3.2 x 10⁻⁴</td>
<td>3.9 x 10⁻⁶</td>
</tr>
<tr>
<td>Basophils: adult, range</td>
<td>0.015-0.15 x 10⁻⁶</td>
<td></td>
<td>- arterial</td>
<td>1.6-2.3 x 10⁻⁴</td>
<td>1.6 x 10⁻⁶</td>
</tr>
<tr>
<td>- median</td>
<td>0.03 x 10⁻⁶</td>
<td></td>
<td>Oxytocin</td>
<td>2 x 10⁻¹²</td>
<td>2 x 10⁻¹²</td>
</tr>
<tr>
<td>Lymphocytes: birth</td>
<td>2-11 x 10⁻⁶</td>
<td></td>
<td>- male</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pediatric</td>
<td>1.5-8.0 x 10⁻⁶</td>
<td></td>
<td>- female, nonlactating</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult, range</td>
<td>1.5-4.0 x 10⁻⁶</td>
<td></td>
<td>- female, pregnant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult, median</td>
<td>2.5 x 10⁻⁶</td>
<td></td>
<td>- pregnant 33-40 wks</td>
<td>32-48 x 10⁻¹²</td>
<td></td>
</tr>
<tr>
<td>Monocytes: birth, range</td>
<td>0.4-3.1 x 10⁻⁶</td>
<td></td>
<td>Pancreatic polypeptide</td>
<td>5-20 x 10⁻¹¹</td>
<td></td>
</tr>
<tr>
<td>- birth, median</td>
<td>1.05 x 10⁻⁶</td>
<td></td>
<td>Pantothentic acid</td>
<td>1.5-4.5 x 10⁻⁷</td>
<td>6-35 x 10⁻⁸</td>
</tr>
<tr>
<td>- pediatric</td>
<td>0.4 x 10⁻⁶</td>
<td></td>
<td>Para-aminobenzoic acid</td>
<td>3-4 x 10⁻⁸</td>
<td></td>
</tr>
<tr>
<td>- adult, range</td>
<td>0.21-1.05 x 10⁻⁶</td>
<td></td>
<td>Parathyroid hormone (PTH)</td>
<td>2-4 x 10⁻¹⁰</td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>In Whole Blood</td>
<td>In Plasma or Serum</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------</td>
<td>--------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(gm/cm³)</td>
<td>(gm/cm³)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pentose, phosphorated</td>
<td>2-2.3 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenol, free</td>
<td>7-10 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>8-12 x 10⁻⁶</td>
<td>1.1-4 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phospholipid</td>
<td>2.25-2.85 x 10⁻³</td>
<td>5-12 x 10⁻¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphatase, acid, prostatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphorus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- inorganic, adult</td>
<td>2-3.9 x 10⁻⁵</td>
<td>2.3-4.5 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- inorganic, children</td>
<td>4.0-7.0 x 10⁻⁵</td>
<td>1.5-10 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- total</td>
<td>3.5-4.3 x 10⁻⁴</td>
<td>1-1.5 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phytic acid</td>
<td>&lt;3 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelets (#/cm³)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- range</td>
<td>1.4-4.4 x 10⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- median</td>
<td>2.5 x 10⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Platelet Derived</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Growth Factor</td>
<td>5.0 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polysaccharides, total</td>
<td>7.3-13.1 x 10⁻⁴</td>
<td>1.4-2.2 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Potassium</td>
<td>1.6-2.4 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnenolone</td>
<td>3-20 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progesterone (male)</td>
<td>12-20 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progesterone (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- follicular</td>
<td>0.4-0.9 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- midluteal</td>
<td>7.7-12.1 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 16-18</td>
<td>30-66 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 28-30</td>
<td>70-126 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- pregnancy, weeks 38-40</td>
<td>131-227 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proinsulin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- fasting</td>
<td>0.5-5 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- mean</td>
<td>1.42-1.70 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prolactin (male)</td>
<td>&lt;20 x 10⁻⁹</td>
<td>1-7 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- while awake</td>
<td>9-20 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prolactin (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- follicular</td>
<td>&lt;23 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- luteal</td>
<td>5-40 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proline</td>
<td>1.2-5.7 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prostaglandins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- PGE</td>
<td>3.55-4.15 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- PGF</td>
<td>1.26-1.56 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 15-keto-PGF₂₅₆</td>
<td>5 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- 15-keto-PGF₂</td>
<td>&lt;5 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein, total</td>
<td>1.9-2.1 x 10⁻¹</td>
<td>6.0-8.3 x 10⁻²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protoporphyrin</td>
<td>2.7-6.1 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PSA (prostate-specific antigen)</td>
<td>0-5 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoglobulin I</td>
<td>8-19 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudoglobulin II</td>
<td>2-8 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purine, total</td>
<td>9.5-11.5 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyrimidine nucleotide</td>
<td>2.6-4.6 x 10⁻⁵</td>
<td>2-12 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyridoxine (Vitamin B₆)</td>
<td>3.6-90 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyruvic acid</td>
<td>3-10 x 10⁻⁶</td>
<td>3-12 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RANTES</td>
<td>7 x 10⁻¹¹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relxin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- day &lt;100 preparturition</td>
<td>&lt;2 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- day 100 to 2 days preceding</td>
<td>5-40 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- day preceding parturition</td>
<td>100-200 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- day following parturition</td>
<td>&lt;2 x 10⁻⁹</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Retinol (Vitamin A)</td>
<td>1-8 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Riboflavin (Vitamin B₂)</td>
<td>1.5-6 x 10⁻⁷</td>
<td>2.6-3.7 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA</td>
<td>5-8 x 10⁻⁴</td>
<td>4-6 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secretin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serine</td>
<td>2.9-4.5 x 10⁻¹¹</td>
<td>3-20 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serotonin (5-hydroxytryptamine)</td>
<td>1.55-1.81 x 10⁻⁷</td>
<td>0.8-2.1 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silicon</td>
<td>1.4-2.95 x 10⁻⁶</td>
<td>2.2-5.7 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Solids, total</td>
<td>2-2.5 x 10⁻¹</td>
<td>8-9 x 10⁻²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatotropin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(growth hormone)</td>
<td>4-140 x 10⁻¹⁰</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sphingomyelin</td>
<td>1.5-1.85 x 10⁻³</td>
<td>1-4 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Succinic acid</td>
<td></td>
<td>5 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sugar, total</td>
<td>7-11 x 10⁻⁴</td>
<td>8-12 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfates, inorganic</td>
<td></td>
<td>3.1-3.8 x 10⁻²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulfur, total</td>
<td>3.8-5 x 10⁻²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taurine</td>
<td></td>
<td>3-21 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testosterone (male)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- free</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testosterone (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- free</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Testosterone (female)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- free</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thiocyanate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- nonsmoker</td>
<td>1-4 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- smoker</td>
<td>3-12 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroglobulin (Tg)</td>
<td>1.3-2 x 10⁻⁵</td>
<td>0.9-3.2 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroid hormone</td>
<td>&lt;5 x 10⁻⁸</td>
<td>4-8 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyrotropin-releasing hormone</td>
<td>5-60 x 10⁻¹²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine (FT₄)</td>
<td>8-24 x 10⁻¹²</td>
<td>4-12 x 10⁻⁸</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine-binding prealbumin</td>
<td>2.8-3.5 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thyroxine-binding globulin</td>
<td>1.0-3.4 x 10⁻⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tin</td>
<td>0-4 x 10⁻⁷</td>
<td>5-20 x 10⁻⁶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α-Tocopherol (Vitamin E)</td>
<td>1.5-2 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transferrin</td>
<td>1.3-2.75 x 10⁻⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- newborn</td>
<td>2.2-4 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult</td>
<td>1.8-3.8 x 10⁻³</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- age &gt;60 yrs</td>
<td>2.5-30 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>2-4 x 10⁻⁴</td>
<td>2.8-4 x 10⁻⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
<td>Component</td>
<td>In Whole Blood (gm/cm³)</td>
<td>In Plasma or Serum (gm/cm³)</td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>-----------</td>
<td>------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Uric acid</td>
<td>6-49 x 10⁻⁶</td>
<td>2-7 x 10⁻⁵</td>
<td>Vasointestinal peptide (VIP)</td>
<td>6-16 x 10⁻¹²</td>
<td></td>
</tr>
<tr>
<td>- child</td>
<td></td>
<td>3.5-7.2 x 10⁻⁵</td>
<td>Vasopressin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- adult, male</td>
<td>2.6-6.0 x 10⁻⁵</td>
<td>- hydrated</td>
<td>- hydrated</td>
<td>4.5 x 10⁻¹³</td>
<td></td>
</tr>
<tr>
<td>- adult, female</td>
<td>2.0-5.5 x 10⁻⁵</td>
<td>- dehydrated</td>
<td>- dehydrated</td>
<td>3.7 x 10⁻¹²</td>
<td></td>
</tr>
<tr>
<td>Valine</td>
<td>2-2.9 x 10⁻⁵</td>
<td>1.7-4.2 x 10⁻⁵</td>
<td>Zinc</td>
<td>5-13 x 10⁻⁶</td>
<td>7-15 x 10⁻⁷</td>
</tr>
</tbody>
</table>
Appendix C. Catalog of Distinct Cell Types in the Adult Human Body

I. Keratinizing Epithelial Cells
   1. Epidermal keratinocyte (differentiating epidermal cell)
   2. Epidermal basal cell (stem cell)
   3. Keratinocyte of fingernails and toenails
   4. Nail bed basal cell (stem cell)
   5. Medullary hair shaft cell
   6. Cortical hair shaft cell
   7. Cuticular hair shaft cell
   8. Cuticular hair root sheath cell
   9. Hair root sheath cell of Huxley’s layer
  10. Hair root sheath cell of Henle’s layer
  11. External hair root sheath cell
  12. Hair matrix cell (stem cell)

II. Wet Stratified Barrier Epithelial Cells
   13. Surface epithelial cell of stratified squamous epithelium of cornea, tongue, oral cavity, esophagus, anal canal, distal urethra and vagina
   14. Basal cell (stem cell) of epithelia of cornea, tongue, oral cavity, esophagus, anal canal, distal urethra and vagina
   15. Urinary epithelium cell (lining bladder and urinary ducts)

III. Exocrine Secretory Epithelial Cells
   16. Salivary gland mucous cell (polysaccharide-rich secretion)
   17. Salivary gland serous cell (glycoprotein enzyme-rich secretion)
   18. Von Ebner’s gland cell in tongue (washes taste buds)
   19. Mammary gland cell (milk secretion)
   20. Lacrimal gland cell (tear secretion)
   21. Ceruminous gland cell in ear (wax secretion)
   22. Eccrine sweat gland dark cell (glycoprotein secretion)
   23. Eccrine sweat gland clear cell (small molecule secretion)
   24. Apocrine sweat gland cell (odoriferous secretion, sex-hormone sensitive)
   25. Gland of Moll cell in eyelid (specialized sweat gland)
   26. Sebaceous gland cell (lipid-rich sebum secretion)
   27. Bowman’s gland cell in nose (washes olfactory epithelium)
   28. Brunner’s gland cell in duodenum (enzymes and alkaline mucus)
   29. Seminal vesicle cell (secretes seminal fluid components, including fructose for swimming sperm)
   30. Prostate gland cell (secretes seminal fluid components)
   31. Bulbourethral gland cell (mucus secretion)
   32. Bartholin’s gland cell (vaginal lubricant secretion)
   33. Gland of Littre cell (mucus secretion)
   34. Uterus endometrium cell (carbohydrate secretion)
   35. Isolated goblet cell of respiratory and digestive tracts (mucus secretion)
   36. Stomach lining mucous cell (mucus secretion)
   37. Gastric gland zymogenic cell (pepsinogen secretion)
   38. Gastric gland oxyntic cell (HCl secretion)
   39. Pancreatic acinar cell (bicarbonate and digestive enzyme secretion)
   40. Paneth cell of small intestine (lysozyme secretion)
   41. Type II pneumocyte of lung (surfactant secretion)
   42. Clara cell of lung

IV. Hormone Secreting Cells
   43. Anterior pituitary cell secreting growth hormone
   44. Anterior pituitary cell secreting follicle-stimulating hormone
   45. Anterior pituitary cell secreting luteinizing hormone
   46. Anterior pituitary cell secreting prolactin
   47. Anterior pituitary cell secreting adrenocorticotropin hormone
   48. Anterior pituitary cell secreting thyroid-stimulating hormone
   49. Intermediate pituitary cell secreting melanocyte-stimulating hormone
   50. Posterior pituitary cell secreting oxytocin
   51. Posterior pituitary cell secreting vasopressin
   52. Gut and respiratory tract cell secreting serotonin
   53. Gut and respiratory tract cell secreting endorphin
   54. Gut and respiratory tract cell secreting somatostatin
   55. Gut and respiratory tract cell secreting gastrin
   56. Gut and respiratory tract cell secreting secretin
   57. Gut and respiratory tract cell secreting cholecystokinin
   58. Gut and respiratory tract cell secreting insulin
   59. Gut and respiratory tract cell secreting glucagon
   60. Gut and respiratory tract cell secreting bombesin
   61. Thyroid gland cell secreting thyroid hormone
   62. Thyroid gland cell secreting calcitonin
   63. Parathyroid gland cell secreting parathyroid hormone
   64. Parathyroid gland oxyphil cell
   65. Adrenal gland cell secreting epinephrine
   66. Adrenal gland cell secreting norepinephrine
   67. Adrenal gland cell secreting steroid hormones (mineralcorticoids and glucocorticoids)
68. Leydig cell of testes secreting testosterone
69. Theca interna cell of ovarian follicle secreting estrogen
70. Corpus luteum cell of ruptured ovarian follicle secreting progesterone
71. Kidney juxtaglomerular apparatus cell (renin secretion)
72. Macula densa cell of kidney
73. Peripolar cell of kidney
74. Mesangial cell of kidney

V. Epithelial Absorptive Cells (Gut, Exocrine Glands and Urogenital Tract)
75. Intestinal brush border cell (with microvilli)
76. Exocrine gland striated duct cell
77. Gall bladder epithelial cell
78. Kidney proximal tubule brush border cell
79. Kidney distal tubule cell
80. Ductulus efferens nonciliated cell
81. Epididymal principal cell
82. Epididymal basal cell

VI. Metabolism and Storage Cells
83. Hepatocyte (liver cell)
84. White adipocyte (fat cell)
85. Brown adipocyte (fat cell)
86. Liver lipocyte

VII. Barrier Function Cells (Lung, Gut, Exocrine Glands and Urogenital Tract)
87. Type I pneumocyte (lining air space of lung)
88. Pancreatic duct cell (centroacinar cell)
89. Nonstiated duct cell (of sweat gland, salivary gland, mammary gland, etc.)
90. Kidney glomerulus parietal cell
91. Kidney glomerulus podocyte
92. Loop of Henle thin segment cell (in kidney)
93. Kidney collecting duct cell
94. Duct cell (of seminal vesicle, prostate gland, etc.)

VIII. Epithelial Cells Lining Closed Internal Body Cavities
95. Blood vessel and lymphatic vascular endothelial fenestrated cell
96. Blood vessel and lymphatic vascular endothelial continuous cell
97. Blood vessel and lymphatic vascular endothelial splenic cell
98. Synovial cell (lining joint cavities, hyaluronic acid secretion)
99. Serosal cell (lining peritoneal, pleural, and pericardial cavities)
100. Squamous cell (lining perilymphatic space of ear)
101. Squamous cell (lining endolymphatic space of ear)
102. Columnar cell of endolymphatic sac with microvilli (lining endolymphatic space of ear)
103. Columnar cell of endolymphatic sac without microvilli (lining endolymphatic space of ear)
104. Dark cell (lining endolymphatic space of ear)
105. Vestibular membrane cell (lining endolymphatic space of ear)
106. Stria vascularis basal cell (lining endolymphatic space of ear)
107. Stria vascularis marginal cell (lining endolymphatic space of ear)
108. Cell of Claudius (lining endolymphatic space of ear)
109. Cell of Boettcher (lining endolymphatic space of ear)
110. Choroid plexus cell (cerebrospinal fluid secretion)
111. Pia-arachnoid squamous cell
112. Pigmented ciliary epithelium cell of eye
113. Nonpigmented ciliary epithelium cell of eye
114. Corneal endothelial cell

IX. Ciliated Cells with Propulsive Function
115. Respiratory tract ciliated cell
116. Oviduct ciliated cell (in female)
117. Uterine endometrial ciliated cell (in female)
118. Rete testis ciliated cell (in male)
119. Ductulus efferens ciliated cell (in male)
120. Ciliated ependymal cell of central nervous system (lining brain cavities)

X. Extracellular Matrix Secretion Cells
121. Ameloblast epithelial cell (tooth enamel secretion)
122. Planum semilunatum epithelial cell of vestibular apparatus of ear (proteoglycan secretion)
123. Organ of Corti interdental epithelial cell (secreting tectorial membrane covering hair cells)
124. Loose connective tissue fibroblasts
125. Corneal fibroblasts
126. Tendon fibroblasts
127. Bone marrow reticular tissue fibroblasts
128. Other (nonepithelial) fibroblasts
129. Blood capillary pericyte
130. Nucleus pulposus cell of intervertebral disc
131. Cementoblast/cementocyte (tooth root bonelike cementum secretion)
132. Odontoblast/odontocyte (tooth dentin secretion)
133. Hyaline cartilage chondrocyte
134. Fibrocartilage chondrocyte
135. Elastic cartilage chondrocyte
136. Osteoblast/osteocyte
137. Osteoprogenitor cell (stem cell)
138. Hyalocyte of vitreous body of eye
139. Stellate cell of perilymphatic space of ear

XI. Contractile Cells
140. Red skeletal muscle cell (slow)
141. White skeletal muscle cell (fast)
142. Intermediate skeletal muscle cell
143. Muscle spindle — nuclear bag cell
144. Muscle spindle — nuclear chain cell
145. Satellite cell (stem cell)
146. Ordinary heart muscle cell
147. Nodal heart muscle cell
148. Purkinje fiber cell
149. Smooth muscle cell (various types)
150. Myoepithelial cell of iris
151. Myoepithelial cell of exocrine glands
XII. Blood and Immune System Cells
152. Erythrocyte (red blood cell)
153. Megakaryocyte
154. Monocyte
155. Connective tissue macrophage (various types)
156. Epidermal Langerhans cell
157. Osteoclast (in bone)
158. Dendritic cell (in lymphoid tissues)
159. Microglial cell (in central nervous system)
160. Neutrophil
161. Eosinophil
162. Basophil
163. Mast cell
164. Helper T lymphocyte cell
165. Suppressor T lymphocyte cell
166. Killer T lymphocyte cell
167. IgM B lymphocyte cell
168. IgG B lymphocyte cell
169. IgA B lymphocyte cell
170. IgE B lymphocyte cell
171. Killer cell
172. Stem cells and committed progenitors for the blood and immune system (various types)

XIII. Sensory Transducer Cells
173. Photoreceptor rod cell of eye
174. Photoreceptor blue-sensitive cone cell of eye
175. Photoreceptor green-sensitive cone cell of eye
176. Photoreceptor red-sensitive cone cell of eye
177. Auditory inner hair cell of organ of Corti
178. Auditory outer hair cell of organ of Corti
179. Type I hair cell of vestibular apparatus of ear (acceleration and gravity)
180. Type II hair cell of vestibular apparatus of ear (acceleration and gravity)
181. Type I taste bud cell
182. Olfactory neuron
183. Basal cell of olfactory epithelium (stem cell for olfactory neurons)
184. Type I carotid body cell (blood pH sensor)
185. Type II carotid body cell (blood pH sensor)
186. Merkel cell of epidermis (touch sensor)
187. Touch-sensitive primary sensory neurons (various types)
188. Cold-sensitive primary sensory neurons
189. Heat-sensitive primary sensory neurons
190. Pain-sensitive primary sensory neurons (various types)
191. Proprioceptive primary sensory neurons (various types)

XIV. Autonomic Neuron Cells
192. Cholinergic neural cell (various types)
193. Adrenergic neural cell (various types)
194. Peptidergic neural cell (various types)

XV. Sense Organ and Peripheral Neuron Supporting Cells
195. Inner pillar cell of organ of Corti
196. Outer pillar cell of organ of Corti
197. Inner phalangeal cell of organ of Corti
198. Outer phalangeal cell of organ of Corti
199. Border cell of organ of Corti
200. Hensen cell of organ of Corti
201. Vestibular apparatus supporting cell
202. Type I taste bud supporting cell
203. Olfactory epithelium supporting cell
204. Schwann cell
205. Satellite cell (encapsulating peripheral nerve cell bodies)
206. Enteric glial cell

XVI. Central Nervous System Neurons and Glial Cells
207. Neuron cell (large variety of types, still poorly classified)
208. Astrocyte glial cell (various types)
209. Oligodendrocyte glial cell

XVII. Lens Cells
210. Anterior lens epithelial cell
211. Crystallin-containing lens fiber cell

XVIII. Pigment Cells
212. Melanocyte
213. Retinal pigmented epithelial cell

XIX. Germ Cells
214. Oogonium/oocyte
215. Spermatocyte
216. Spermatogonium cell (stem cell for spermatocyte)

XX. Nurse Cells
217. Ovarian follicle cell
218. Sertoli cell (in testis)
219. Thymus epithelial cell
This glossary includes words useful in the study of nanomedicine, mostly from the fields of physics, engineering, chemistry, biochemistry, biology, anatomy and medicine. Specialists are cautioned that terms may be described, but not rigorously defined, in order to encourage the quickest possible cross-disciplinary understanding by the nonspecialist reader. A very small number of the entries identify new terms introduced for the first time in this text. Much of the material was compiled or modified from pre-existing sources including the following References: 9, 10, 750, 869, 996, 997, 1095, 1259, 1736, 2219-2221, 2223, 2224.

**Ab initio** — from the beginning; from first principles.

**Abscess** — a circumscribed collection of pus appearing in acute or chronic localized infection, and associated with tissue destruction and frequently with swelling; a cavity formed by liquefaction necrosis within solid tissue.

**Abstraction** — specific removal of bonded atom from a structure. A reaction that removes an atom from a structure.

**AC** — alternating current.

**Accommodation** — voluntarily thickening the lens of the eye to focus diverging rays of light from nearby objects on the fovea of the retina; visual focusing on nearby objects.

**Acellular** — cell-free.

**Acetylcholine** — a chemical neurotransmitter.

**Acetylcholinesterase** — enzyme that rapidly degrades acetylcholine.

**Acidophilic** — acid-loving.

**Acoustomechanical conversion** — conversion of acoustic energy into mechanical energy.

**Action potential** — complete sequence of electrical events accompanying and following the nerve impulse.

**Active site** — the restricted part of a protein to which a substrate binds.

**Adduct** — in chemistry, an addition product or complex; in biomechanics, to draw together physically separated components.

**Adiabatic** — change in pressure or volume without loss or gain of heat.

**Adipocyte** — fat cell.

**Adipose** — fatty; pertaining to fat.

**ADP** — adenosine diphosphate; has one energy-rich phosphate bond.

**Adrenergic** — activated or energized by adrenalin (epinephrine).

**Adsorption** — attachment of a substance to the surface of another material.

**Aerobic** — needing oxygen to live or function.

**Aerobots (aerobotics)** — aerial (flying) robots.

**Afferent** — in relation to nerves or blood vessels, conducting toward structure or organ; carrying impulses toward a center, as when sensory nerves carry sensory information toward the brain or spinal cord.

**Affinity (constant)** — the strength of the binding of a ligand to a receptor, or the reciprocal of the dissociation rate constant; a measure of the binding energy of a ligand in a receptor; the greater the affinity, the more securely the receptor binds the ligand.

**AFM** — see Atomic Force Microscope.

**Aft** — toward the rear.

**Agglutinin** — an antibody present in the blood that attaches to antigens, such as those found on the surfaces of red blood cells, causing them to clump together; agglutinins cause transfusion reactions when blood from a different group is given.

**Agonist** — in pharmacology, a drug which binds to a receptor and thus stimulates the receptor’s function, possibly mimicking the body’s own regulatory function. Compare antagonist.

**Albumin** — one of a group of simple proteins widely distributed in plant and animal tissues.

**ALC** — airborne lethal concentration.

**Algorithm** — in general, a formula or set of rules for solving a particular problem; in medicine, a set of steps used in diagnosing and treating a disease.

**Alimentary** — pertaining to the digestive tract.
Alimentography — a physical description (and mapping) of the human alimentary canal.

Aliquot — a portion obtained by dividing the whole into equal parts without a remainder; loosely, any one of two or more samples of something, of the same volume or weight.

Alkali — strongly basic substance, especially the metal hydroxides, usually associated with the alkali metals (e.g. sodium and potassium).

Allele — one of several alternative forms of a gene occupying a given locus on paired chromosomes.

Alloantigen — a substance present in certain individuals that stimulates antibody production in other members of the same species, but not in the original donor. See also antiserum.

Allometric scaling laws — in biology, scaling laws that involve a biological variable that is an exponential function of the mass of the organism.

Allosteric control — the ability of an interaction at one site of a protein to influence the activity of another site.

Allotropic — pertains to the existence of a chemical element or compound in two or more distinct forms with different physical and chemical properties (e.g. diamond and graphite are allotropes of C).

Alphanumeric — able to contain both alphabetic and numeric characters.

Alveolus (alveolar) — in anatomy, a small cell or cavity; a saclike dilation. Most commonly, a small air sac found at the lowest levels of the branching tube system comprising the lungs.

AM — amplitude modulation.

Amide — a molecule containing an amine bonded to a carboxyl group (e.g. CONH$_2$). Amide bonds link amino acids in peptides and proteins.

Amine — a molecule containing N with a single bond to C and two other single bonds to H or C (but not an amide); the amine group or moiety (e.g. -NH$_2$).

Amino acid — a molecule containing both an amine and a carboxylic acid group; there are 20 genetically encoded amino acids in biology.

Amniotic — pertaining to the amnion (the innermost of the fetal membranes).

Amphipathic — molecular structures which have two surfaces or ends, one of which is hydrophilic and the other of which is hydrophobic. Lipids are amphipathic, and some protein regions may form amphipathic helices with one charged face and one neutral face.

Anabolism — the constructive phase of metabolism and the opposite of catabolism; in anabolism, a cell takes from the blood the substances required for repair and growth, building them into a cytoplasm, thus converting nonliving material into the living cytoplasm of the cell.

Anaerobic — able to live or function without oxygen.

Analgesia — absence of normal sense of pain.

Analog — pertaining to data measurable and representable through continuously variable physical quantities. Compare digital.

Anaphase — the phase of mitosis (cell division) beginning with centromere division and the movement of chromosomes away from the metaphase plate toward opposite spindle poles.

Anaphylactoid-type reaction — a physiological response similar to anaphylaxis.

Anaphylaxis — the immediate transient kind of immunologic (allergic) reaction characterized by contraction of smooth muscle and dilation of capillaries due to release of pharmacologically active substances (e.g. histamine, bradykinin, serotonin, etc.); a powerful allergic response. Anaphylaxis is classically initiated by the combination of antigen (allergen) with mast cell-fixed, cytophilic antibody (chiefly IgE immunoglobulin), but can also be initiated by relatively large quantities of serum aggregates (antibody-antigen complexes, and other) that seemingly activate complement leading to production of anaphylatoxin.

Anastomose — to open one structure into another directly or by connecting channels, usually said of blood vessels, lymphatics, and hollow viscera; to unite by means of an anastomosis, or a connection between formerly separate structures.

AND gate — a logical gate that returns a high (1) output if and only if both input signals are high (1). See bit.

Anergic — unresponsive.

Aneutronic — without neutrons.

Angioedema — a condition characterized by development of urticaria (hives) and edematous (swollen with excessive fluid) areas of skin, mucous membranes, or viscera.

Angiogenesis — growth of new blood vessels, especially capillaries.

Anion — a negatively charged ion.

Anisotropic — not isotropic.

Anode — the positive pole of an electrical source.

ANS — see autonomic nervous system.

Antagonist — in pharmacology, a drug that prevents receptor function. Compare agonist.

Anterior — the front of the human body, on or nearest the abdominal surface; the front of something.

Anteroinferior — in front and below.
Anteroposterior — passing from front to rear.

Anthropogenic — caused by human activity.

Antibody — a protein (immunoglobulin) produced by B-lymphocyte cells that recognizes a particular foreign antigen, thus triggering the immune response.

Antigen — any molecule or foreign substance that, when introduced into the body, provokes synthesis of an antibody, thus stimulating an immune response.

Antiserum — serum that contains demonstrable antibody or antibodies specific for one (monovalent) or more (polyvalent) antigens.

Aorta — the largest artery in the human body, leading away from the heart.

Apheresis — removal of blood from an individual patient, separating certain elements (e.g. red cells, platelets, white cells) for use elsewhere, and reintroducing the remaining components into the patients; also known as cytopheresis, hemapheresis, leukapheresis, pheresis, and plasmapheresis, depending on the type of cells being harvested.

Apical — pertaining to the apex (e.g. the point of a cone) of a structure.

Apoptosis — an orderly disintegration of eukaryotic cells into membrane-bound particles that may then be phagocytosed by other cells.

Aqueous humor — transparent liquid contained in the anterior chamber of the eyeball in front of the lens.

Aromatic compounds — in chemistry, ring or cyclic compounds related to benzene, many having a fragrant odor.

Arrhythmia — irregularity or loss of rhythm, especially of the heartbeat.

Arteriovenous — relating to both arteries and veins.

Artery — in anatomy, a blood vessel that sends blood to the tissues from the heart.

Aseptic — characterized by the absence of living pathogenic organisms; a state of sterility.

Asperities — protruding elements of roughness on a surface, e.g., burrs or spurs.

Asphyxia — condition of hypoxia caused by insufficient oxygen intake.

Assembler — see molecular assembler.

Asymptotic — in geometry and mathematics, a curve that approaches closer and closer, but never quite reaches, another curve or line.

Asynchronous — not synchronized in time.

atm — atmosphere, a unit of pressure; mean air pressure at Earth’s surface is 1 atmosphere (~1.01 x 10^5 N/m^2).

Atomic Force Microscope (AFM) — an instrument that uses atomic forces between a sample and a sharp scanning needle tip to image surfaces to molecular accuracy by mechanically probing their surface contours; the AFM measures the tiny upward and downward motions of the tip as the tip is dragged over the surface, producing an atomic-resolution topographic map of the surface. The AFM has also been used to physically manipulate individual molecules.

Atom laser — a laserlike device that uses beams of coherent atoms rather than photons.

ATP — adenosine triphosphate; has two energy-rich phosphate bonds.

Auscultation — the process of listening for sounds within the body, usually to the sounds of thoracic or abdominal viscera, in order to detect some abnormal condition, or to detect fetal heart sounds, or to diagnose vascular abnormalities such as arteriovenous fistulas.

Autogenous control — in medical nanorobotics, the conscious control of in vivo nanorobotic systems by the human user or patient; in biochemistry, the action of a gene product that either inhibits (negative autogenous control) or activates (positive autogenous control) expression of the gene coding for it.

Autonomic (nervous system) — the part of the nervous system that is concerned with the control of involuntary bodily functions.

Automated engineering — Engineering design done by a computer system, with detailed designs generated from broad specifications with little or no human help, a specialized form of artificial intelligence.

Automated manufacturing — molecular manufacturing that requires little human labor.

Autosomes — all the chromosomes except the sex chromosomes; a diploid cell has two copies of each autosome.

Avascular — having no blood vessels.

Avulsion — a tearing away forcibly of a part or structure.

Axisymmetric — symmetric relative to a geometric axis.

Axon — the (usually long and straight) protoplasmic process of a neuron that conducts impulses away from the cell body. There is usually one to a neuron cell.

Axoplasm — the cytoplasm (neuroplasm) of an axon that encloses the neurofibrils.

Bacterium — a single-celled prokaryotic microorganism, typically ~1-10 micron (~1000 nm) in diameter or length.

Bacteriophages — viruses that infect selected bacteria; often abbreviated as phages.

Barographics — pressure mapping of the human body.
Baronatation — in medical nanorobotics, locomotion through a frozen fluid by applying mechanical pressure along the path travelled to induce melting ahead, followed by regelation (refreezing) behind.

Basal lamina — basement lamina, e.g. basement membrane.

Basal Metabolic Rate (BMR) — a measure of the metabolic rate taken with the patient fasting and at rest. The oxygen consumed in breathing under these conditions indicates the minimum rate of chemical reactions in the body.

Basement membrane (basement lamina) — a thin layer of delicate noncellular material of a fine filamentous texture underlying the epithelium; its principal component is collagen.

Base pair (bp) — a complementary purine-pyrimidine hydrogen-bonded residue pair, one from each strand of DNA double helix, designating one unit (bp) of sequence. A partnership of adenine (A) with thymine (T) or of cytosine (C) with guanine (G) in a DNA double helix; other pairs can be formed in RNA under certain circumstances.

Basophil — a type of granulocytic white blood cell comprising less than 1% of all leukocytes, that is essential to the nonspecific immune response to inflammation because of its important role in releasing histamine and other chemicals that act on blood vessels.

Bearing — A mechanical device that permits the motion of a component (ideally, with minimal resistance) in one or more degrees of freedom while resisting motion (ideally, with a stiff restoring force) in all other degrees of freedom.

Beriberi — a thiamine deficiency disease, characterized by peripheral neurologic, cerebral, and cardiovascular abnormalities; early deficiency produces fatigue, irritation, poor memory, sleep disturbances, precordial pain, anorexia, abdominal discomfort, and constipation.

Biaxial — pertaining to two distinct geometric or spatial axes.

Bicuspid — having two cusps.

Bifurcate — to separate into two separate branches.

Bilobate — having two lobes.

Billion — this book follows the American convention in which a billion is $10^9$.

Bimorph — a simple mechanical actuator capable of flexing in a single plane.

Binding — the process by which a molecule (or ligand) becomes bound, that is, confined in position (and often orientation) with respect to a receptor. Confinement occurs because structural features of the receptor create a potential well for the ligand; van der Waals and electrostatic interactions commonly contribute.

Binding energy — in chemistry, the reduction in the free energy of a system that occurs when a ligand binds to a receptor. In nuclear physics, the energy which binds neutrons and protons together in an atomic nucleus, which may be calculated from the decrease in mass that would occur if the appropriate numbers of neutrons and protons were brought together to form the nucleus in question; Fe$^{56}$ has the maximum binding energy of any nucleus. Generally used to describe the total energy required to remove something, or to take a system apart into its constituent particles.

Binding site — the active region of a receptor; any site at which a chemical species of interest tends to bind.

Biocompatibility — the ability of the body to tolerate the presence or implantation of foreign objects.

Bioinformatics — most generally, the study and compilation of the information content of biological materials, especially genetic materials.

Biomimetic — mimicking biology.

Biosensor — A device that senses and analyzes biological information (e.g. temperature, pressure, chemical content, etc.), commonly using a biological recognition mechanism combined with a physical transduction technique.

Biotechnology — The application of biological systems and organisms to technical and industrial processes, with production carried out using intact organisms (e.g. yeasts and bacteria) or natural substances from organisms (e.g. enzymes), or by modifying the genetic structure of organisms (generic engineering); most generally, the engineering of all biological systems, even completely artificial organic living systems, using biological instrumentalities.

Birefringence — in optics, a crystal having a different index of refraction for different polarizations of incident light (e.g. calcite).

Bit — a binary digit. In a binary code, one of the two possible characters, usually 0 (e.g. low voltage) and 1 (e.g. high voltage); in information theory, the fundamental unit of information.

Blackbody radiator — an absorber which ideally absorbs all electromagnetic energy incident upon it; such an ideal absorber is also an ideal emitter of such radiant energy.

Blepharitis — inflammation of the edges of the eyelids involving hair follicles and glands that open onto the surface.

Blockade — prevention of the action of something, such as the effect of a drug or of a body function (e.g. halting immune system blood cleansing, by overloading the RES).

B-lymphocytes (B cells) — thymus-independent white blood cells responsible for synthesizing antibodies.

Boolean functions — see logic gate.

Bose-Einstein condensate — low-temperature collection of bosons (particles with integral spin, e.g. deuterium nuclei and alpha particles) in their ground state.

Bottom-up — an approach to nanotechnology that aims to construct nanodevices molecule by molecule, making larger and larger objects with atomic precision.

Boutons terminaux — bulblike expansions at the tips of axons that come into synaptic contact with the cell bodies of other neurons.
bp — see base pair.

Brachiation — locomotion by alternately swinging the arms (e.g. swinging hand-over-hand).

Bradycardia — slowness of heartbeat or heart action.

Bronchial — pertaining to the lungs.

Bronchiography — a physical description (and mapping) of the human bronchial system.

Brownian assembly — Brownian motion in a fluid brings molecules together in various positions and orientations; if molecules have suitable complementary surfaces, they can bind, assembling to form a specific structure. The term also may apply to the self-assembly of nanoscale parts having complementary surfaces.

Brownian motion — random motion of small particles in a fluid owing to thermal agitation, first observed in 1827 by Robert Brown; originally attributed to a vital force, Brownian motion plays a vital role in the assembly and activity of the molecular structures of life.

Bruit — an adventitious (arising sporadically) sound of venous or arterial origin, heard upon auscultation, such as murmurs at the carotid (a major artery) bifurcation.

Bubonic plague — the usual form of plague; primarily a disease of rodents that is transmitted to humans by fleas that have bitten infected animals.

Buccal — pertaining to the cheeks (mouth).

Buckyballs — ball-like molecules of fullerene carbon, C_{60}.

Bulimia — excessive and insatiable appetite for food.

Bulk technology — technology in which atoms and molecules are manipulated in bulk, rather than individually.

Bumpers — in medical nanorobotics, expansible surfaces placed at interfaces between adjacent nanorobots to achieve a tight seal between nanodevices.

Byte — a sequence of N adjacent bits operated on as a unit for the sake of convenience and frequently used as a measure of memory or information, with one byte corresponding to a single alphabetic character or some other symbol; N may equal 8, 16, 32, 64, or 128, in different 20th century computers.

c (velocity) — the speed of light, in vacuo.

CAD — Computer-Aided Design.

Caliber — inside diameter of a tube or cylinder.

Calmodulin — a 17,000-dalton protein that binds calcium ions in eukaryotic cells, thereby becoming the agent for many or most of the cellular effects ascribed to calcium ions.

CAM — Computer-Aided Manufacturing.

Cam — in mechanical engineering, a mechanical component that translates or rotates to move a contoured surface past a follower; the contours impose a sequence of motions (potentially complex) on the follower. See follower.

cAMP — cyclic AMP (adenosine monophosphate), an intracellular messenger molecule.

Capillary — in anatomy, a thin-walled tiny blood vessel, averaging 8 microns in diameter.

Capsid — the external protein coat of a virus particle.

Carbohydrates — a group of chemical substances, including sugars, glycogen, starches, dextrins, and celluloses, that contain only C, H, and O; usually the ratio of H:O is 2:1. Glucose and its polymers (including starch and cellulose) represent the most abundant organic chemical compounds on Earth.

Carbonyl — a chemical moiety consisting of O with a double bond to C (e.g. -CO-). If the C is bonded to N, the resulting structure is termed an “amide”; if it is bonded to O, it is termed a carboxylic acid or an ester linkage.

Carboxylic acid — a molecule that includes a C having a double bond to O and a single bond to OH (e.g. -COOH).

Carbyne — a chain of carbon atoms with alternating single and triple bonds.

Cardiac — pertaining to the heart.

Cardiostasis — pertaining to a heart that has stopped beating.

Carnot efficiency — in thermodynamics, the energy efficiency of a heat engine (a device that converts thermal energy into mechanical energy).

Cartotaxis — in medical nanorobotics, in vivo nanorobotic navigation by following a series of landmarks recorded on a previously drawn map.

Cartilage — specialized type of dense connective tissue consisting of cells embedded in a firm, compact fibrous collagenous matrix.

Catabolism — the destructive phase of metabolism, the opposite of anabolism. Catabolism includes all processes in which complex substances are converted into simpler substances (the end products commonly being excreted), usually with the release of energy.

Catalysis — increase in the speed of a chemical reaction or process produced by the presence of a catalyst (a substance that is not consumed in the net chemical reaction or process).

Catalyst — a chemical species or other structure that facilitates a chemical reaction without itself undergoing a permanent change.

Catacholamines — biologically active amines (e.g. epinephrine, norepinephrine) derived from the amino acid tyrosine, that have marked effects on the nervous and cardiovascular systems, on metabolic rate and temperature, and on smooth muscle.
Nanomedicine • Volume I

Cathode — the negative pole of an electrical source.

Cation — a positively charged ion.

Caudal (caudad, inferior) — away from the head or the lower part of a structure (literally means “toward the tail”).

Cauterize — to apply a cautery (an agent or device used for scar-ring, burring, or cutting the skin or tissues by means of heat, electrical current, or caustic chemicals).

Cavitation — in physics, the formation of bubbles in a fluid during high-power sonication of that fluid; in medicine, formation of a cavity by either normal or pathological biological processes.

Cell — a small biological structural unit, surrounded by a plasma membrane, making up living things.

Cell engineering — deliberate artificial modifications to biological cellular systems on a cell-by-cell basis.

Cell surgery — in medical nanorobotics, modifying cellular structures using medical nanomachines.

Cell typing — a method of identifying a cell’s type by comparing it to a typology of cell characteristics.

Cellular immunity — immunity resulting from activation of sensitized T-lymphocytes.

Cellular topographics — description of the cell and its parts.

Central Processing Unit (CPU) — the main computational engine of a computer, responsible for interpreting and executing instructions to process information.

Centrioles — small hollow cylinders consisting of microtubules, residing within the centrosomes, that become located near the poles during mitosis (cell division).

Centromere — a constricted region of a chromosome that divides the chromosome into two arms; the attachment point for the spindle fiber concerned with chromosome movement during mitosis (cell division).

Centrosomes — the regions from which microtubules are organized at the poles of a mitotic (dividing) cell; see also MTOC.

Centrum — middle part.

Cephalic (cephalad, superior) — nearest or toward the head or the upper part of a structure; also, pertaining to the head or top.

Cervical — usually pertaining to, or in the region of, the neck (i.e. below the skull); more generally, pertaining to the neck of an organ, such as the uterine cervix.

C fiber — an unmyelinated nerve fiber, 0.4-1.2 microns in diameter, conducting nerve impulses at 0.7-2.3 m/sec.

Chalcogenide — a glassy material containing arsenic and selenium.

Chaperone — a molecular chaperone is a protein needed for the assembly or proper folding of some other target protein, but which is not itself a component of the target protein complex.

Chemical force microscopy — an AFM with a functionalized tip, used to probe the chemical characteristics (especially adhesiveness) of surface-bound molecules.

Chemoergic — pertaining to the conversion of chemical energy into forms of energy of various other kinds.

Chemoblomography — chemical mapping of the human body.

Chemosurgery — conversion of chemical energy into mechanical energy.

Chimeric — pertaining to chemotaxis.

Chemosurgery — in medical nanorobotics, a nanosensor used to determine the chemical characteristics of surfaces, possibly configured as a pad coated with an array of reversible, perhaps reconfigurable, artificial molecular receptors.

Chemosurgery — the movement of additional white blood cells to an area of inflammation in response to the release of chemical mediators by neutrophils, monocytes, or injured tissue.

Chiral — a chiral molecule is an asymmetric molecule that cannot be superimposed on its mirror image. The molecule has two forms, or isomers, called enantiomers, that are mirror images of each other, the solutions of which rotate the plane of polarized light in different directions, left (levo) or right (dextro), making the enantiomers levorotatory or dextrorotatory. Only levorotatory amino acids are present in biological systems.

Chirurgeon — surgeon (obsolete).

Cholinergic — activated or energized by acetylcholine.

CHON — Carbon (C), Hydrogen (H), Oxygen (O), and Nitrogen (N).

Chonic — congenital absence of the choroid of the eye; progressive degeneration of the choroid, leading to complete blindness, due to X-linked chromosome inheritance.

Chromatin — the complex of DNA and protein in the nucleus of the interphase eukaryotic cell; individual chromosomes cannot be distinguished in it.

Chromomorphic — in medical nanorobotics, a nanorobot capable of altering its external surface color or other external surface optical characteristics.

Chromophore — a color-producing chemical substance.

Chromosome — a discrete unit of the genome carrying many genes; a structure in the nucleus of a eukaryotic cell containing a long
linear molecule of duplex DNA (and an approximately equal mass of proteins) which conveys genetic information.

**Chronobiology** — that aspect of biology concerned with the timing of biological events, especially repetitive or cyclic phenomena in individual organisms; the study of biological clocks.

**Chronocyte** — in individual organisms; the study of biological clocks.

**Chronometer** — a clock or other timekeeping device.

**Ciliated** — in anatomy, hairlike processes projecting from epithelial cells, as in the bronchi, for propelling mucus, pus, and dust particles, or in the inner ear, for sensing fluid motion to allow hearing; in microbiology, a means of manipulation and propulsion for motile microorganisms.

**Circadian** — pertaining to physiological events that occur at approximately 24-hour intervals.

**Circumcorporeal** — around or surrounding the human body.

**Circumvascular** — around or surrounding a blood vessel.

**Cisterna** — a reservoir or cavity.

**Clinical** — founded on actual observation and treatment of patients, as distinguished from data or facts obtained by experimentation or pathology; pertaining to a clinic.

**Clinicopathological correlation** — correlating (1) the signs and symptoms manifested by a patient, plus the results of any laboratory studies, with (2) the findings of the gross and histological examination of the patient’s tissue, to arrive at a correct diagnosis.

**Clyster** — an enema (obsolete).

**Cochlea** — the coiled, fluid-filled structure of the inner ear that transduces sound, allowing hearing.

**Codon** — in DNA and RNA, a sequence of three nucleotidyl residues designating an amino acid to be placed in polypeptide sequence during translation.

**Coenzyme** — a substance that enhances or is necessary for the action of enzymes; coenzymes are of smaller molecular size than the enzymes themselves, are dialyzable and relatively heat-stable, and are usually easily dissociable from the protein portion of the enzyme. See also vitamins.

**Collagen** — the major protein of the white fibers of connective tissue, cartilage, and bone, rich in glycine, alanine, proline, and hydroxyproline (amino acids), low in sulfur, and completely lacking in tryptophan (another amino acid); the collagen family comprises ~25% of all mammalian protein.

**Colloid** — see macromolecule.

**Communion** — breaking into fragments.
Cytoskeleton — the internal structural framework of a cell consisting of at least three types of filaments (microfilaments, microtubules, and intermediate filaments), forming a dynamic framework for maintaining cell shape and motion and allowing rapid changes in the three-dimensional structure of the cell.

Cytosol — the liquid that fills all of the cytoplasmic region, except for fluids filling the interior of the cell organelles.

Cytotomography — tomographic imaging of an individual cell.

Cytotoxic — tending to kill cells.

Cytovehicle — in medical nanorobotics, a living cell that has been commandeered by a medical nanorobot for use during cytocarriage.

Dalton — unit of molecular weight (1 dalton = 1 proton).

DC — direct current.

Decoction — liquor extract, from vegetable matter boiled in water.

Deglutition — swallowing.

Degrees of Freedom (DOF) — distinct axes of allowed mechanical motion; the number of directions that a mechanism is free to translate or rotate.

Demarcation — in medical nanorobotics, a crude form of functional navigation in which artificial conditions detectable by in vivo medical nanorobots are created at or near the target treatment site, such as warm or cold spots, pressure spots, or injected chemical plumes.

Denaturation — conversion of a protein from the physiological conformation to some other (possibly inactive) conformation.

Dendrimers — large, regularly-branching molecules.

Dendrite — a branched protoplasmic process of a neuron that conducts impulses toward the cell body. There are usually many to a cell, forming synaptic connections with other neurons.

De novo — created anew.

Deoxyribonucleic acid (DNA) — a complex molecule of very high molecular weight encoding genetic information. DNA consists of deoxyribose (a sugar), phosphoric acid, and four bases (purines or pyrimidines), arranged as two long chains that twist around each other to form a double helix joined by bonds between the complementary purine and pyrimidine components (analogous to rungs on a twisted ladder). DNA is present in the chromosomes of all cells and is the chemical basis of heredity and the carrier of genetic information for almost all organisms (e.g. except the RNA virus, etc.).

Dermatoglyphics — the study of dermal ridge patterns of fingers, toes, palms and soles.

Dermis — inner layer of the skin that lies below the epidermis.

Desiccate — removal of water; dehydration.

Desquamation — shedding epidermis in scales or shreds.

Diagnosis — the determination of the cause and nature of a disease, in order to provide a logical basis for treatment and prognosis.

Dialysis — the passage of a solute through a membrane; process of diffusing blood across a semipermeable membrane to remove toxic materials and to maintain fluid, electrolyte, and acid-base balance in cases of impaired kidney function.

Diametral — pertaining to a diameter.

Diamondoid — structures that resemble diamond in a broad sense; strong, stiff structures containing dense, three-dimensional networks of covalent bonds, formed chiefly from first and second row atoms with a valence of three or more. Many of the most useful diamondoid structures will be rich in tetrahedrally coordinated carbon.

Diamondophagy — eating diamond.

Diapedesis — transendothelial migration (passing through blood vessel endothelial coated walls) to exit the bloodstream and enter the surrounding tissues.

Diaphragm — in human anatomy, the musculomembranous partition between the abdominal and thoracic cavities, whose motions induce lung movements, enabling respiration.

Diaphysis — the shaft or middle part of a long cylindrical bone.

Diastereomeric — having two stereoisomers.

Diastole — the normal period in the heart cycle during which the muscle fibers loosen and lengthen, the heart dilates, and the cavities fill with blood; roughly, the period of relaxation alternating with systole or contraction.

Diathermy — local elevation of temperature within the tissues, produced by high-frequency (~MHz) current, ultrasonic waves, or microwave radiation.

Dielectric — a material capable of storing electrical energy; the interior of a capacitor.

Diels-Alder cycloadDITION — in chemical synthesis, the cycloadDITION of a conjugated diolefin (an olefin or alkene is a hydrocarbon possessing one or more double bonds in the carbon chain). Diels-Alder reactions are highly stereospecific.

Differentiation — acquisition of character or functions that are different from those of the original type; specialization of cell type within a cell line of increasingly specialized types, by a change in physical form of a cell.

Diffusion — a process by which populations of molecules intermingle and become mixed as a result of their incessant thermal motions.

Digital — pertaining to the use of combinations of bits to represent all quantities that arise during a problem or computation. Compare analog.
**Dimer** — in chemistry, a combination of two similar or identical molecules, usually by elimination of H₂O or a similar small molecule between the two, or by simple noncovalent association.

**Diode** — an electronic device that allows current to flow in only one direction.

**Diploe** — in anatomy, spongy tissue that lies between two layers of the compact bone of the skull.

**Diploid** — a set of chromosomes that contains two copies of each autosome plus two sex chromosomes; generally found in the somatic body cells.

**Dipole** — two equal and opposite charges separated by a distance. See also electric dipole moment.

**Disassembler** — in molecular nanotechnology, a nanomachine or system of nanomachines able to take an object apart while at each step recording the structure and composition of that object at the molecular level.

**Disassimilation** — in biology, changing assimilated material into less complex compounds for the production of energy.

**Disease** — see volitional normative model of disease.

**Disequilibration** — in medical nanorobotics, maintenance or inducement of a state of perpetual ionic, chemical, or energetic disequilibrium in a living cell by a medical nanorobot, usually for the purpose of inducing cytocide.

**Dispersion force (London dispersion force)** — see van der Waals forces.

**Dissociation** — in chemistry, the separation of a molecular complex into simpler molecules due to lytic reaction, heat, or ionization.

**Dissociation constant** — For systems in which ligands of a particular kind bind to a receptor in a solvent, there will be a characteristic frequency with which existing ligand-receptor complexes dissociate as a result of thermal excitation, and a characteristic frequency with which empty receptors bind ligands as a result of Brownian encounters, forming new complexes. The frequency of binding is proportional to the concentration of the ligand in solution. The dissociation constant is the magnitude of the ligand concentration at which the probability that the receptor will be found occupied is 0.5.

**Distal** — away from a source or a point of attachment or origin; in the extremities, farthest from the trunk.

**Diurnal** — daily.

**DNA** — see deoxyribonucleic acid.

**DNase** — an enzyme that attacks bonds in DNA.

**DNA polymerase** — an enzyme that synthesizes a daughter strand of DNA under direction from a DNA template; may be involved in repair or replication.

**DOF** — see Degrees of Freedom.
(low electronegativity) whereas fluorine (F) tends to withdraw electron density (high electronegativity); nitrogen (N) and oxygen (O) are also electronegative atoms.

**Electrophoresis** — the movement of charged colloidal particles through the medium in which they are dispersed as a result of changes in electrical potential; used in the analysis of protein mixtures because protein particles move with different characteristic velocities dependent principally on the number of charges carried by each particle.

**Electroporation** — insertion of macromolecules (e.g. DNA) into cells by employing a brief intense pulse of electricity to open cellular pores.

**Electrostatic Force Microscope (EFM)** — a kind of SPM that images electrostatic forces.

**Electrostatic Force** — a force arising between charged bodies produced by electrostatic fields.

**Embolus** — a mass of undissolved matter (solid, liquid, or gaseous) present in a blood or lymphatic vessel, brought there by the blood or lymph current.

**Emesis** — vomiting; may be chemically induced using an emetic.

**emf** — in physics, “electromotive force” (typically, the electrical potential established across the terminals of a battery), measured in volts.

**Emissivity** — relative intensity of emission of electromagnetic radiation by a heated body, as compared to a blackbody radiator whose emissivity is defined as 1.

**EMP** — electromagnetic pulse (e.g. as typically generated during a nuclear explosion).

**Enantiomer** — in chemistry, a special class of stereoisomers whose structure is not superimposable on its mirror image; a chiral molecule. Two enantiomers of a compound have identical chemical properties, but differ in a characteristic physical property, the ability to rotate the plane of plane-polarized light.

**Endocrine gland** — a gland that secretes biochemical substances (especially hormones) directly into the bloodstream.

**Endocytosis** — a process by which proteins arriving at the surface of a cell are internalized, being transported inside the cell within membranous vesicles.

**Endogenous** — originating inside an organ, part, or system.

**Endohedral** — lying entirely within a (fullerene) cage molecule.

**Endometrial** — pertaining to the lining of the uterus.

**Endoplasmic reticulum** — in cell biology, a highly convoluted sheet of membranes, extending from the outer layer of the nuclear envelope into the cytoplasm.

**Endosome** — the vacuole formed when material is absorbed into a cell by the process of endocytosis; the vacuole fuses with lysosomes.

**Endothelium** — a form of squamous epithelium consisting of flat cells that line the blood and lymphatic vessels, the heart, and various other body cavities.

**Endotheliocyte** — in medical nanorobotics, a theorized (nanorobotic) device capable of performing repairs of an injured vascular luminal surface.

**Endothermic** — a transformation that absorbs energy in the form of heat. A typical endothermic reaction increases both entropy and molecular potential energy, and thus is analogous to a gas expanding while absorbing heat and compressing a spring. Opposite of exothermic.

**Energy** — in physics, a conserved quantity that can be interconverted among many forms, including kinetic energy, potential energy, and electromagnetic energy.

**Engulf formation** — in medical nanorobotics, a configuration that is designed to capture the desired object (e.g. a cell) by the nanorobot.

**Enthalpy** — in thermodynamics, the internal energy of a system plus the product of its volume and the external pressure.

**Entropy** — in the physical sciences, a measure of uncertainty regarding the state of a system; free energy can be extracted by converting a low-entropy state to a high-entropy state. In other contexts, the term is often used by analogy to describe the extent of randomness and disorder in a system and the consequent lack of knowledge or information about it.

**Enucleated cell** — a cell from which the nucleus has been removed.

**Enzyme** — a protein molecule that often acts as a specific catalyst, facilitating specific chemical or metabolic reactions necessary for cell growth and reproduction; a biological chemosynthetic molecular machine.

**Eosinophil** — a type of granulocytic white blood cell comprising 1%-4% of all leukocytes, that is known to destroy parasitic organisms and to play a major role in allergic reactions (some of the major chemical mediators that cause bronchoconstriction in asthma are released by eosinophils).

**Epidermis** — the outer epithelial portion of the skin.

**Epitaxial methods** — a class of methods for inducing crystal growth.

**Epithelium** — the avascular layer of cells forming the epidermis of the skin and the surface layer of mucous (secreting mucus) and serous (secreting serum or serumlke fluid) membranes, including the glands. The cells rest on a basement membrane and lie closely approximated to each other with little intercellular material between them.

**Epitope** — any component of an antigen molecule that functions as an antigenic determinant by permitting the attachment of certain antibodies.
Equilibrium — a system is said to be at equilibrium (with respect to some set of feasible transformations) if that system has minimal free energy. A system containing objects at different temperatures is in disequilibrium, because heat flow can reduce the free energy (causing the different temperatures to converge). Springs have equilibrium lengths, reactants and products in solution have equilibrium concentrations, thermally excited systems have equilibrium probabilities of occupying various states, and so forth.

Ergoacoustic — pertaining to the conversion of energy of various kinds into acoustic energy.

Ergooptical (ergophotonic) — pertaining to the conversion of energy of various kinds into optical (photonic) energy.

Eructations — producing gas from the stomach, usually with a characteristic sound; belching.

Erysipelas — acute febrile disease with localized inflammation, eruption and redness of skin and subcutaneous tissue, accompanied by systemic signs and symptoms, caused by a hemolytic streptococcus.

Erythrocyte — red blood cell.

Erythropoietin — a hormone that controls the production rate of red blood cells in the human body.

Euchromatin — all of the genome in the interphase nucleus, except for the heterochromatin.

Eukaryote — an organism or cell that contains its genome within a nucleus.

Eutactic — characterized by precise molecular order.

Excimer laser — a kind of high-energy ultraviolet (UV) laser based on excited state transitions.

Excision — an act of cutting away or taking out.

Excluded volume — the presence of one molecule or moiety reduces the physical volume available for other molecules or moieties to occupy, thus partially crowding them out.

Excoriation — abrasion of the epidermis or of the coating of any organ of the body by trauma, chemicals, burns, or other causes.

Exfusion — generally, to remove from the body; distinguished from infusion.

Exocrine — external secretion of a gland.

Exocytosis — the process of secreting proteins from a cell into the surrounding medium, by transport in membranous vesicles from the endoplasmic reticulum, through the Golgi, to storage vesicles, and finally (upon a regulatory signal) through the plasma membrane.

Exodermal — lying outside of the skin surface.

Exoergic — a transformation that releases energy; such a reaction decreases molecular potential energy. Opposite of endoergic.

Exogenous — originating outside an organ, part, or system.

Exohedral — lying entirely outside a (fullerene) cage molecule.

Exons — portions of genes represented in mRNA.

Exothermic — a transformation that releases energy in the form of heat. Exoergic reactions in solution are commonly exothermic. Opposite of endothermic.

Exploratory engineering — design and analysis of systems that are theoretically possible but cannot be built yet, owing to the limitations of currently available tools.

Exponential — refers to a mathematical function (e.g. \( y = 10^x \)) which has an exponented quantity as its independent variable (e.g. \( x \)); more informally, may denote a very steeply rising (or falling) quantity, function or variable.

Extracellular — outside of the cell.

Extracellular matrix (ECM) — an extracellular fibrous scaffolding that helps organize cells into tissues.

Extrasomatic — originating or existing outside of the human body.

Extravasation — exiting the bloodstream.

Exudate — accumulation of a fluid in a cavity; matter that penetrates through vessel walls into adjoining tissue; the production of pus or serum.

Ex vivo — outside of the living human body.

Facultative — having the ability to do something which is not mandatory.

Fascia — a fibrous membrane covering, supporting, and separating muscles; also, unites the skin with underlying (e.g. muscular) tissue.

Fatty acid — a hydrocarbon in which one of the H atoms has been replaced by a carboxyl group.

Febrile — pertaining to fever.

Fenestrated — having openings.

Fibroblast — a stellate or spindle-shaped motile cell with cytoplasmic processes present in connective tissue, capable of forming collagen fibers.

Fission (fissile) — in physics, the release of energy when an atomic nucleus more massive than Fe56 (see binding energy) is split into two or more less massive fragments; the fission of nuclei lighter than Fe56 is typically endoergic. In microbiology, a method of asexual reproduction in bacteria, protozoa, and other lower forms of life; in cell biology, the partition of one organelle into two, as for example the fissioning mitochondrion.

Fistula — in anatomy, an abnormal tubelike passage from a normal cavity or tube to a free surface or to another cavity; may be due to
congenital incomplete closure of parts, or may result from abscesses, injuries, or inflammatory processes.

**Flagellum** — in cell biology, a whiplike locomotory organelle consisting of nine double peripheral microtubules and two single central microtubules.

**Flatus** — gas in the digestive tract; expelling of gas from a body orifice, especially the anus (e.g. to fart).

**FLOP** — in computer science, a floating point operation, an individual computational step as software executes in a computer. A floating point number is a number expressed as a product of a bounded number and an exponential scale factor, as in scientific notation.

**Fluidity** — in cellular biomechanics, a property of cell membranes, indicating the ability of lipids to move laterally within their particular monolayer.

**Flux (fluence)** — generally, a rate of flow.

**FM** — frequency modulation.

**Follower** — in mechanical engineering, a mechanical component in a cam system that is driven through a pattern of displacements as it rests against a moving contoured surface. See cam.

**Foramen magnum** — in anatomy, an opening in the occipital bone through which passes the spinal cord from the brain.

**Formed elements** — the cellular components of the blood, usually including red cells, white cells, and platelets.

**Fourier transform** — in mathematical physics, the expression of a complex waveform as a weighted sum of an infinite series of monofrequency waves.

**Fovea** — the clearest point of vision in the retina of the eye.

**Fractionation** — to separate components of a mixture into distinct molecular submixtures or pure components.

**Free energy** — a measure of the ability of a system to do work, such that a reduction in free energy could in principle yield an equivalent quantity of work. The Helmholtz free energy describes the free energy within a system; the Gibbs free energy does not.

**Fresnel lens** — a glass lens cast in a series of adjacent annular rings of various heights and shapes, commonly employed in lighthouses to produce powerful directed beams of light.

**Fusion** — in physics, the release of energy when atomic particles of total atomic mass less than the atomic mass of Fe56 (see binding energy) join together to form a more massive single particle; the fusion of nuclei heavier than Fe56 is typically endoergic. In cell biology, fusion is the merging of vesicles budded from the ER into the Golgi complex, or of endosomes with lysosomes, or of the contents of two cells by artificial means without the destruction of either, resulting in a heterokaryon that, for at least a few generations, will reproduce its kind (this was once an important method in assigning loci to chromosomes). In sensory physiology, the sensation that results when two different sensory qualities merge to give a third quality, such as the fusion of red and yellow into orange. Most generally, to combine together.

**g** — unit of gravitational acceleration (9.81 m/sec\(^2\)); describes the mean gravitational force experienced by a mass at rest on Earth’s surface.

**Galvanic** — pertaining to electrical direct current, usually chemically generated.

**Gamete** — either type of reproductive or germ cell (e.g. sperm or egg) with haploid chromosome content. Compare somatic cell.

**Ganglion** — a mass of nervous tissue composed principally of nerve-cell bodies and lying outside the brain or spinal cord (e.g. the chains of ganglia that form the main sympathetic trunks, or the dorsal root ganglion of a spinal nerve).

**Gangrene** — a necrosis (death) of tissue, usually due to deficient, obstructed, or lost blood supply; it may be localized to a small area or it may involve an entire organ or extremity.

**Gastro** — pertaining to the stomach.

**Gate** — see logic gate.

**GDP** — guanosine diphosphate.

**Gene** — a segment of a chromosome; an hereditary unit comprised of DNA occupying a locus in the genome which bears base sequential information for regulation, for transcription into RNA, or for transcription into mRNA destined for translation into protein.

**Genome** — the total hereditary material of a cell, containing the entire chromosomal set found in each nucleus of a given species.
Genomics — concerning the genome, the study of genes and how they affect the human body.

Genotype — the basic combination of genes of an organism. The genetic constitution of an organism.

Geodesic — pertaining to the mapping of the surface of the Earth.

Germ line — the cell line from which gametes derive.

GHz — gigahertz; billions of cycles per second.

Gibbs free energy — equals the Helmholtz free energy plus the product of the system volume and the external pressure. Changes in the Gibbs free energy at a constant pressure thus include work done against external pressure as a system undergoes volumetric changes.

Gimbal — in mechanical engineering, a device consisting of a pair of rings pivoted on axes at right angles to each other so that one is free to swing within the other.

Glioblastoma — a neuroglial (brain) cell tumor.

Global Positioning System (GPS) — a system of Earth-orbiting satellites that broadcasts navigational radio signals, allowing a ground receiver to precisely fix its position on the Earth’s surface.

Glottis — the sound-producing apparatus of the larynx, consisting of two vocal cords and the intervening space.

Glucose — simple blood sugar (C₆H₁₂O₆).

Glycocalyx — a thin layer of glycoprotein and polysaccharide that covers the surface of some cells, such as muscle cells, fibroblasts, pericytes, and epithelial cells, and contributes to the basal lamina.

Glycogen — a polysaccharide commonly called animal starch, with chemical formula (C₉H₁₄O₅)n; glycogen is the form in which carbohydrate is stored in the animal body (mostly in the liver and muscles) for future conversion into sugar. The formation of glycogen from carbohydrate sources is called glycogenesis, while the reverse process is called glycogenolysis.

Glycolysis — enzyme-mediated energy-yielding anaerobic hydrolysis of glucose to lactic acid in various cells and tissues, notably muscle.

Glycoprotein — a protein molecule with carbohydrate moieties attached.

Glycosylation — the covalent bonding of carbohydrate moieties to another molecule.

Golgi complex/apparatus — in cell biology, individual stacks of membranes near the endoplasmic reticulum involved in glycosylating proteins and sorting them for transport to different intracellular locations.

Gorget — in surgery, an instrument grooved to protect soft tissues from injury as a pointed instrument is inserted into a body cavity.

G-proteins — guanine nucleotide-binding trimeric proteins that reside in the plasma membrane. When bound by GDP, the trimer separates into a monomer and a dimer, one of which may then activate or repress a target protein.

GPS — see Global Positioning System.

Gradient — a slope or grade; an increase or decrease of varying degrees or the curve that represents such; rate of change of temperature, pressure, chemical concentration, or other physical variable, as a function of time, distance, frequency, etc.

Granules — a small, grainlike body. Small granules may be found in cells, containing stores of nutrients; large granules may be formed in tissues following a granulomatous reaction.

Granulocyte — a granular leukocyte; a polymorphonuclear (nucleus composed of two or more lobes or parts) leukocyte, including basophils, eosinophils, and neutrophils.

Granuloma — a nodular inflammatory lesion, usually small or granular, that is firm, persistent, and contains compactly grouped mononuclear phagocytes.

Granulomatous reaction — producing a granuloma, a granular tumor or growth, usually of lymphoid and epithelioid cells; an encapsulation reaction to the presence of a foreign object in the body that cannot be readily phagocytosed.

Gravimeter — a device for measuring gravitational acceleration.

Ground state — the lowest-energy state of a system. A system in its electronic ground state cannot further reduce its energy by an electronic transition, but may still contain vibrational energy.

GTP — guanosine triphosphate.

Gustatory — pertaining to the sense of taste.

Halogen — pertaining to fluorine, chlorine, bromine, iodine, or astatine.

Haploid — a set of chromosomes that contains one copy of each autosome and one sex chromosome.

Haploid number — the number of chromosomes in sperm or ova (23 in humans), half the number found in somatic (e.g. diploid) cells; characteristic of the gametes of diploid organisms.

Haptic — operated by, or pertaining to, the sense of touch.

Harmonic oscillator — in physics, any system of particles that displays harmonic (periodic oscillating) motion, such as a pendulum or spring with an attached mass; a system in which a mass is subject to a linear restoring force. A harmonic oscillator vibrates at a fixed frequency, independent of amplitude.

Haustra — in anatomy, the sacculated pouches of the colon.

Hawk — to make an audible effort to force up phlegm from the throat.

Hct — see hematocrit.
Heat — as defined in thermodynamics, heat is the energy that flows between two systems as a result of temperature differences, as distinguished from thermal energy. (A system contains neither heat nor work, but can produce heat or do work.)

Heat capacity — the ratio of the heat input to the temperature increase in a system.

Hectomicron — 100 microns.

Helmholtz free energy — the internal energy of a system minus the product of its entropy and temperature.

Hemapheresis — see apheresis.

Hematocrit (Hct) — volume-fraction or bloodstream concentration of erythrocytes (red blood cells), expressed as a percentage.

Hematoma — a swelling or mass of blood (usually clotted), confined to an organ, tissue, or other space, caused by a break in a blood vessel.

Hemobaric — pertaining to blood pressure.

Hemolytic — tending to destroy red blood cells, liberating hemoglobin.

Hemorrhagic — pertaining to bleeding.

Hemostasis — arrest of bleeding.

Hepatic — pertaining to the liver.

Hepatocyte — the most common tissue cell found in the liver.

Heterochromatin — regions of the genome that remain permanently in a highly condensed state and are not genetically expressed.

Histamine — a chemical substance, produced from the amino acid histidine, normally present in the body; exerts a pharmacological action when released from injured cells.

Histaminergic — activated or energized by histamine.

Histology — the study of tissues.

Histonomatation — in medical nanorobotics, locomotion (swimming) through tissues by a nanorobot.

Histonomatation — in medical nanorobotics, navigation through tissues by a nanorobot.

Histones — conserved DNA-binding proteins of eukaryotes that form the nucleosome, the basic subunit of chromatin; histones are rich in arginine and lysine residues.

HLA complex — Histocompatibility Locus Antigens, formerly known as Human Leukocyte Antigen (or Associated) complex.

Holliday junction — a connection between two DNA duplex molecules (as during recombination), wherein one duplex rotates relative to the other, creating a junction with 3-D topology.

Homeodomain — a DNA-binding motif in a transcriptional regulator that identifies, or at least is common in, the genes concerned with early embryonic developmental regulation.

Homeostasis — in physiology, a state of equilibrium of the internal environment of the body that is maintained by dynamic processes of feedback and regulation; homeostasis is a dynamic equilibrium (changing balance), keeping cells within the physical and chemical limits that can support life.

Homeothermic — tending to maintain a constant temperature, despite temperature variations in the surrounding environment.

Homochirality — chirality is a property of certain asymmetric molecules, the property being that the mirror images of the molecules cannot be superimposed one on the other while facing in the same direction. Homochirality is the preference of a process or system for a single optical isomer in a pair of isomers.

Homologous — similar in form (e.g. fundamental structure and origin), but not necessarily in function.

Hormone — a chemical substance that originates in an organ, gland, or part and is conveyed through the blood to another part of the body, stimulating that other part by chemical action to increase functional activity or to increase secretion of another hormone.

Hyaline cartilage — true cartilage; a smooth and pearly layer that covers the articular (place of union between two or more bones) surfaces of bones.

Hydrocarbon — a molecule consisting only of H and C.

Hydrodynamics — in physics, the study of the action of and motion of (and in) water and other liquids.

Hydrogen bond — the weak bond between a positively charged hydrogen atom that is covalently bound to one electronegative atom, and another electronegative atom.

Hydrolysis — a (hydrolytic) reaction in which a covalent bond is broken with the incorporation of a water molecule.

Hyrophilicity — tending to mix with water; wettable. Hydrophilic groups interact with water, so that hydrophilic regions of protein or the faces of a lipid bilayer reside in an aqueous environment.

Hyrophobic force — water molecules are linked by a network of hydrogen bonds; a nonpolar nonwetting surface such as wax cannot form hydrogen bonds, hence repels water.

Hyrophobicity — tending not to mix with water; nonwetting. Hydrophobic groups repel water, so that they interact with one another to generate a nonaqueous environment.

Hydrostatic — pertaining to the pressure of fluids or to fluid properties when in equilibrium.
Hygroscopic — readily absorbing and retaining moisture.

Hypercholesterolemia — presence of an abnormal amount of cholesterol in the cells and plasma of the circulating blood.

Hypergolic — refers to two substances which, although stable when stored separately, spontaneously combust or react when mixed.

Hypogravity — conditions of below-normal gravity (e.g. microgravity in Earth orbit).

Hypotension — low blood pressure.

Hypoxia — a condition in which the tissues are not receiving enough oxygen to sustain their metabolic activity.

Hypothermal limit — the maximum amount of energy that may be released at Earth’s surface, as a result of human technological activities, without significantly altering the natural global energy balance; estimated as $10^{13}$-$10^{15}$ watts.

Hz — hertz (MKS unit of frequency); cycles per second of an oscillation.

Iatrogenic disorder — an adverse condition induced in a patient by the actions of a physician.

Icosahedral — a solid figure with 20 planar faces.

Iliac crest — in anatomy, pertaining to the top of the ilium (the uppermost of the three sections of the hipbone).

Immunoglobulin — see antibody.

Impedance — opposition to flow (e.g. fluid, electrical, etc.) when flow is steady, or the driving pressure per unit flow when flow is changing; the resistance of an acoustic system to being set in motion.

Incision — a cut made with a knife.

In cyto — within a biological cell.

Inertia — in physics, the tendency of a body to remain in its kinematic state (e.g., at rest or in motion) until acted upon by an outside force; in biology, sluggishness or lack of activity.

Infarct — an area of tissue in an organ or part that undergoes necrosis following cessation of blood supply.

Inferior — beneath or lower; often refers to the undersurface of an organ or indicates a structure below another structure. See also caudal.

Infrasonic — inaudible (to human ears); sounds with acoustic frequencies less than ~16 Hz.

Infusion — the introduction of a fluid, other than blood, into a vein.

Inmessaging — in medical nanorobotics, conveyance of information from a source external to the human body, or external to working nanodevices, to a receiver located inside the human body.

Inner ear — the portion of the ear consisting of the cochlea, containing the sensory receptors for hearing, and the vestibule and semicircular canals, which include the receptors for balance and the sense of position.

Innervated — having nerves.

In nucleo — within the nucleus of a cell.

In sanguo — within the bloodstream.

Integral membrane protein — in cell biology, an amphipathic protein embedded in the lipid bilayer of the cell which cannot be extracted from the membrane without disrupting the lipid bilayer; most integral proteins are transmembrane proteins.

Intercalated — inserted between two others, as something interposed.

Intercostal — between the ribs.

Interferometry — measurement of very small spatial or frequency displacements using wave interference patterns.

Intermolecular — an interaction (e.g. a chemical reaction) between different molecules.

Internal energy — the sum of the kinetic and potential energies (including electromagnetic field energies) of the particles that make up a system.

Interphase — in cell biology, the quiescent period of time between mitotic cell divisions.

Interstitial — pertaining to extracellular interstices or spaces within an organ or tissue.

Intracellular fluid — all of the fluid inside a cell; the cytosol plus the fluid inside each organelle, including the nucleus.

Intracorporeal — inside the human body.

Intramolecular — an interaction (e.g. chemical reaction) within a single molecule. Intramolecular interactions between widely separated parts of a molecule resemble intermolecular interactions in most respects.

Intraperitoneal — within the peritoneal (abdominal) cavity.

Intravascular — inside a blood vessel.

In vacuo — in a vacuum (viz. the ablative case of the 2nd-declension Latin adjective “vacuus”).

Invaginate — to place or receive into a sheath; to receive within itself or into another part.

In vivo — inside the living human body.

Ion — an atom or molecule with a net charge.

Ion Beam Lithography (IBL) — lithography using ions rather than light beams.
London, England: a precision time stamp that is broadcast continuously by radio (e.g. NIST’s station WWV).  

**Corium** — the layer of the skin lying immediately under the epidermis.  

**Corneal** — pertaining to the cornea (the clear, transparent anterior portion of the fibrous coat of the eye comprising about one-sixth of its surface).  

**Coronal** — vertical, but at right angles to sagittal sections, dividing the body into anterior (front) and posterior (back) portions.  

**Cortex (cortical)** — the outer layer of an organ or of the body. Compare medulla.  

**Covalent bond** — in chemistry, a bond formed by sharing a pair of electrons between two atoms.  

**CPU** — see Central Processing Unit.  

**Cranial** — nearest or toward the head.  

**Crepitation** — a crackling sound heard in certain diseases, such as the rale heard in pneumonia; a grating sound heard on movement of ends of a broken bone; a clicking or crackling sound often heard in movements of joints, such as the temporomandibular (jaw), elbow, or patellofemoral (knee) joints, due to roughness and irregularities in the articulating surfaces.  

**Crinal** — pertaining to hair.  

**Crista** — in cell biology, a projection, sometimes branched, of the inner wall of a mitochondrion into its fluid-filled cavity; in anatomy, a ridge of sensory cells inside the ampulla, thrusting hair endings into the cupula (the cells respond to rotary acceleration and deceleration of the head).  

**Critical pressure** — the highest pressure at which gas and liquid may exist as separate phases at any temperature.  

**Critical temperature** — the highest temperature at which gas and liquid may exist as separate phases at any pressure.  

**Cross-links** — in biochemistry, additional bonds formed between normally separate parts of a polymer, typically increasing the tensile strength and stiffness of the chain.  

**Cryobiology** — the study of the effects of cold on biological systems.  

**Cryogenic** — producing, of pertaining to, low temperatures, typically the temperature of liquid nitrogen (77 K) or below.  

**Cryonic suspension** — a currently non-standard medical technique that attempts to prevent the permanent cessation of life, usually by promptly immersing the body of a post-pronouncement patient in a storage fluid maintained at cryogenic temperatures. A person who is cryonically suspended cannot be revived by 20th century medical technology because the freezing process does too much damage, but once frozen the patient’s biological condition does not further deteriorate. The procedure is attempted in the belief that future medical technology may allow reversal of tissue damage and the successful revival of the patient.  

**Crystallescence** — in medical nanorobotics, the crystallization of solid solute that is offloaded by nanorobot sorting rotors at a concentration that exceeds the solvation capacity of the surrounding solvent.  

**Cutaneous** — pertaining to the skin.  

**Cyclic** — a structure is termed cyclic if its covalent bonds form one or more rings.  

**Cycloaddition** — a chemical synthesis reaction in which two unsaturated molecules (or moieties within a molecule) bond to form a ring.  

**Cystic** — (usually) pertaining to the gallbladder or urinary bladder.  

**Cytapheresis** — see apheresis.  

**Cytoambulation** — in medical nanorobotics, cell surface walking.  

**Cytocarriage** — in medical nanorobotics, the commandeering of a natural motile cell, by a medical nanorobot, for the purposes of in vivo transport (of the nanorobot), or to perform a herding function (of the affected cell), or for other purposes.  

**Cytocide** — the killing of living cells.  

**Cytography** — a physical description (and mapping) of the living cell.  

**Cytoidentification** — identification of cell type.  

**Cytokine** — a group of extracellular biochemical substances that may be produced by a variety of cells, for the purposes of chemical messaging, regulation, and control; proteins that exert changes in the function or activity of a cell, such as differentiation, proliferation, secretion, or motility.  

**Cytology (cytological)** — the study of biological cells.  

**Cytometrics** — the quantitative measurement of cell sizes, shapes, structures, and numbers.  

**Cytonavigation** — in medical nanorobotics, navigation inside the cell; cellular navigation.  

**Cytopenetration** — in medical nanorobotics, entry into cells by penetrating the plasma membrane.  

**Cytoplasm** — the filling substance between the plasma membrane and the nucleus of a cell.  

**Cytoskeleton** — in medical nanorobotics, purposeful destruction of the cellular cytoskeleton by a nanorobot, for cytocidal purposes.
Ion channels — a large heterogeneous family of voltage-activated proteins that control the permeability of cells to specific ions by opening or closing in response to differences in potentials across the plasma membrane. Ion channels participate in the generation and transmission of electrical activity in the nervous system and in the hormonal regulation of cellular physiology.

Ionic bond — a chemical bond resulting chiefly from the electrostatic attraction between positive and negative ions.

Ionosphere — an upper region of Earth’s atmosphere.

IR — infrared.

Ischemia — local and temporary deficiency of blood supply due to obstruction of the circulation into a body part.

Isoareal — occurring without any change in surface area.

Isobaric — held, or existing, at a constant pressure.

Isomer — one of two or more chemical substances that have the same molecular formula but different chemical and physical properties due to a different arrangement of the atoms in the molecule; for example, dextrose is an isomer of levulose. Isomers may be geometric, optical, or structural. See also chiral, enantiomer, stereochemistry, and stereoisomer.

Isomagnetic — having the same magnetic field strength.

Isometric — in general, having equal dimensions; in physiology (or biomechanics), denoting a condition wherein the ends of a contracting muscle (or motor protein) are held fixed so that contraction produces increased tension at constant overall length.

Isoporous — having pores of the same size.

Isostatic — denoting a condition in which there is equal pressure on every side; in hydrostatic equilibrium.

Isotherm — region, section, or surface of constant temperature.

Isothermal — held, or existing, at a constant temperature.

Isotonic — animal cells containing a solution which exerts an osmotic pressure approximately equal to that of the surrounding fluid are isotonic or isoosmotic to that fluid. Stronger solutions that cause cells to shrink are hypertonic; weaker solutions that cause cells to swell are hypotonic.

Isotopic — any of two or more forms of the same chemical element that have nearly identical chemical properties but which differ in the number of neutrons contained in each atomic nucleus; many isotopes are radioactive.

Isotropic — the same in all directions.

Isovolemic — occurring without any change in volume.

IV — intravenous (inserted into a vein).

J — joule (MKS unit of energy).

Kelvins (K) — kelvin degrees (MKS unit of temperature).

Keratin — a sulfur-rich scleroprotein or albuminoid present largely in cuticular (pertaining to cuticles) structures.

Keratinocyte — a cell of the epidermis, and parts of the mouth, that produces keratin.

kg — kilogram (MKS unit of mass).

KHz — kilohertz; thousands of cycles per second.

Kinematic — pertaining to motion.

Kinesthesia — sensations of position, movement, direction and extent from the limbs, neck and body trunk.

Kinetic energy — energy resulting from the motion of masses.

Kupffer cells — macrophages lining the sinusoids of the liver.

Labial — pertaining to the lips.

Lacrimal — pertaining to the tears (eye fluid).

Lamellipodium — a cytoplasmic veil produced on all sides of a migrating polymorphonuclear leukocyte (granulocyte).

Laminar (Poiseuille) flow — fluid flow that moves exclusively along separate and independent parallel flow planes (i.e. streamlines), generally with an axisymmetric parabolic profile if in a tube. Laminar flow minimizes the impedance (resistance) and energy dissipation of fluid flow.


Lateral — on the side, or farthest from, the midsagittal plane; away from the midline of the body (to the side).

Lathe — in mechanical engineering, a machine for shaping an object, by holding and turning the object rapidly against the edge of a cutting tool.

LD50 — a dose of or exposure to a toxic influence that produces death in 50% of organisms exposed to it.

LED — light emitting diode.

Leukapheresis — see apheresis.

Leukocytes (white blood cells) — the primary effector cells that respond to infection and tissue damage in the human body. There are two types: granulocytes (including basophils, eosinophils, and neutrophils) and agranulocytes (including monocytes and lymphocytes). Leukocytes are formed from two stem cell populations in the bone marrow. The myeloid stem cell line produces granulocytes and monocytes, while the lymphoid stem cell line produces lymphocytes. Lymphoid cells travel to the thymus, spleen and lymph nodes, where they mature and differentiate into active, antigen-specific lymphocytes.

Leukokine — a cytokine secreted by a leukocyte.
**Leukotaxis** — active amoeboid movement of leukocytes, especially neutrophils, either toward (positive leukotaxis) or away from (negative leukotaxis) certain microorganisms or substances that frequently are formed in inflamed tissue; the property of attracting or repelling leukocytes.

**Ligand** — in protein chemistry, a small molecule that is (or can be) bound by a larger molecule; in organometallic chemistry, a moiety bonded to a central metal atom (this latter definition is more common in general chemistry).

**Ligation** — the formation of a phosphodiester bond to link two adjacent bases separated by a nick in one strand of a double helix of DNA; the term can also be applied to blunt-end ligation and to the joining of RNA.

**Linear** — in control theory, describes systems in which an output is directly proportional to an input.

**Lingual** — pertaining to the tongue.

**Lipid bilayer** — in cell biology, the form taken by a concentration of lipids in which the hydrophobic fatty acids occupy the interior and the hydrophilic polar heads face the exterior; primary constituent of the plasma membranes of cells.

**Lipids** — molecules having hydrophilic polar heads, containing phosphate (phospholipid), sterol (such as cholesterol), or saccharide (glycolipid) connected to a hydrophobic tail consisting of fatty acid.

**Lipofuscin** — brown pigment granules representing lipid-containing nondegradable residues of lysosomal digestion.

**Lipophilic** — having an affinity for lipids (fats).

**Lipophobic** — repulsed by lipids (fats).

**Liposomes** — the sealed concentric microscopic shells formed when certain lipid substances are in an aqueous solution. See also micelles.

**Lithography** — a technique for surface patterning commonly used to make semiconductor devices.

**Lithotomy** — in surgery, incision of a duct or organ, especially of the bladder, to remove a calculus (any abnormal concretion within the animal body, usually composed of mineral salts, commonly called stone).

**Lithotripsy** — crushing of a stone in the bladder or urethra.

**Load error** — in control theory, minimum range of variation in a control variable that is necessary to provoke a response from a control system.

**Logic gate** — in digital logic, a component that can switch the state of an output depending on the states of one or more inputs; a device implementing any of the elementary logic (or Boolean) functions (e.g. AND, NAND, NOR, NOT, OR, XOR). A logic gate is characterized by the relationship between its inputs and outputs, and not by its internal mechanisms.

**Loin (lumbar)** — lower part of the back and sides, between the thorax (ribs) and the pelvis (hips).

**LPS** — lipopolysaccharide, the lipid used to construct the outer leaflet of the outer bilayer membrane of Gram-negative bacteria.

**LR oscillator** — an electrical oscillator circuit comprised of resistors and inductors (solenoids).

**Lumbar** — see loin.

**Lumen** — the interior, especially of a compartment bounded by membranes, as for instance the endoplasmic reticulum or the mitochondrion.

**Luminal** — pertaining to the interior of a cavity, tube, or vessel.

**Lymph** — an alkaline fluid found in the lymphatic system.

**Lymphatic system** — includes all structures involved in the conveyance of lymph from the tissues to the bloodstream, including lymph capillaries, lacteals, lymph nodes, lymph vessels, main lymph ducts, and cisterna chyli.

**Lymphocyte** — a morphologically distinct variety of leukocytes, comprising 20%-44% of all white blood cells. But only ~2% of all lymphocytes present in the human body are in the bloodstream; most reside elsewhere, particularly in the lymph and the lymph nodes. B-lymphocytes differentiate into antibody-secreting plasma cells, whereas T-lymphocytes play diverse regulatory roles in the immune response.

**Lysis (lytic)** — in microbiology, the death of a bacterium at the end of a bacteriophage infective cycle when the bacterium bursts open to release the progeny of an infecting phage; also applies to eukaryotic cells, as for example infected cells that are attacked by the immune system. More generally, dissolution or decomposition.

**Lysosomes** — small bodies inside cells, enclosed by membranes, that contain hydrolytic enzymes that are part of the cell’s digestive apparatus.

**Lysozyme (muramidase)** — an enzyme that is destructive to cell walls of certain bacteria, found in white blood cells of the granulocytic and monocytic series.

**m** — meter (MKS unit of length)

**M** — see molarity.

**Macromolecule** — a molecule of colloidal size, typically 1-100 nm in diameter or length, consisting most notably of proteins, nucleic acids, and polysaccharides.

**Macrophage** — a monocyte that has left the circulation and settled and matured in a tissue; found in large numbers in the spleen, lymph nodes, alveoli, and tonsils, with ~50% found in the liver as Kupffer cells. Along with neutrophils, macrophages are the major phagocytic cells of the immune system, able to recognize (and then ingest) foreign antigens via chemical receptors on the surface of their cell membranes. Macrophages also serve a vital role by processing
antigens and presenting them to T cells, activating the specific immune response.

**Macroscopic** — easily visible to the human naked eye; typically ~1 mm³ or larger.

**Macrosensing** — in medical nanorobotics, the detection of global somatic states (inside the human body) and extrasomatic states (sensory data originating outside of the human body) by in vivo nanorobots.

**Magnetic Force Microscope (MFM)** — a kind of SPM that images surfaces using magnetic forces.

**Maillard reaction** — in food science, the "browning" reaction that occurs between proteins and reducing sugars as they are heated.

**Major histocompatibility complex (MHC)** — the complex of HLA genes on the short arm of human chromosome 6.

**Manometer (manometry)** — a device for measuring liquid or gaseous pressure such as blood, spinal fluid, or atmosphere (air).

**Manustupration** — masturbation.

**Margination** — adhesion of leukocytes to endothelial cells lining the walls of a blood vessel, during the relatively early stages of inflammation; more generally, the process of differential radial migration among suspended particles of different sizes during fluid flow through a tube.

**Massometer** — in medical nanorobotics, a nanosensor device for measuring the mass of individual molecules or small physical objects to single-proton resolution.

**Mast cells** — cells resident in connective tissue just below epithelial surfaces, serous cavities, and around blood vessels, that synthesize, store, and release (upon stimulation) histamine and other local chemical mediators of inflammation (e.g. leukotrienes).

**Mastication** — chewing.

**Mechanical nanocomputers** — nano-sized computers that use mechanical rather than chemical or electronic logic switches.

**Mechanomechanical conversion** — conversion of one form of mechanical energy into another form of mechanical energy.

**Mechanochemistry** — the chemistry of processes in which mechanical systems operating with atomic-scale precision guide, drive, or are driven by chemical transformations; in more general usage, the chemistry of processes in which energy is converted from mechanical to chemical form, or vice versa (aka, piezochemistry).

**Mechanosynthesis** — chemical synthesis controlled by mechanical systems operating with atomic-scale precision, enabling direct positional selection of reaction sites; synthetic applications of mechanochemistry.

**Medial** — middle or nearest to the midsagittal plane; toward the midline of a body or vessel.

**Medulla (medullary)** — the inner or central portion of an organ. Compare cortex.

**Medulla oblongata** — enlarged portion of the spinal cord in the cranium, after the cord enters the foramen magnum of the occipital bone; the lower portion of the brain stem.

**Melanoma** — a malignant, darkly pigmented mole or tumor of the skin.

**Membrane** — in cell biology, an asymmetrical lipid bilayer that has lateral fluidity and contains proteins; in anatomy, a thin, soft, pliable layer of tissue that lines a tube or cavity, covers an organ or structure, or separates one part from another.

**Membrane proteins** — in cell biology, plasma membrane proteins that have hydrophobic regions that allow part or all of the protein structure to reside within the membrane; the bonds involved in this association are usually noncovalent.

**Membranolytic** — causing the physical failure of a membrane.

**MEMS** — see microelectromechanical systems.

**Meniscus** — the curved upper surface of a liquid in contact with a container.

**-mer** — see oligomer.

**Mesenchyme** — a diffuse network of cells forming the embryonic mesoderm and giving rise to connective tissues, blood, and blood vessels, the lymphatic system, and cells of the RES.

**Mesentery** — a peritoneal fold encircling the greater part of the small intestines and connecting the intestine to the posterior abdominal wall.

**Mesoderm** — a tissue layer in the embryo from which arises all connective tissues, including the muscular, skeletal, circulatory, lymphatic, and urogenital systems, and the linings of the body cavities.

**Mesoscopic** — lying midway in size between the nanoscale (~nanometers) and the microscale (~microns).

**Mesothelium** — the layer of cells derived from the mesoderm that lines the primitive body cavity; in the adult, it becomes the epithelium covering the serous membranes.

**Messenger molecule** — a chemically recognizable molecule which can convey information after it is received and decoded by an appropriate chemical sensor.

**Messenger RNA (mRNA)** — the RNA whose sequence corresponds to that of exons in the transcribed gene, which embodies the codons and is translated into the protein gene product.

**Metabolism** — the sum of all physical and chemical changes that take place within an organism; all energy and material transformations that occur within living cells, including anabolism and catabolism.
Metamorphic — in medical nanorobotics, capable of adopting multiple physical configurations via smooth changes from one configuration to another.

Metaphase — stage of mitosis (cell division) in which all chromosomes are aligned equidistant between the spindle poles.

Metastable state — a high energy state that requires an input of energy to relax to the ground state.

Metastasize — usually refers to the manifestation of a malignancy (e.g. of cancerous body cells) as a secondary growth arising from the primary growth, but in a new location.

Metazoa — all multicellular life.

Metrology — the science of weights and measures.

MeV — million electron volts (unit of energy).

MHC — see major histocompatibility complex.

MHz — megahertz; millions of cycles per second.

Micelle — a self-assembling hollow spheroidal aggregate of amphipathic lipids in a polar liquid (e.g. aqueous) medium.

Michaelis-Menten enzyme — an enzyme whose saturation kinetics obey the Michaelis-Menten equation, which describes the half-maximal reaction velocity as a function of substrate concentration; certain enzymes and other ligand-binding proteins such as hemoglobin may obey sigmoid (e.g., Hill equation) substrate saturation kinetics instead.

Microbarom — a low-amplitude, slowly-varying atmospheric pressure wave.

Microbiotagraphics — mapping the microbiotic populations present in the human body.

Microelectromechanical systems (MEMS) — micron-scale devices combining electronic and mechanical elements.

Micron — one-millionth of a meter; a micrometer.

Micro-opto-mechanical systems (MOMS) — micron-scale devices combining electronic, optical and mechanical elements.

Mole — a number of instances of something (e.g. molecular objects) equal to $6.023 \times 10^{23}$ objects.

Molar — in anatomy, pertaining to a grinding (back) tooth.

Molarity (M) — in chemistry, moles of solute per liter of solvent.

Molecular assembler — a general-purpose device for molecular manufacturing, able to guide chemical reactions by positioning individual molecules to atomic accuracy (e.g. mechanosynthesis) and to construct a wide range of useful and stable molecular structures according to precise specifications.

Molecular Beam Epitaxy (MBE) — a technique for producing single-layer crystals.

Molecular electronics — any system of atomically precise electronic devices of nanometer dimensions, especially if made of discrete molecular parts rather than the continuous bulk materials found in 20th century semiconductor devices.

Molecular machine — a mechanical device that performs a useful function using components of nanometer scale and a well-defined molecular structure; may include both artificial nanomachines and naturally occurring devices found in biological systems.

Molecular machine system — a system of molecular machines.

Molecular manipulator — a manipulator device able to position molecular tools with high precision, for example to direct varying sequences of mechanosynthetic steps; a major component of a molecular assembler.

Molecular manufacturing — manufacturing using molecular machinery, giving molecule-by-molecule control of products via positional chemical synthesis, to produce complex molecular structures manufactured to precise specifications.
**Molecular medicine** — a variety of pharmaceutical techniques and gene therapies that address specific molecular diseases or molecular defects in biological systems.

**Molecular mill** — a mecanochemical molecular transport and processing system characterized by limited motions and repetitive operations without programmable flexibility.

**Molecular nanotechnology** — thorough, inexpensive control of the structure of matter based on molecule-by-molecule control of products and byproducts; the products and processes of molecular manufacturing, including molecular machinery; a technology based on the ability to build structures to complex, atomic specifications by mechanosynthesis or other means; most broadly, the engineering of all complex mechanical systems constructed from the molecular level.

**Molecular recognition** — a chemical term referring to processes in which molecules adhere in a highly specific way, forming a larger structure; a possible enabling technology for nanotechnology.

**Molecular sorting rotor** — a class of nanomechanical device capable of selectively binding (or releasing) molecules from (or to) solution, and of transporting these bound molecules against significant concentration gradients.

**Molecular surgery (molecular repair)** — in medical nanorobotics, the analysis and physical correction of molecular structures in the body using medical nanomachines.

**Molecular systems engineering** — design, analysis, and construction of systems of molecular parts working together to carry out a useful purpose.

**Molecule** — group of atoms held together by chemical bonds; the typical unit manipulated by molecular nanotechnology.

**Monkeywrenching** — in medical nanorobotics, the mechanical or chemical jamming of cellular equilibrium processes, with a cytocidal objective. See also disequilibration.

**Monocyte** — a mononuclear phagocytic white blood cell derived from the myeloid stem cells, that is short-lived (~1 day half-life) and circulates in the bloodstream from which it moves into tissues, at which point it matures into a macrophage (which is long-lived). Monocytes represent 3%-8% of all white blood cells.

**Monomer** — any molecule that can be bound to similar molecules to form a polymer.

**Morcellation** — fragmentation of biological materials as they are excised from the body.

**Morphogen** — a (biochemical) factor that induces development of particular cell types in a manner that depends on its concentration.

**Morphology** — the science of structure and form, without regard to function.

**Motif** — main element, theme, or design feature that may occur in slightly different forms.

**Motile** — capable of voluntary movement. Opposite of sessile.

**MRI** — magnetic resonance imaging.

**MRFM** — magnetic resonance force microscope.

**mRNA** — see messenger RNA.

**MTOC** — microtubule organizing center; a cellular structure from which microtubules may be extended. The major MTOCs in a mitotic (dividing) cell are the centrosomes.

**Mucosa** — a mucous membrane; the moist tissue layer that lines a hollow organ or body cavity.

**Muller cells** — neuroglial cells with fine fibers that form supporting elements of the retina.

**Multivalent** — in chemistry, able to form more than one covalent or ionic bond; in biomedicine, efficacious in more than one direction, or active against several strains of an organism.

**Mutagenesis, site-directed** — a biotechnological technique whereby a gene encoding an enzyme is mutated, then overexpressed in a unicellular host, then characterized. By altering the base sequence of a single codon, site-directed mutagenesis can replace a given amino acid with any other protein amino acid, and can also generate multiple point mutants.

**Mycelium** — the mass of hyphae (branching tubular cells characteristic of the growth of filamentous fungi) making up a colony of fungi.

**Myelin** — a fatlike substance forming a sheath around the axons of certain nerves; composed of lipids and protein.

**Myocardium** — heart muscle.

**Myoepithelium** — tissue containing contractile epithelial cells (e.g. contractile spindle-shaped cells arranged longitudinally or obliquely around sweat glands and around the secretory alveoli of the mammary and salivary glands).

**Myopotential** — pertaining to an action potential in an axon innervating a muscle cell.

**N** — in physics, newton (MKS unit of force); in solution chemistry, see Normal.

**Naked DNA** — DNA that is not surrounded by an outer protein envelope.

**Nanapheresis** — in medical nanorobotics, the removal of bloodborne medical nanorobots from the body using apheresis-like processes.

**NAND gate** — a logical gate equivalent to a negated (NOT-) AND gate.

**nano** — a prefix meaning one billionth (1/1,000,000,000); more specifically, may refer to sizes of ~10^{-9} meters or to physical
objects constructed or processes performed at the atomic or molecular scale.

**Nanocentrifuge** — in medical nanorobotics, a proposed nanodevice that can spin materials at very high speed, imparting rotational accelerations of up to one trillion gravities (g’s), thus permitting rapid sortation.

**Nanochronometer** — in medical nanorobotics, a proposed clock or timing mechanism constructed of nanoscale components.

**Nanoclusters** — nanoscale aggregates of atoms or molecules.

**Nanocomputer** — a proposed computer with parts built on a molecular scale, possibly employing molecular electronics, mechanical rod logic, or biological molecules.

**Nanocrit (Nct)** — in medical nanorobotics, volume-fraction or bloodstream concentration of medical nanorobots, expressed as a percentage.

**Nanolithography** — the process of transferring nanometer-sized surface patterns.

**Nanomachine** — functional machine systems on the scale of nanometers; an artificial mechanical device constructed with precise molecular order using nanometer-scale components; any molecular structure large and complex enough to function as a machine.

**Nanomanipulator** — a nanorobotic manipulator device.

**Nanomanufacturing** — see molecular manufacturing.

**Nanomechanical** — pertaining to nanomachines.

**Nanomedicine** — (1) the comprehensive monitoring, control, construction, repair, defense, and improvement of all human biological systems, working from the molecular level, using engineered nanodevices and nanostructures; (2) the science and technology of diagnosing, treating, and preventing disease and traumatic injury, of relieving pain, and of preserving and improving human health, using molecular tools and molecular knowledge of the human body; (3) the employment of molecular machine systems to address medical problems, using molecular knowledge to maintain and improve human health at the molecular scale.

**Nanometer** — a billionth of a meter, roughly the diameter of 3-7 atoms.

**Nanorobot** — a computer-controlled robotic device constructed of nanometer-scale components to molecular precision, usually microscopic in size (often abbreviated as “nanobot”).

**Nanosensor** — a chemical or physical sensor constructed using nanoscale components, usually microscopic or submicroscopic in size.

**Nanosieving** — in medical nanorobotics, a nanodevice that can sort molecules or other nanoscale objects by physical sieving.

**Nanosystem** — a set of nanoscale components, characterized by precise molecular order, working together to serve a set of purposes; complex nanosystems can be of macroscopic size.

**Nanotechnology** — engineering and manufacturing at nanometer scales; any technology related to features of nanometer scale, including thin films, fine particles, chemical synthesis, advanced microlithography, and so forth, as well as complex mechanical systems constructed from the molecular level.

**Nanotubes** — hollow fullerene tubes, including but not limited to single- and multi-walled carbon nanotubes, with submicroscopic, often nanoscale, diameters and a wide range of continuous lengths.

**NASA** — National Aeronautics and Space Administration.

**Nasal septum** — in anatomy, the cartilage, covered with mucous membrane, that divides the two nostrils and nasal passages.

**Nasopharynx** — in anatomy, the nasal passages, mouth, and upper throat.

**Natation** — swimming.

**Naturophilia** — an exclusive love of Nature, disdaining everything that is artificial or technological.

**Nauseogenic** — tending to produce nausea.

**Navicyte** — in medical nanorobotics, a mobile, mass-storage (nanorobotic) device, similar to a communicyte, that may be used to establish a navigational network inside the human body.

**N/C** — numeric control machining (e.g. computer-aided manufacturing).

**Nct** — see nanocrit.

**Necrosis (necrotic)** — the death of areas of tissue or bone, surrounded by healthy parts.

**Neoplastic** — pertaining to, or of the nature of, new and abnormal tissue (i.e. neoplasm) formation and growth.

**Neurofibrils** — a filamentous aggregation of microfilaments and microtubules found in the main cell body, dendrites, and axon of the nerve cell, and sometimes in its synaptic endings.

**Neuron** — a nerve cell, the principal structural and functional unit of the nervous system.

**Neuropeptide** — any of a variety of neurotransmitter peptides found in neural tissue (e.g. endorphins, enkephalins).

**Neurotransmitter** — a biochemical substance released when the axon terminal of a presynaptic neuron is excited. The substance travels across the synapse to act on the target cell to either inhibit or excite it.

**Neutrophil** — the most common type of granulocytic white blood cell. Neutrophils are responsible for much of the body’s protection against infection. Comprising ~60% of all white blood cells, neutrophils play the primary role in inflammation, easily recognizing foreign antigens and destroying them through phagocytosis. Neutrophils also may overreact to stimuli and become involved in tissue
destruction, as in rheumatoid arthritis, myocardial reperfusion injury, respiratory distress syndrome, and ulcerative colitis.

NMR — see Nuclear Magnetic Resonance.

NIST — National Institute of Standards and Technology.

Nociceptors — pain receptors.

Nodes of Ranvier — constrictions of the myelin sheath of a myelinated nerve fiber.

Nonpolar liquid — a liquid consisting of molecules that have zero electric dipole moment (e.g. ethane, benzene, liquefied carbon dioxide). Compare polar liquid.

NOR gate — a logical gate equivalent to a negated (NOT-) OR gate.

Normal (N) — in solution chemistry, 1 liter of solution which contains 1 gram-equivalent of solute.

Nosogenic — tending to cause disease.

NOT gate — a logical gate that inverts or negates the input signal, as its output.

NSOM — Near-field Scanning Optical Microscope.

Nuclear cortex — a proteinaceous layer on the inside of the cellular nuclear envelope, consisting of up to three kinds of laminin proteins.

Nuclear envelope — a layer consisting of two membranes surrounding the cellular nucleus, penetrated by nuclear pores and bounded on the interior by the nuclear cortex.

Nuclear lamina — see nuclear cortex.

Nuclear Magnetic Resonance (NMR) — in physics and radiology, an analytical technique that relies on the absorption of certain electromagnetic frequencies by atomic nuclei.

Nuclear matrix — a network of fibers surrounding and penetrating the cell nucleus.

Nuclear pores — holes in the cellular nuclear envelope used for transport of macromolecules.

Nucleic acids — polymers of nucleotidyl residues.

Nucleoelectric conversion — conversion of nuclear energy into electrical energy.

Nucleography — a physical description (and mapping) of the nucleus of a human cell.

Nucleolus — an RNA-rich spherical body made up of dense fibers and granules within the cell nucleus, associated with chromosomal loci for rRNA genes, created by the transcription of rRNA genes.

Nucleoplasm — the protoplasm of a cell nucleus.

Nucleosides — compounds consisting of a purine or pyrimidine attached to ribose or deoxyribose at the 1’ carbon.

Nucleosome — the basic structural subunit of chromatin, consisting of 146 base pairs of DNA wrapped around a core of eight histone protein molecules.

Nucleotide (nucleotidyl) — phosphorylated nucleosides; nucleotidyl residues are the monomeric units of RNA and DNA. The building blocks of nucleic acids. Each nucleotide is composed of sugar (ribose or deoxyribose), phosphate, and one of four nitrogen bases (purine or pyrimidine). The sequence of the bases within the nucleic acid determines what proteins will be made.

Nucleus — in physics, the positively-charged core of an atom, an object of ~0.00001 atomic diameters containing >99.9% of the atomic mass; in cell biology, a cellular organelle containing the genome, separated from the cytoplasm by a membrane.

Nyctalopia — a vitamin A deficiency disease, characterized by decreased ability to see in reduced illumination or at night.

Obligate — to make necessary or require, without alternative.

Occipital — pertaining to the back part of the head.

Occlusal surface — in dental medicine, the masticating surface of the premolar and molar teeth.

Olfaction (olfactory) — pertaining to the sense of smell.

Oligomer — any short polymeric molecular chain consisting of well-defined subunits; N-mer refers to an oligomer that has exactly N subunits.

Oligonucleotide — a polymeric chain that consists of a small number of (possibly different) nucleotides.

Opsonin — a biochemical substance that coats foreign antigens, making them more susceptible to macrophages and other leukocytes, thus increasing phagocytosis of the organism. Complement and antibodies are the two main opsonins in human blood.

Opsonization — the coating action of opsonins, thus facilitating phagocytosis.

Organelle — most commonly described subcellular compartment, located in the cytoplasm, that is surrounded by a membrane (e.g. lysosome, mitochondrion).

Organography — a physical description (and mapping) of the body organs.

OR gate — a logical gate that returns a high (1) output if either input signal is high (1). See bit.

Orthogonal — at right angles; mutually perpendicular.

Orthotropic — capable of alternating between bent and straight configurations.

Osmosis — the passage of solvent through a semipermeable membrane that separates solutions of different concentrations.
Osmotic pressure — the pressure that would develop if a solution is enclosed in a solvent-permeable membrane that is impermeable to all solutes present, and is then surrounded by pure solvent.

Osseous — pertaining to bone.

Ossicle — any small bone, especially one of the three bones of the ear.

Osteoblast — a bone-forming cell derived from mesenchyme to form the osseous matrix in which it becomes enclosed as an osteocyte.

Osteoclast — a giant multinuclear cell with abundant acidophilic cytoplasm, formed in the bone marrow of growing bones, which functions to absorb and remove unwanted osseous tissue.

Osteocyte — a mesodermal bone-forming cell that has become entrapped within the bone matrix, helping to maintain bone as living tissue.

Osteography — a physical description (and mapping) of the human skeletal system.

Osteomalacia — a vitamin D deficiency disease, characterized by a gradual and painful softening and bending of the bones.

Ostium — small opening, especially one into a tubular organ.

Otolith — a crystalline particle of calcium carbonate and protein, adhering to the gelatinous membrane of the maculae (thickened areas in the utricle and saccule) of the inner ear.

Ouabain — a glycoside (C_{29}H_{44}O_{12}•H_{2}O) obtained from the wood of *Acocanthera ouabaio* or from the seeds of *Strophanthus gratus*; its action is qualitatively identical to that of the digitalis glycosides (e.g. increasing the contractility of cardiac muscle).

Outmessaging — in medical nanorobotics, conveyance of information from a transmitter located inside the human body, especially from working nanodevices, to the patient or to a recipient external to the human body.

Ovariotomy — incision into an ovary.

Oxidation — in chemistry, a combination with oxygen, or an increase in positive valence of an ion, atom, or molecule, either by the loss from it of hydrogen or by the loss of one or more electrons. In bacteriology, the aerobic disassimilation of substrates with the production of energy and water; in contrast to fermentation, the transfer of electrons is accomplished via the respiratory chain, which utilizes oxygen as the final electron acceptor.

Oxidation-reduction reaction (redox) — chemical interaction in which one substance is oxidized and loses electrons, thus is increased in positive valence, while another substance gains an equal number of electrons by being reduced, thus is decreased in positive valence.

Oxyglucose — pertaining to a chemoergic energy conversion system involving the glucose/oxygen reaction; a glucose (fuel) and oxygen (oxidant) supply or mixture.

Osteograph — a conceptual model, pattern, or system.

Parenchyma — the essential parts of an organ that are concerned with its function as opposed to its framework; opposite of stroma. The distinguishing or specific cells of a gland or organ, contained within and supported by the connective tissue framework.

Paracrine control — a general form of bioregulation in which one cell type in a tissue selectively influences the activity of an adjacent cell type by secreting chemicals that diffuse into the tissue and act specifically on cells in that area.

Paradigm — pertaining to the palate (roof of the mouth).

Papilla — in anatomy, a small nipple-like protuberance or elevation.

Parotid — pertaining to, or forming, the walls of a cavity; often specifically refers to the parietal bone, one of two bones that together form the roof and sides of the skull.

Pariental — denoting any medication route other than the alimentary canal, such as intravenous, subcutaneous, intramuscular, or mucosal.

Paradigm — a conceptual model, pattern, or system.

Parenchyma — the essential parts of an organ that are concerned with its function as opposed to its framework; opposite of stroma. The distinguishing or specific cells of a gland or organ, contained within and supported by the connective tissue framework.

Paracrine control — a general form of bioregulation in which one cell type in a tissue selectively influences the activity of an adjacent cell type by secreting chemicals that diffuse into the tissue and act specifically on cells in that area.

Parturbation — the act of giving birth to young; childbirth.

Passivation — the covalent bonding of a layer of atoms to a surface, in order to neutralize (occupy) any dangling surface bonds, thus chemically stabilizing the surface.

Patency — the state of being freely open.

Pathogen — a microorganism or agent capable of producing disease.

Pathogramonic — characteristic or indicative of a disease; relating to one or more of the typical symptoms of a disease.

Pathological — diseased or due to a disease; more informally, pertaining to an adverse condition.

PEG — polyethylene glycol.

Pellagra — a niacin deficiency disease, characterized by gastrointestinal disturbances, erythema (inflammatory skin redness) followed by desquamation (shedding epidermis in scales or shreds), and nervous and mental disorders.

Peptide — a short chain of amino acids joined by amide bonds, up to 100 residues in length.

Peptidergic — activated or energized by neuropeptides.

Peptidoglycan layer — the dense material consisting of cross-linked polysaccharide chains present in the cell wall of most bacteria.

Periaxonal — near or around the neural axon.
**Perineum** — in anatomy, the area between the thighs extending from the coccyx to the pubis and lying below the pelvic diaphragm; the structures forming the pelvic floor.

**Perinuclear space** — in cell biology, the space that lies between the inner and outer membranes of the nuclear envelope.

**Periosteum** — near the portal end.

**Periportal** — in anatomy, the fibrous membrane forming the investing covering of bones except at their articular (joint) surfaces.

**Peripheral** — pertaining to a part of a body or object that is away from the center.

**Peripheral protein** — a non-amphipathic protein bound to the hydrophilic region of an integral membrane protein, or bound to the hydrophilic heads of plasma membrane lipids, of a cell; most peripheral proteins are located near the cytoplasmic side of the plasma membrane.

**Peristalsis** — a progressive wavelike movement that occurs involuntarily in hollow tubes of the body, especially the alimentary canal; it is characteristic of tubes possessing longitudinal and circular layers of smooth muscle fibers.

**Peritoneum** — in anatomy, the serous membrane reflected over the viscera and lining the abdominal cavity.

**Periurethral** — in anatomy, located near or around the urethra (which discharges urine).

**Permittivity (relative)** — in physics, the dielectric constant, a measure of the ability of a material to store electrical energy.

**Peroxisome** — in cell biology, an organelle found in vertebrate animal cells that contains a great number and variety of enzymes important in cell metabolism.

**PET** — Positron Emission Tomography; see Computerized Tomography.

**Phage** — see bacteriophage.

**Phagocyte** — a cell with the ability to ingest and destroy particulate substances such as bacteria, protozoa, cells and cell debris, dust particles, and colloids.

**Phagocytosis** — ingestion and digestion of bacteria and particles by phagocytes.

**Pharmacodynamics** — the study of the action of drugs on living organisms, and the physiological and clinical changes produced in organisms by those drugs.

**Pharmacogenetics** — the study of the influence of hereditary factors on the response of individual organisms to drugs; a branch of genetics dealing with hereditary aspects of drug metabolism.

**Pharmacokinetics** — study of the metabolism of drugs with particular emphasis on the time required for absorption, duration of action, distribution in the body, and method of excretion.

**Pharmacyte** — in medical nanorobotics, a theorized (nanorobotic) device capable of delivering precise doses of biologically active chemicals to individually-addressed human body tissue cells (e.g. cell-by-cell drug delivery).

**Pharynx** — the passageway for air from the nasal cavity to the larynx (also acting as a resonating cavity), and for food from the mouth to the esophagus; more specifically, a musculomembranous tube extending from the base of the skull to the level of the 6th cervical vertebra, where the tube becomes continuous with the esophagus.

**Phase** — fraction of a periodic cycle, often expressed as an angle.

**Phasic (muscular)** — refers to muscles with origins and insertions on skeleton or skin, that have rapid, brief contraction cycles. Compare tonic.

**Phenotype** — the appearance or other characteristics of an organism, resulting from the interaction of its genetic constitution with the environment; any observable characteristic that expresses the genotype of an individual.

**Pheresis** — see apheresis.

**Pheromone** — a substance permitting chemical communication between animals, and between certain insects, of the same species; may affect development, reproduction, or behavior of like individuals.

**Phonons** — in physics, a quantum of acoustic energy, analogous to the quantum of electromagnetic energy, the photon. Thermal excitations in a crystal (lattice vibrations) or in an elastic continuum can be described as a population of phonons (analogous to black-body electromagnetic radiation).

**Photolithography** — a surface patterning technique that employs light, used to make semiconductor devices.

**Physiognomy** — the countenance (physical features) of the face.

**Phytotoxic** — pertaining to a poisonous plant.

**Piezoelectric** — the property of some materials to generate a voltage when mechanical stress is applied, and vice versa.

**Pinocytosis** — the process by which cells absorb or ingest nutrients and fluid, in which minute incuppings or invaginations are first formed in the surface of the plasma membrane and then close to form fluid-filled vesicles; resembles phagocytosis.

**Pixel** — abbreviation for picture element; a single square cell within a two-dimensional geometric grid or array.

**Plague** — see bubonic plague, pneumonic plague.

**Planetary gear** — in machinery, an epicyclic train of gears as found in an automobile transmission, for the purpose of transforming the rotational speeds of rotating shafts.

**Plasma** — in anatomy, the fluid (noncellular) part of the lymph and of the blood, usually distinguished from the serum obtained
after coagulation; in cell biology, the part of the protoplasm (cell substance) outside of the nucleus.

**Plasma membrane** — the outermost membrane of a cell, with cell contents on one side and the extracellular environment on the other side; the continuous membrane defining the boundary of every cell.

**Plasmapheresis** — see apheresis.

**Plasmid** — an autonomous self-replicating extrachromosomal circular DNA molecule present intracellularly and symbiotically in most bacteria, encoding a protein product that confers drug resistance or some other advantageous phenotype. Plasmids reproduce inside the bacterial cell but are not essential to its viability, and can influence a great number of bacterial functions.

**Plasticity** — a property of an object or material, wherein the object or material suffers a permanent deformation in its original shape after a force is applied and then removed.

**Platelet** — a round or ovoid 2-4 micron disk found in the blood of vertebrates; platelets play an important role in blood coagulation and hemostasis.

**Plateletocrit** — volume-fraction or bloodstream concentration of platelets, expressed as a percentage.

**Pleura** — serous membranes that enfold both lungs and are reflected on the walls of the thorax and diaphragm, and are moistened with a serous secretion that reduces friction during the respiratory movements of the lungs.

**Pluripotent** — concerning an embryonic stem cell that can differentiate into many different kinds of cells.

**Plexus** — an extensive two-dimensional meshwork.

**Pneumatic** — driven to motion or extension by pressurized gas or fluid.

**Pneumonic plague** — highly virulent and frequently fatal form of plague, having an extensive involvement of the lungs.

**Poiseuille flow** — laminar (nonturbulent) fluid flow.

**Polarization** — in optical physics, a change effected in a ray of light passing through certain media, whereby the transverse vibrations of the emerging ray occur in one plane only, instead of in all planes as in the ordinary incident light; in biology and electrical physics, the development of differences in electrical potential between two points on an object, such as between the inside and outside of a cell wall or along the length of a piezoelectric bone subjected to shear stress.

**Polar liquid** — a liquid consisting of molecules that have an electric dipole moment (e.g. water, ethanol, liquefied ammonia). Compare nonpolar liquid.

**Polymer** — a long molecular chain of well-defined linked subunits.

**Polynucleotides** — polymers of nucleotidyl residues, as in RNA and DNA.

**Polysaccharide** — complex carbohydrates of high molecular weight; one of a group of carbohydrates that upon hydrolysis yields more than two molecules of simple sugars.

**Polysome (polyribosome)** — an mRNA associated with a series of ribosomes engaged in translation.

**Polyyne** — see carbyne.

**Portal** — in anatomy, pertaining to any porta or entrance to an organ, but especially to the portal vein through which nutrient-laden blood is carried from the gastrointestinal tract to the liver for filtration.

**Positional navigation** — in medical nanorobotics, a form of nanorobotic navigation in which nanodevices know their exact location inside the human body to ~micron accuracy continuously at all times.

**Positional synthesis** — in molecular nanotechnology, control of chemical reactions by precisely positioning specific moieties; possibly a basic component of molecular assemblers.

**Positron Emission Tomography (PET)** — see Computerized Tomography.

**Posterior** — the backside of the human body; the backside of something.

**Postpartum** — after childbirth.

**Potential energy** — the energy associated with a configuration of particles, as distinct from their motions.

**Power** — the rate at which energy is produced or consumed.

**Precession** — movement of the axis of rotation of a fixed-pivot gyroscope under the influence of gravity, describing a conical surface.

**(to) Present** — used as verb in medicine, meaning “to exhibit” a symptom or a condition (a medical idiom).

**Presentation semaphore** — in medical nanorobotics, a mechanical device used to display specific antigens, chemical ligands, or other molecular objects to the external environment, with the purpose of selectively modifying the chemical or other surface characteristics of a nanorobot exterior.

**Process** — used as a noun in cell biology and medicine; a projection or outgrowth of bone or tissue.

**Prognosis** — a judgement or forecast, based upon a correct diagnosis, of the future course of a disease or injury, and the patient’s prospects for partial or full recovery.

**Prokaryote** — in microbiology, an organism or cell that lacks a nucleus.
Prometaphase — stage of mitosis (cell division) beginning with a fragmentation of the nuclear membrane and marked by kinetochore formation on the centromeres of the condensed chromosomes.

Promoter — a region of DNA involved in the binding of RNA polymerase to initiate transcription.

Pronouncement — a legally sufficient assertion of the biological death of a patient, usually attested by a physician.

Prophase — stage of mitosis (cell division) during which the chromosomes begin to condense within the still-intact nucleus, the nucleolus disperses, and the mitotic spindle begins to develop.

Prophylaxis — that which helps to prevent disease.

Proprioception — sensations of position and movement, including kinesthesia and the vestibular senses.

Protease — a class of enzymes that break down, or hydrolyze, the peptide bonds that join the amino acids in a protein.

Proteasomes — an extra-lysosomal ATP-driven system for selectively degrading endogenous ubiquinated proteins in virtually all human cells.

Protein — a long chain of amino acids joined by amide bonds, exceeding 100 residues in length; shorter chains are peptides. More generally, living cells contain many molecules that consist of amino acid polymers folded to form more-or-less definite three-dimensional structures, termed proteins. Short polymers lacking definite three-dimensional structures are termed peptides. Molecular objects made of protein form much of the molecular machinery of living cells.

Protein engineering (protein design) — design and construction of new artificial proteins with desired functions; a possible enabling technology for molecular manufacturing.

Proteolytic — hastening the hydrolysis (breakdown) of proteins, usually by enzyme action, into simpler substances.

Proteomics — the study of proteins and how they affect the human body.

Protoplasm — a thick, viscous colloidal substance that constitutes the physical basis of all living activities; the entire contents of a living cell.

Protozoa — the simplest animals, mostly unicellular although some are colonial.

Proximal — near the source or point of attachment or origin; in the extremities, closer to the trunk.

Proximal probes — a generic name for devices that exploit interactions between a sample and a sharp probe tip; a family of devices capable of fine positional control and sensing, including scanning tunneling microscopes and atomic force microscopes; a possible enabling technology for molecular manufacturing.

Pseudopod — in microbiology, a temporary protruding protoplasmic process in protozoa for the purpose of taking up food and aiding in locomotion.

Pseudostratified — apparently composed of layers.

Psychosomatic — pertaining to the influence of the mind or of higher functions of the brain upon the functions of the body, especially in relation to bodily disorders or disease.

Puerperal — relating to the period after childbirth.

Pulmonary — pertaining to the lungs.

Purgative — an agent used for purging the bowels.

Purine — parent molecule of a group of heterocyclic compounds (e.g. adenine, caffeine, guanine, uric acid, xanthine), which have fused five- and six-member carbon-nitrogen rings. Purines are the end products of nucleoprotein digestion; they may be synthesized in the body, and break down to form uric acid.

Pyemia — septicemia due to pyogenic (pus-forming) organisms causing multiple abscesses.

Pylorus — the lower orifice of the stomach, opening into the duodenum.

Pyrimidine — parent molecule of a group of heterocyclic compounds (e.g. cytosine, thymine, uracil), which have a six-member carbon-nitrogen ring.

Pyroelectric — producing electrical energy directly from heat.

Q (resonator or circuit) — in engineering, a measure of the sharpness of a resonance peak.

Quantum computer — a computer that relies upon quantum effects such as superposition and interference.

Quantum confinement — restriction of particle to small physical regions so that quantum effects are exhibited. See also tunneling.

Quantum dot — a zero-dimensional quantum system.

Quantum dot lasers — lasers that exploit the energy levels of quantum dots.

Quantum Hall effect — phenomenon in which certain semiconductors at low temperatures display quantized resistance.

Quantum mechanics — theory of matter and radiation in which certain physical properties are quantized.

Quantum uncertainty — a measurement indeterminacy associated with the dual wave/particle nature of matter; e.g. the Heisenberg uncertainty principle.

Qubits — quantum bits; the equivalent of bits, for quantum computers.
Racemic (mixture) — a mixture of two optical isomers in equal amounts; a racemic mixture does not show optical activity.

Radiolarian — an order of single-celled sea animals with long slender pseudopodia and a perforated outer skeleton of silica.

Radionuclide — a radioactive isotope.

Rale — an abnormal sound heard upon auscultation of the chest. Rales are produced by the passage of air through bronchi that contain secretion or exudate, or that are constricted by spasm or a thickening of bronchial walls.

RAM — random access memory (a data storage device for computers).

Raoult’s law — in an ideal solution, molar weights of non-volatile non-electrolytes, when dissolved in a definite weight of a given solvent under the same conditions, lower the solvent’s freezing point, elevate its boiling point, and reduce its vapor pressure equally for all such solutes.

Ratchet — a hinged catch (pawl) arranged so as to engage with a toothed wheel or bar whose teeth slope in one direction, thus imparting forward movement and preventing backward movement.

RBC — red blood cell (erythrocyte).

RC time — the characteristic exponential decay time of an RC (resistor-capacitor) circuit; the capacitive time constant of a circuit.

Reaction — in chemistry, a process that transforms one or more chemical species into others; typical reactions make or break bonds, while others change the state of ionization or other properties taken to distinguish chemical species. In medicine, a response of a living organism to a chemical, mechanical, or other stimulus.

Reagent — a chemical species that undergoes change as a result of a chemical reaction.

Receptor — most generally, a structure that can capture a molecule (often of a specific type in a specific orientation) owing to complementary surface shapes, charge distributions, and so forth, without forming a covalent bond. In biology, a receptor is a transmembrane protein, located in the plasma membrane, that binds a ligand in a domain on the extracellular side, and as a result has a change in activity of the cytoplasmic domain of the protein.

Recombination — in genetics, the inclusion of a chromosomal part or extrachromosomal element of one microbial strain in the chromosome of another; the interchange of chromosomal parts between different microbial strains. Most generally, the joining together of gene combinations in the offspring that were not present in the parents.

Red blood cell (RBC) — see erythrocyte.

Redox — see oxidation-reduction reaction.

Reduced mass — many dynamical properties of a system consisting of two interacting masses m₁ and m₂ are equivalent to those of a system in which one mass is fixed in space and the other has a mass (the reduced mass) with the value m₁ m₂ / (m₁ + m₂). The reduced mass description has fewer dynamical variables.

Reference man — a man 22 years of age, weight 70 kg, engaged in light physical activity in a 20˚C environment, consuming ~2800 Kcal/day.

Reference woman — a woman 22 years of age, weight 58 kg, engaged in light physical activity in a 20˚C environment, consuming ~2000 Kcal/day.

Refractive index — in optics, the ratio of the sine of the angle of incidence to the sine of the angle of refraction when light is transmitted through a refracting substance (that bends the light beam).

Regioselectivity — selection of the precise position of a ligand group addition to a molecule, e.g. ortho-, meta-, and para- positions in substituted benzenes.

Register — a temporary storage location for an array of bits of data within a digital logic system.

Relaxation time — the measure of the rate at which a disequilibrium distribution decays toward an equilibrium distribution.

Renal — pertaining to the kidney.

Replicator — any system that can build copies of itself when provided with the appropriate raw materials and energy.

Repression — the ability of bacteria to prevent synthesis of certain enzymes when the products of those enzymes are present; more generally, the inhibition of transcription (or translation) by the binding of repressor protein to a specific site on DNA or mRNA.

Repressor protein — a protein molecule that binds to an operator on DNA or RNA to prevent transcription or translation, respectively.

RES — see reticuloendothelial system.

Resection — the partial excision of a bone or other structure.

Residue — in biochemistry, a portion of a single amino acid or nucleotide, linked as part of a peptide/protein or polynucleotide polymer chain, respectively.

Resistor — in electrical engineering, an electronic device that resists the flow of current.

Resonance — in engineering and in control theory, the tendency of an object or system to vibrate or oscillate at a characteristic frequency.

Resonant tunneling — increased probability of quantum tunneling under favorable resonance conditions.

Respirocrit — in medical nanorobotics, the volume-fraction or bloodstream concentration of respirocyte nanorobots, expressed as a percentage.

Respirocyte — in medical nanorobotics, a theorized bloodborne spherical 1-micron (nanorobotic) device having a 1000-atm pressure vessel
with active pumping powered by endogenous serum glucose, that serves as a mechanical artificial red blood cell.1400

**Reticular** — relating to a reticulum.

**Reticulum** — a fine network formed by cells, or formed of certain structures within cells, or formed of connective tissue between cells.

**Reticuloendothelial system (RES)** — in anatomy, the network of fixed and mobile phagocytes that engulf (and dispose of) foreign antigens and cell debris found inside the human body.

**Reticuloendothelium** — tissue of the reticuloendothelial system (RES); the system of mononuclear phagocytes located in the reticular connective tissue of the body that is responsible for the phagocytosis of damaged or old cells, cellular debris, foreign substances, and pathogens, removing them from the circulation.

**Reticuloplasm** — the fluid that fills the lumen of the endoplasmic reticulum.

**Reticulitis** — inflammation of the retina of the eye.

**Reversible computer** — a computer capable of reversing its computational steps (after storing the final answer), thus recovering most of the energy that was employed during the computation.

**Reynolds number** — the ratio of inertial to viscous forces in fluid flow. Macroscopic objects and flows typically experience Reynolds numbers >> 1, where mass and inertia dominate object motions; microscopic and especially nanoscale objects and flows typically experience Reynolds numbers << 1, where the viscosity of the environment dominates object motions.

**rf** — radio frequency.

**Rheology** — study of the deformation and flow properties of materials, especially fluids, such as blood.

**Ribonucleic acid (RNA)** — the ribonucleotide polymer into which DNA is transcribed.

**Ribosomal RNA (rRNA)** — the RNA components of ribosomes.

**Ribosome** — a naturally occurring molecular machine that manufactures proteins according to instructions derived from the cell’s genes; a cytoplasmic ribonucleoprotein complex that serves as the site of translation in the cell. Each ribosome has a large and a small subunit, 60S and 40S in eukaryotes. These subunits dissociate and reassociate in a cycle related to their functions, during translation.

**Rickets** — a vitamin D deficiency disease characterized by overproduction and deficient calcification of osteoid (bony) tissue, with associated skeletal deformities, enlargement of the liver and spleen, profuse sweating, and general tenderness of the body when touched.

**Rigidity** — see stiffness.

**RMS** — root mean square. Refers to a kind of average molecular speed; the speeds of individual molecules vary over a wide range of magnitude, with a characteristic temperature-dependent distribution of molecular speeds.

**RNA** — see ribonucleic acid.

**RNase** — an enzyme that attacks bonds in RNA.

**RNA polymerase** — an enzyme that synthesizes RNA under direction from a DNA template (formally described as DNA-dependent RNA polymerase).

**Robot** — a programmable device usually consisting of mechanisms for sensing and mechanical manipulation, often connected to (or including) a computer that provides control.

**Rough ER** — regions of endoplasmic reticulum that are associated with ribosomes.

**Rouleaux** — stack-of-coins configuration of a cluster of red blood cells.

**rRNA** — see ribosomal RNA.

**Rugosity** — condition of being folded or wrinkled; surface roughness.

**Sacculation** — formation into a sac or sacs; group of sacs, collectively.

**Saccule** — the smaller of the two membranous sacs in the vestibule (central cavity) of the osseous labyrinth (the inner ear).

**Sagittal** — in anatomy, a vertical plane or section that divides the body into right and left portions.

**Salt bridge** — in biochemistry, an ionic bond between charged groups that are part of larger covalent structures; salt bridges occur in many proteins.

**SAM** — in microscopy, a Scanning Acoustic Microscope; in polymer chemistry, a self-assembled monolayer or a spontaneously assembled monolayer.

**Sanguinatation** — in medical nanorobotics, locomotion (especially swimming by a nanorobot) through the bloodstream.

**Sapphirophagy** — eating sapphire (corundum).

**Sarcoplasmic reticulum** — specialized, elaborate smooth ER, found in muscle cells.

**Saturated** — a molecule which is a closed-shell species lacking double or triple bonds; forming a new bond to a saturated molecule requires the cleavage of an existing bond.

**Scanning Electron Microscope (SEM)** — a kind of microscope that uses electrons rather than light.

**Scanning Force Microscope** — see Atomic Force Microscope.

**Scanning Probe Microscope (SPM)** — a class of microscope (including AFMs and STMs) that exploits physical interactions that are sensitive to the separation of a sample and a sharp tip.

**Scanning Tunneling Microscope (STM)** — an instrument able to image conducting surfaces to atomic accuracy. A sharp conductive tip is moved across a conductive surface close enough (typically a
nanometer or less) to permit a substantial tunneling current. In a common mode of operation, a constant voltage is established and the current is monitored and kept constant by controlling the height of the tip above the surface using a feedback circuit, ultimately producing an atomic-resolution map of the surface reflecting a combination of topography and electronic properties. Increasing the voltage enables a researcher to move atoms around, pile them up, or trigger chemical reactions.

Schistosomiasis — a parasitic disease due to infestation with blood flukes; endemic throughout Asia, Africa, and tropical America.

Sclerosis (sclerotic) — hardening of a tissue or organ, especially due to excessive growth of fibrous tissue; also, thickening and hardening of the tissue layers comprising the walls of an artery.

Scurvy — a vitamin C deficiency disease characterized by hemorrhagic manifestations and abnormal formation of bones and teeth, including spongy condition of the gums, sometimes with ulceration.

Self-assembly — see Brownian assembly.

Semi-circular canals — superior, posterior, and inferior passages forming part of the inner ear.

Semilunar — shaped like a crescent.

Sepsis — the presence of various pus-forming and other pathogenic organisms, or their toxins, in the blood or tissues.

Septic — pertaining to or caused by sepsis.

Septicemia — septic fever; systemic disease caused by the multiplication of microorganisms in the circulating blood.

Septum — a wall dividing two cavities (e.g. the interalveolar septa between the alveoli).

Serotonin — a biochemical substance, 5-hydroxytryptamine (5-HT), that is present in platelets, gastrointestinal mucosa, mast cells, and in carcinoid tumors. Serotonin is a potent vasoconstrictor involved in neural mechanisms important in sleep and sensory perception.

Serous membrane — a membrane lining a serous cavity, specifically the pleural (lung), peritoneal (abdominal), and pericardial (heart) cavities.

Serum — the watery portion of the blood after coagulation; a fluid found when clotted blood is left standing long enough for the clot to shrink. More generally, any serous fluid, especially the fluid that moistens the surfaces of serous membranes.

Sessile — incapable of voluntary movement. Opposite of motile.

Shear — a shear deformation is one that displaces successive layers of a material transversely with respect to one another, like a crooked stack of cards. Shear is a dimensionless quantity measured by the ratio of the transverse displacement to the thickness over which it occurs.

Shear modulus — shear stress divided by shear strain (units of force per unit area).

Shirr — a series of parallel rows of short, running stitches with gathers between rows.

Sinusoidal node — a node at the junction of the superior vena cava with the right cardiac atrium (one of the upper chambers of the heart), regarded as the starting point of the heartbeat.

Sinusoid — resembling a sinus (a cavity having a relatively narrow opening); a minute blood vessel found in such organs as the liver, spleen, adrenal glands, and bone marrow, that is slightly larger than a capillary and has a lining of reticuloendothelium.

Site-directed mutagenesis — see mutagenesis.

Smooth ER — in cell biology, regions of endoplasmic reticulum that are devoid of ribosomes.

SNR — signal-to-noise (power) ratio.

Solenoid — in physics and electrical engineering, a coil of wire carrying an electric current; an electromagnet.

Solute — the dissolved substance in a solution.

Solvation — the process of dissolving a solute in a solvent, making a solution.

Solvent — a substance, usually a liquid, that can hold another substance in solution.

Somatic — in general, relating to the body, as opposed to the mind or soul; corporeal.

Somatic cell — in cell biology, all the cells of an organism except those of the germ line (gametes).

Somatography — a physical description (and mapping) of the human body.

Somesthesis — sensations from the skin and viscera, including pressure, warmth, cold, and pain.

Sonication — to bombard with high-energy acoustic waves, often for the purpose of fragmenting or destroying the sonicated object.

Sonolucent — in ultrasonography, the condition of not reflecting the ultrasound waves back to their source.

Sonoluminescence — generation of visible light during high-powered sonication of a fluid, usually water.

Sortation — the act of separating or sorting, especially of molecular species.

Sorting rotor — see molecular sorting rotor.

Species — in chemistry, a distinct kind of molecule, ion, or other structure.
**Specific gravity** — relative density, as a ratio to the density of pure water.

**Specificity** — in biochemistry, defines the degree to which a receptor can distinguish between similar ligands.

**Specular reflection** — in physics, reflection as from a mirror, requiring that the wavelength of the reflected wave is smaller than the characteristic dimension of the reflector.

**Sphincter** — in anatomy, circular muscle constricting an orifice. In normal tonic (i.e. under tension or in contraction) condition, it closes the orifice; the muscle must relax in order to allow the orifice to open.

**Spindle** — in cell biology, the reorganized structure of a eukaryotic cell that is passing through mitosis (cell division); the nucleus has been dissolved and chromosomes are attached to the spindle by microtubules.

**Splenic** — pertaining to the spleen.

**SPM** — see Scanning Probe Microscope.

**Sputum** — substance expelled by coughing or clearing the throat.

**Squamous cell** — a flat, scaly epithelial cell.

**SQUID** — superconducting quantum interference device.

**Stable** — in the physical sciences, a system is termed stable if no rearrangement of its parts can produce a system having lower free energy; in biology, a living system is stable so long as homeostasis can be maintained.

**Stationkeeping** — active maintenance of location in a specific physical place, area, or volume.

**Stellate** — star-shaped.

**Steric** — pertaining to the spatial relationships among atoms in a molecular structure; in particular, pertaining to the space-filling properties of a molecule.

**Steric hindrance** — in chemistry, slowing of the rate of a chemical reaction owing to the presence of molecular structures possessed by the reagents that mechanically interfere with the motions associated with the reaction, typically by obstructing the reaction site; in hemodynamics, the reduction in hematocrit near small blood vessel bifurcations due to the elongation and orientation of red cells along the direction of shear flow.

**Stereochromy** — the branch of chemistry concerned with the three-dimensional spatial relationships among atoms in molecules and the effects of such relationships on the properties (especially optical rotation) of molecules.

**Stereoisomers** — isomeric forms that have different configurations in space; may be geometric (non-enantiomeric, achiral) or enantiomeric (chiral).

**Steroids** — a large family of chemical substances, comprising many hormones, vitamins, body constituents, and drugs, each containing the tetracyclic cyclopentophenanthrene skeleton.

**Stewart platform** — in engineering, a manipulator employing a 6-DOF mobile platform.

**Sticky ends** — in biochemistry, complementary single strands of DNA that protrude from opposite ends of a duplex or from ends of different duplex (two-stranded) molecules; can be generated by staggered cuts in duplex DNA.

**Stiffness** — in mechanical engineering, the stiffness of a system with respect to a deformation (e.g. the stiffness of a spring with respect to stretching) is the second derivative of the energy with respect to the corresponding displacement. Positive stiffness is associated with stability, and a large stiffness can result in a small positional uncertainty in the presence of thermal excitation. Negative stiffnesses correspond to unstable locations on the potential energy surface.

**Stirling engine** — invented by Robert Stirling in 1816; generates power not by burning fuels explosively in an otherwise empty cylinder, but by heating and cooling working gas from the outside of a cylinder which always contains a pressurized working gas.

**STM** — see Scanning Tunneling Microscope.

**Stoichiometric** — in chemistry, pertaining to the precise quantities of reagents required to complete a chemical reaction; in particular, to the exact amounts needed to balance the chemical reaction equation.

**Strain** — in mechanical engineering, strain (a dimensionless quantity) is a measure of the deformation resulting from stress (an applied force per unit area); the displacement of one point with respect to another, divided by their equilibrium separation in the absence of stress. In chemistry, a molecular fragment generally has some equilibrium geometry (e.g. bond lengths, interbond angles) when the rest of the molecular structure does not impose special constraints (e.g. bending bonds to form a small ring); deviations from this molecular equilibrium geometry are described as strain, and increase the energy of the molecule. Strain in the mechanical engineering sense also causes strain in the chemical sense.

**Stratum corneum** — in anatomy, the outermost (horny) layer of the epidermis, consisting of many layers of flat, keratinized, enucleated cells.

**Streamline** — the movement or flow of a small portion of fluid, especially in relation to a solid body which lies in the path of this flow. Compare turbulence.

**Stress** — in mechanical engineering, the force per unit area applied by one part of an object to another. Pressure is an isotropic compressive stress. Suspending a mass from a fiber places the fiber in tensile stress. Gluing a layer of rubber between two plates and then sliding one plate over the other (while holding their separation constant) places the rubber in shear stress. Twisting one end of a rod while holding fixed the other end places the rod in torsional stress.
**Striated** — striped; marked by streaks or striae (lines or bands elevated above or depressed below surrounding tissue, or differing in color or texture).

**Stroma** — foundation-supporting tissues of an organ, defining the framework of an organ; opposite of parenchyma.

**Sublimation** — in chemistry, passing directly from solid to vapor state.

**Submaxillary** — below the maxilla (upper jaw).

**Substrate** — in biology, the substance acted upon and changed by an enzyme; in replicator theory, the set of all material inputs to the replication process.

**Superconductivity** — in physics, a physical phenomenon in which the electrical resistance of certain materials drops to zero at moderate or low temperatures.

**Superficial** — near the surface of the skin, as opposed to inside the body.

**Superior** — upper or higher than; situated above something else. See also cephalic.

**Supine** — lying on the back, with the face up.

**Suppuration** — producing or associated with the generation of pus, often involving an infection with pyogenic (pus-forming) bacteria.

**Supraglottal** — located above the glottis.

**Supramolecular chemistry** — the study of interactions between molecules; chemistry beyond the molecule; the chemistry of the noncovalent bond.

**Surfactant** — in physical chemistry, a chemical agent that lowers surface tension.

**Sympathetic (nervous system)** — a major component of the autonomic nervous system, consisting of ganglia, nerves (mostly motor, some sensory), and plexuses that supply the involuntary muscles.

**Synapse** — the point of junction between two neurons in a neural pathway, where the termination of the axon of one neuron comes into close proximity with the cell body or dendrites of another neuron.

**Synovial membrane** — a membrane lining the capsule of a joint.

**Synthesis** — in chemistry, the production of a specific molecular structure by a series of chemical reactions.

**System** — in engineering usage, a set of components working together to serve a common set of purposes.

**Systems theory** — in clinical medicine, an approach that considers the human being as a whole as opposed to his parts; human beings are considered as open systems constantly exchanging information, matter, and energy with the environment.

**Systole** — the normal period in the heart cycle during which the muscle fibers tighten and shorten, the heart constricts, and the cavities empty of blood; roughly, the period of contraction alternating with diastole or relaxation. Occurs in the interval between the first and second heart sounds during which blood is surged through the aorta and pulmonary artery.

**Tachometer** — a device for measuring the speed of rotation.

**Tachyatria** — the art of curing quickly.

**Teleoperation** — remote control.

**Telepresence** — in control engineering, teleoperation with full sensor feedback (see virtual reality system).

**Telomerase** — in biochemistry, the ribonucleoprotein enzyme that creates repeating units of one strand at the telomere, by adding individual bases.

**Telomere** — the natural end of a chromosome; the telomeric DNA sequence consists of a simple repeating unit (in humans, TTAGGG) with a protruding single-stranded end that may fold into a hairpin.

**Telophase** — the final stage of mitosis (cell division), in which chromosomes at opposite spindle poles begin to decondense as spindle fibers disappear and the nuclear membrane reforms.

**Temperature** — a system in which internal vibrational modes have equilibrated with one another can be said to have a particular temperature; two systems are at different temperatures if heat flows between them when they are brought into physical contact. The most common measurement units of temperature are Centigrade or Celsius (°C), Fahrenheit (°F), and Kelvin (K).

**Temporal** — pertaining to time.

**Tensile** — pertaining to extension or stretching.

**Teragravity** — one trillion times normal terrestrial gravity.

**Tessellation** — in physical geometry, laid out or paved in a mosaic pattern of small squares or blocks, usually with the objective of completely covering a surface or completely filling a volume.

**Tetanic** — in medicine, tending to produce spasms characteristic of tetanus.

**Tether** — a cable, tube, or other physical linkage that allows information, matter, or energy to flow between two or more objects.

**Thermal conductivity** — transport of thermal energy due to a temperature gradient; the energy flux (W/m²) per unit of spatial temperature gradient (K/m) equals the coefficient of thermal conductivity (W/m-K).

**Thermal energy** — the internal energy present in a system as a result of the energy of thermally equilibrated vibrational modes and other motions (including both kinetic energy and molecular potential energy); the mean thermal energy of a classical harmonic oscillator is kT.
Thermal expansion coefficient — the rate of change of length with respect to temperature for a particular material.

Thermogenesis — the production of heat, especially in the body (e.g. by shivering).

Thermogenic limit — in medical nanorobotics, the maximum amount of waste heat that may safely be released by a population of in vivo medical nanorobots that are operating within a given tissue volume.

Thermography — temperature mapping of the human body.

Thiol — in chemistry, an -SH group, or a molecule containing such a group; also known as a sulfhydryl or mercapto group.

Thoracic — in anatomy, pertaining to the chest or thorax.

Thorax — that part of the body between the base of the neck superiorly and the diaphragm inferiorly.

Thrombus — blood clot.

THz — terahertz; trillions of cycles per second.

Tight-receptor structure — in molecular nanotechnology, a molecular receptor structure in which a bound ligand of a particular kind is confined on all sides by repulsive interactions. A tight-receptor structure discriminates strongly against all molecules larger than the target.

Tinnitus — a subjective ringing or tinking sound in the ear.

Titer — in analytical chemistry, the standard of strength of a volumetric test solution; assay value of an unknown measure by volumetric means.

T-lymphocytes (T cells) — White blood cells that are produced in the bone marrow but later mature in the thymus. T cells are important in the body’s defense against certain bacteria and fungi, help B cells make antibodies, and assist in the recognition and rejection of foreign tissues.

Tomography — a noninvasive imaging technique designed to show detailed images of structures in a selected plane of tissue by blurring images of structures in all other planes.

Tonic (osmotic) — see isotonic.

Tonic (muscular) — refers to muscles arranged around hollow structures that contract slowly and can hold for a long period of time. Compare phasic.

Tonic (physiological) — pertaining to or characterized by tension or contraction; in a state of continuous action, denoting especially a muscular contraction.

Top-down — an approach to nanotechnology that aims to construct nanodevices by progressive miniaturization of existing bulk components.

Torque — in physics, a force that produces torsion.

Torsional — relating to, producing, or resulting from twisting.

Toxic shock — a disease caused by the release of toxins produced by certain strains of various bacteria.

Trabecular — inner part of organ or body.

Transcription — synthesis of RNA on a DNA template.

Transcutaneous (percutaneous) — effected through the skin.

Transdermal — through the skin.

Transducer — any mechanism or device that can convert energy or signals from one physical form to another.

Transduction — in physics and engineering, the conversion of energy or signals from one form to another; in biology and biotechnology, a phenomenon causing genetic recombination in bacteria in which DNA is carried from one bacterium to another, usually by a bacteriophage.

Transendothelial migration — see diapedesis.

Transfer RNA (tRNA) — the adaptor RNA that carries amino acid residues into polypeptide linkage during translation, as the codon sequence in mRNA is read.

Transgenic organism — in biotechnology, an organism modified by the insertion of foreign genetic material into its germ line cells. Recombinant DNA techniques are commonly used to produce transgenic organisms.

Translation — in biochemistry, the synthesis of protein on the mRNA template; the process of reading the codon sequence in mRNA to synthesize the corresponding polypeptide with the involvement of ribosomes, tRNA, and many enzymes.

Transmembrane — crossing the plasma membrane of a living cell.

Transmembrane protein — in cell biology, a membrane component, wherein a hydrophobic region or regions of the protein resides within the membrane, and hydrophilic regions are exposed on one or both sides of the membrane. See integral membrane protein.

Transtegumental — crossing or passing through the skin or covering of a body.

Transvenue outmessaging — in medical nanorobotics, outmessaging from the nanorobots present within the body of one patient to the nanorobots present in another (physically separated) human body.

Transverse — horizontal, hence at right angles to both sagittal and coronal sections, dividing the body into upper and lower portions; more generally, extending crossways or from side to side.

Triage — in emergency medicine, the screening of sick and wounded patients during war, disasters, or other emergencies. During triage, patients are classified into one of three groups: (1) those who will not survive even if given treatment, (2) those who will recover without
treatment, and (3) those who need treatment in order to survive (the priority treatment group).

**Tribology** — the study of friction.

**Trichinosis** — a pathogenic disease caused by ingestion of *Trichinella spiralis* in raw or insufficiently cooked infected pork.

**Trillion** — this book follows the American convention in which a trillion is $10^{12}$.

**Trimer** — in chemistry, a compound of three similar or identical molecules.

**Triple point** — a single condition of temperature and pressure in which three distinct phases of a substance (e.g. solid, liquid, and gas) can exist simultaneously and in equilibrium.

**tRNA** — see transfer RNA.

**Trypanosomiasis** — any disease caused by trypanosomes (asexual protozoan flagellates parasitic in the blood plasma of many vertebrates).

**Tubulin** — a protein present in the microtubules of cells, which are polymers of α-tubulin (~53,000 dalton) and β-tubulin (~55,000 dalton) dimers.

**Tunneling** — in quantum physics, a probabilistic effect in classically-forbidden transitions that is a consequence of the wave nature of matter. In classical physics, a particle or system cannot penetrate regions of negative energy (e.g. barrier regions in which the potential energy is greater than the system energy). In quantum physics, a wave function of significant amplitude may extend into and beyond such regions; if the wave function extends into another region of positive energy, the barrier is crossed with some nonzero probability, a process called tunneling (since the barrier is penetrated rather than climbed).

**Turbulence** — in hydrodynamics, fluid flow which does not follow parallel streamlines, which has a blunt (nonparabolic) profile in tube flow, and often involves eddies, vortices, and significant variations in fluid velocities, accelerations and shear stress between adjacent fluid elements. Turbulence dissipates more energy, and presents more resistance to flow, than laminar flow.

**Turgid** — swollen.

**Turing machine** — a programmable computing device.

**Tympanic** — pertaining to the tympanum (the middle ear or tympanic cavity; eardrum).

**Ubiquitin** — a small protein present in eukaryotic cells that combines with other proteins and makes those other proteins susceptible to destruction; this protein is also important in promoting the functions of proteins that make up the ribosomes.

**Ullage** — the amount by which a fluid container falls short of being full.

**Ultrasonic** — inaudible (to human ears) sounds with frequencies greater than ~20,000 Hz.

**Unsaturated** — possessing double or triple bonds; capable of making additional covalent bonds.

**Urethra** — a canal for the discharge of urine extending from the bladder to the outside of the body.

**Urticaria** — vascular reaction of the skin characterized by the eruption of pale evanescent wheals (round elevations of the skin, white in the center with a pale red periphery), which are associated with severe itching; hives.

**UTC** — see coordinated universal time.

**Utricle** — the larger of the two membranous sacs in the vestibule (central cavity) of the osseous labyrinth (the inner ear).

**UV** — ultraviolet.

**Uveitis** — a nonspecific term for any intraocular inflammatory disorder, usually of the uveal tract structures (iris, ciliary body, and choroid, forming the pigmented layer) although nonuveal parts such as the retina and cornea may also be involved.

**Valence** — in chemistry, regarding covalent compounds, the valence of an atom is the number of bonds that the atom forms to other atoms. In immunology, antiserum contains antibodies that may have specificity for one (monovalent) or more (polyvalent) antigens.

**Vallate papilla** — one of a group of papillae (small nipple-like protruberances or elevations) forming a V-shaped row on the posterior dorsal surface of the tongue.

**Van der Waals forces** — weak electrostatic forces between atoms and molecules; any of several intermolecular attractive forces not resulting from ionic charges; also known as the London dispersion force.

**Vascular** — containing, or pertaining to, blood or lymph vessels.

**Vasculography** — a physical description (and mapping) of the human vascular system.

**Vasoconstriction** — in physiology, a decrease in the diameter of blood vessels.

**Vasodilation** — in physiology, an increase in the diameter of blood vessels.

**Vaults** — in cell biology, barrel-shaped protein particles found in the cytoplasm, that are of the right size size and shape to be able to dock at the nuclear pore complex.

**Vein** — in anatomy, generally, a blood vessel that returns blood from the tissues to the heart.

**Velum palatinum** — the soft palate.

**Vena cava** — the largest vein in the human body, leading toward the heart.

**Venesection** — blood-letting (phlebotomy, venotomy).
Ventral — the front of the human body, hence on or nearest the abdominal surface; pertaining to the belly.

Ventricle — either of two lower chambers of the heart.

Ventricular fibrillation — the primary mechanism and arrhythmia seen in sudden cardiac arrest (cessation of blood circulation).

Vernier — a short, graduated scale, permitting accurate measurement of the relative movement of mechanical parts.

Vertigo — the sensation of moving around in space; sometimes used as a synonym for dizziness, lightheadedness, or giddiness.

Vesicles — small bodies bounded by membrane, derived by budding from one membrane and often able to fuse with another membrane.

Vesicles (endocytotic) — membranous particles that transport proteins through endocytosis; also known as clathrin-coated vesicles, having on their surfaces a layer of the protein clathrin.

Vesicles (exocytotic) — membranous particles that transport and store proteins during exocytosis.

Vestibular senses — sensory receptors in the labyrinth of the human inner ear that respond to human head position and movement, imparting a sense of balance.

Virion — a physical virus particle.

Virtual reality system — a combination of computer and interface devices (goggles, gloves, etc.) that presents a user with the illusion of being in a three dimensional world of computer-generated objects. These three dimensional environments and force-feedback systems can aid in the visualization of complex molecules, and in telepresence systems.

Virucide — the destruction of active or dormant virus particles.

Virus — A parasite (consisting primarily of genetic material enclosed in a protein capsid shell) that invades cells and takes over their molecular machinery in order to copy itself.

Vitamins (biology) — a group of organic substances, present in minute amounts in natural foodstuffs, that are essential to normal metabolism, growth, and development of the body. Vitamins serve principally to regulate metabolic processes and to play a role in energy transformations, usually acting as coenzymes in enzymatic systems. See also coenzyme.

Vitamins (engineering) — in machine replication theory, vitamin parts are components of a self-replicating machine which the machine is incapable of producing itself, therefore these vital parts must be supplied from an external source.

Vitreous humor — in anatomy, a delicate network enclosing in its meshes a clear watery fluid filling the interior of the eyeball behind the lens.

Volitional normative model of disease — in medical nanorobotics, disease is said to be present in a human being upon either: (1) the failure of optimal physical (e.g. biological) functioning, or (2) the failure of desired (by the patient) functioning.

Vomeronasal organ — in anatomy, a small tubular epithelial sac lying on the anteroinferior surface of the nasal septum.

Voxel (volume pixel) — abbreviation for volume element; a single cubic cell within a three-dimensional geometric solid grid or array.

W — watt (MKS unit of power).

Watt governor — in control engineering, a governor is a mechanism by which the speed of an engine, turbine, wheel, or motor may be regulated. The Watt governor employs rotating weights attached to a pivoted lever arm; the weights extend as the pair rotates faster, depressing the attached lever arm and throttling the motor down, thus slowing the rotation, and vice versa.

Wave function — in quantum mechanics, a complex mathematical function extending over the configuration space of a material system.

WBC — white blood cell (see leukocyte).

White blood cell (WBC) — see leukocyte.

Whitlow — in medicine, a suppurative inflammation at the end of a finger or toe.

Work — in physics, the energy transferred by applying a force across a distance; for example, lifting a mass does work against gravity and stores gravitational potential energy.

WWV — a radio station that continuously broadcasts the exact Coordinated Universal Time.

Xenotransplantation — transplantation of cells or organs from an organism of a different species.

Xerophthalmia — a vitamin A deficiency disease, characterized by extreme dryness of the conjunctiva with keratinization of epithelium.

Young's modulus — in mechanical engineering, a modulus relating tensile (or compressive) stress to strain in a rod that is free to contract or expand transversely. The relevant measure of strain is the elongation divided by the initial length (see also strain and stress).

Zippocytes — in medical nanorobotics, a theorized medical nanorobot that can rapidly perform incision-wound repairs to the dermis and epidermis; dermal zippers.

Zwitterions — dipolar ions that contain positive and negative charges of equal strength, and are therefore not attracted to either
anode or cathode; in a neutral solution, some amino acids function as zwitterions.  

Zymogenic — pertaining to a substance (a zymogen or proenzyme) that develops into an enzyme capable of producing or causing fermentation or digestion (e.g., pepsinogen, trypsinogen); a cell that produces zymogens (proenzymes).
Basic Capacities • References


1694. H. Eyring, “The activated complex and the absolute rate of chemical reactions,”


1696. Thomas Rufer Barnard Robinson, Francis Alan Burnett Ward, “Clock,”


2103. Luis del Peso, Maribel Gonzalez-Garcia, Carmen Paez, Roman Herrera, Gabriel Nunez, “Interleukin-3-Induced Phosphorylation of BAD Through the Protein Kinase Ale,” Science 278(24 October 1997):687-689.


3306. A. Milojevic, L.J. Dawson, "Salivary factors in vomiting bulimics with and without
3307. S. Hamlet, J. Choi, M. Zormeier, F. Shamsa, R. Stachler, J. Muz, L. Jones,
3308. R.O. Dantas, W.J. Dodds, "Influence of the viscosity of the swallowed food bolus
3309. J.G. Ruseler-van Embden, L.M. Van Lieshout, D.J. Binnema, M.P. Hazenberg,
3311. M.C. Lin, T.C. Tsai, Y.S. Yang, "Measurement of viscosity of human semen with
3312. S. Hamlet, J. Choi, M. Zormeier, F. Shamsa, R. Stachler, J. Muz, L. Jones,
3314. M.C. Lin, T.C. Tsai, Y.S. Yang, "Measurement of viscosity of human semen with
3315. G.I. Leitch, S.A. Harris-Hooker, I.A. Udezulu, "Movement of
3316. G. Cornelissen, F. Halberg, "Enhanced circasemisepan
3318. H. Engelberg-Kulka, M. Reches, S. Narasimhan, R. Schoulaker-Schwarz, Y. Maibach, "Time-dependent variations of the skin barrier function in humans:
3319. J. Poschl, T. Weiss, C. Diehm, O. Linderkamp, "Periodic variations in skin
3320. H. Engelberg-Kulka, M. Reches, S. Narasimhan, R. Schoulaker-Schwarz, Y. Maibach, "Time-dependent variations of the skin barrier function in humans:
3321. L. Snyder, "Phage-exclusion enzymes: a bonanza of biochemical and cell biology
3322. H. Engelberg-Kulka, M. Reches, S. Narasimhan, R. Schoulaker-Schwarz, Y. Maibach, "Time-dependent variations of the skin barrier function in humans:
3323. K. Okazaki, Y. Yamamoto, K. Ito, "Endoscopic measurement of papillary sphincter
3324. M.C. Lin, T.C. Tsai, Y.S. Yang, "Measurement of viscosity of human semen with
3325. G. Cornelissen, F. Halberg, "Enhanced circasemisepan
3326. M. Blank, O. Denisova, G. Cornelissen, F. Halberg, "Enhanced circasemisepan
3328. M. Miura, J. Okada, "Non-thermal vasodilation by radio frequency burst-type
3329. A. Weydahl, F. Halberg, "Daily spot-checking versus chronobiologic monitoring of
3330. R.J. Seymour, W.E. Lucefield, "Wheelchair cushion effect on pressure and skin
3331. G.D. Callin, "A shower spray facility for accurate control and rapid changes of
3332. S.J. Park, H. Tokura, "Effects of different types of clothing on circadian rhythms of
3334. J.W. Doust, "Periodic homeostatic fluctuations of skin temperature in the sleeping
3335. S.J. Park, H. Tokura, "Effects of different types of clothing on circadian rhythms of
3336. J. Poschl, T. Weiss, C. Diehm, O. Linderkamp, "Periodic variations in skin
3337. K. Krauchi, A. Wirz-Justice, "Circadian rhythm of heat production, heart rate,
3338. J. Poschl, T. Weiss, C. Diehm, O. Linderkamp, "Periodic variations in skin
3339. S.J. Park, H. Tokura, "Effects of different types of clothing on circadian rhythms of
3340. K. Sudo, "Progress in biotechnology and its clinical application to isozyme
3342. S.J. Park, H. Tokura, "Effects of different types of clothing on circadian rhythms of
3343. W.H. Fishman, "Clinical and biological significance of an isozyme tumor marker
3348. J. Huijgen, G.T. Sanders, R.W. Koster, J. Vreeken, P.M. Bossuyt, "The clinical
3350. H.J. Huijgen, G.T. Sanders, R.W. Koster, J. Vreeken, P.M. Bossuyt, "The clinical
3351. N. Calander, K.A. Karlsson, P.G. Nyholm, I. Pascher, "On the dissection of
3353. N. Calander, K.A. Karlsson, P.G. Nyholm, I. Pascher, "On the dissection of
3354. P.V. Bala ji, P.K. Qasba, V.S. Rao, "Molecular dynamics simulations of
3355. N. Calander, K.A. Karlsson, P.G. Nyholm, I. Pascher, "On the dissection of
3356. H.J. Huijgen, G.T. Sanders, R.W. Koster, J. Vreeken, P.M. Bossuyt, "The clinical
3385. D.B. Callerio, P. Revelant, L. Di Filippo, "Cristae of a triangular aspect in the
3386. B. Fernandez, I. Suarez, C. Gianonatti, "Fine structure of astrocytic mitochondria
3388. F.P. Prince, "Mitochondrial cristae diversity in human Leydig cells: A revised look
3387. L.G. Baum, J.C. Paulson, "Sialyloligosaccharides of the respiratory epithelium in the
3390. C. Ruhrberg, F.M. Watt, "The plakin family: versatile organizers of cytoskeletal
3391. J.A. Gutmans, D.J. Mulholland, A.W. Yogi, "Plectin is concentrated at intercellular
junctions and at the nuclear surface in morphologically differentiated rat Sertoli
3392. K. Andra, B. Nikolic, M. Stocher, D. Dreckhahn, G. Wiche, "Not just scaffolding:
3393. B.J. Foerts, J.J. van den Oord, V.J. Desmet, L. Misseren, "Cytokeletal filament
3398. J. Southgate, P. Harnsun, L.K. Tiejdiesiewicz, "Cytokeratin expression patterns in
normal and malignant urothelium: a review of the biological and diagnostic
3400. M.H. Stromer, "The cytoskeleton in skeletal, cardiac and smooth muscle cells,"
3405. L.F. Pemberton, G. Blobel, J.S. Rubenstum, "Transport routes through the nuclear
architecture to functional dynamics," Curr. Opin. Cell Biol. 11(June 1999):391-
401.
3407. H. Ris, "High-resolution field-emission scanning electron microscopy of the
3413. L.G. Kossi, "Characteristics of chromosomes in polarized normal human bronchial
distributions of early and late replicating chromatin in interphase chromosome
3420. M.A. Mancini, D. He, L. Ouspenski, B.R. Brinkely, "Dynamic continuity of nuclear
3607. C.T. Brighon et al., “The biochemical pathway mediating the proliferative
response of bone cells to a mechanical stimulus,” J. Bone Joint Surg. Am. 31(30
“Mechanical strain and its effect on cultured endothelial cells,” Heart Vessels 122
suppress matrix metalloproteinase-1 expression by human vascular smooth muscle
3604. M. Soskhe et al., “Mechanotransduction and intracellular signaling mechanisms of
stretch-induced remodeling in endothelial cells,” Hum. Heart Vessels 122 (2000):
979-985.
3607. C.T. Brighton et al., “The biochemical pathway mediating the proliferative
response of bone cells to a mechanical stimulus,” J. Bone Joint Surg. Am. 31
3611. O. Thoumine, A. Ott, “Time scale dependent viscoelastic and contractile regimes
in fibroblasts probed by microplate manipulation,” J. Cell Sci. 110(Suppl. 1
3626. W.A. Klimke, L.S. Frost, “Genetic analysis of the role of the transfer gene, traN,”
3628. G.M. Dunny, B.A. Leonard, P.J. Hedberg, “Pheromone-inducible conjugation in
3624. A. Tozer en, “Cell-cell, cell-substrate adhesion:  theoretical and experimental
3621. O. Thoumine, A. Ott, “Time scale dependent viscoelastic and contractile regimes
in fibroblasts probed by microplate manipulation,” J. Cell Sci. 110(Suppl. 1
3611. O. Thoumine, A. Ott, “Time scale dependent viscoelastic and contractile regimes
in fibroblasts probed by microplate manipulation,” J. Cell Sci. 110(Suppl. 1
3626. W.A. Klimke, L.S. Frost, “Genetic analysis of the role of the transfer gene, traN,”
3628. G.M. Dunny, B.A. Leonard, P.J. Hedberg, “Pheromone-inducible conjugation in
3624. A. Tozer en, “Cell-cell, cell-substrate adhesion:  theoretical and experimental
3621. O. Thoumine, A. Ott, “Time scale dependent viscoelastic and contractile regimes
in fibroblasts probed by microplate manipulation,” J. Cell Sci. 110(Suppl. 1
INDEX

A
Accelerative onset, 99-100
Accelerometer, 41, 98-100
Accommodation, 191, 203
Accordion configuration, 130
Acetylcholinesterase, 79, 115, 197
Acoustic amplitude absorption coefficients, 117
Acoustic attenuation, 117, 167
Acoustic attenuation, function of temperature, 375
Acoustic communication, 181-182, 184
Acoustic energy, reflection at interfaces, 163
shadows/concentration, 162
transducers, 145-146
Acoustic impedance, 118, 296
values, 118, Appendix A
Acoustic microscopy, 112-113
Acoustic power, available from physiological sources, 146
radiators, 181, 182
safe exposure limits, 161
transmission, 145-146, 160-163, 167
Acoustic sensing, 105-106, 112-113, 116-118
Acoustic threshold, audible, 117, 162
pain, 162, 172
Acoustic transmission line, diamond, 184
Actin, see microfilament
Actin polymerization, 324
Actomyosin motor, see myosin motor
Acyclovir, 366
Addiction modules, 361
Addis, Thomas, 24
Adelman, Leonard M., 350
Adhesion antenna, 291, 371
Adhesion forces, 275-278
Adiabatic demagnetization, 375
Adipocytes, 247, 261
Adociasulfate-2, 328, 363
Adrenalin, see epinephrine
Aerobotics, 330-335; see also flying nanorobots
Affinity (ligand-receptor), 84-85
Affymetrix, 16
AFM (Atomic Force Microscope), 16, 45, 46, 56-60, 102, 112, 144, 154
AFM tip technology, 58-59
Agelas, 3
Aggregation,
aerial nanorobots, 332
platelets, 306
red blood cells, 300-301, 306
Airborne nanorobots, see flying nanorobots
Akay, Christopher W., 78
Alarm signal (chemical), 181
Albedo, mean Earth, Appendix A
Albers, Bruce, 252, 259, 263, 264, 268, 269, 270, Appendix C
ALC, 179
Alchemy, 64
Alcmaeon of Croton, 5
Alcohol (beverage), see ethanol
Alexander the Great, 3, 6
Alimentary system, human, quantification, 222
Alimentography, human, 222-227
Alivisatos, A. Paul, 45
Alkaliphilus, 69
Allbutt, Sir Clifford, 10
Allen, L., 214
Allotropy-transition cooling, 375
Allport, G.W., 25
Alvarez, L.W., 159
Ambilinid, David B., 50
Ambulances, 10
American Academy of Anti-Aging Medicine, 16
American Institute of Ultrasound in Medicine (AIUM), 161
American Physiological Society, 232
Amoeba, 28, 116, 258, 291, 316, 317-318, 324
Amoebidial, 366
Ampere's law, 110
Amphipathic lipids, 257, 258
Amphotericin B, 366
Analog Devices, 41, 98
Analytical Engine, 343
Anaphylaxis, 179, 181
Anastomosis, 213, 214
Anatomy, history of, 10-11
Anchorage, in cell membrane, 313, 315-316
Andrade's formula, 299, 372
Anesthesia,
electronic, 197
history of, 11-12
surgery without, 11
Angular momentum, proton, 113, Appendix A
Anomalous radiative transfer, see near-field anomalous radiative transfer
Ansell, Richard J., 88
Antarctic fish enzyme (cold-tolerant), 374
Antibiotics, 15, 364
Antifungal agents, 366
Antiseptics, history of, 12-14
Antiviral agents, 366
Aorta, human, 210, 212, 243, 245, 302, 319
Apoptosis (programmed cell death), 359-361
Appendix, vermiform (anatomical), 19, 227
Applied Modern Technologies Corp., 191
Archimedes, 6
Arenobufagin, 365
Arey, Leslie Brainerd, 226
Aristotle, 6, 35
Arney, Susanne C., 61
Arnold, Marc, 67
Arnold of Villanova, 12
Aromatic pi bonds, 84
Arteries, 210-211
Artificial proteins and base pairs, 44
Artificial symptoms, 204
Aschoff, Von Jurgen, 239
Asclepiades of Bithynia, 6
Asclepius, 5
Ashkin, A., 369
Asimov, Isaac, 28, 29, 30
Assembler, see molecular assembler
Association rate constant, 85
Astrocytes, 253, 254, 264
Astumian, R. Dean, 110, 149
Asynchronous control, 337
Ataka, Manabu, 293
Atherstone, W.G., 12
Atmosphere (Earth),
composition, Appendix A
standard pressure, Appendix A
Atom and ion diameters, Appendix A
Atom lithography, 42
Atomic clocks, 341
Atomic Force Microscope, see AFM
Atomic mass unit (amu), Appendix A
ATP, 114, 141, 146, 147, 148, 245, 259, 262, 263, 264, 326
AT&T Bell Laboratories, 50
Aubrey, John, Afterword
Audible frequencies, maximum, 162
Audible pressure, minimum, 105-106
Audible power intensity, 117, 199, 207
Australian sensor, see ion channel switch sensor
Australopithecus, 1
Autocrine mediators, 193
Autogenous control, 32, 188, 189-192, 197-206
Autolysis, 262, 361
A.V.I.D. Identification Systems, 353
Avogadro’s number, Appendix A
Avouris, Phaedon, 60

B
Babbage, Charles, 343
Backward chaining, Afterword
Bacteria, 123, 124, 329-330, 370
chemical composition, 371
galvanotropic, 110
in human colon, 227, 247
in therapeutic pills, 69
largest known, 370
magnetoractic, 110, 307
minimum size of motive genera, 308
number density on the skin, 329-330
optimal shapes, 123, 124
smallest known, 370
square, 123
Bacterial actin-based motility, 324
Bacterial exhaust plumes, 246
Bacterial sex pili, 137, 280
Bacteriocide, mechanical, 356, 357, 367-368, 370-372
Bacteriocins, 364
Bacteriophages, 364
Bag inflation pressure, 132-133
Balasubramaniam, L., 62
Ball, Philip, 43
Ballast control, 358-359
Bandgap (electronic), Appendix A
Banting, Sir Frederick, 15
Bar magnet, 110, 192
Barbee, J.H., 303
Barber surgeons, 4, 8
Barbiturates, 15
Barnard, Christiaan, 15
Barnard, Theodor, 12
Barographicyte, 243-244
Barography, 119
Baronatation, 373
Basketane, 48, 49
Basophils, 325
Batteries,
chemical, 140
nuclear, 156
Beardsley, T.M., 31
Bearing, molecular, 61
Becker, R.S., 57
Becker, Wayne M., 248, 260, 264, 286, 371
Becquerel, Alfred, 10
Bedside practice, history of, 23-24
Beilstein Information Systems, 64
Bejan, A., 373
Bekenstein, Jacob D., 353
Bekenstein Bound, 353
Bellevue Hospital (New York), 10
Bending moment, axoneme, 286
Bending stiffness, 287, 315; see also rigidity and stiffness
Bennett, Charles H., 349, 350
Bennett, J.C., 24
Bentham, Jeremy, 11
Benzylpenicillin, 364
Berg, Howard C., 73, 95, 96
Berger, Douglas, Preface
Bergstrom, D.E., 45
Bernal, J.D., 27
Bernard, Claude, 14
Bernard de Gordon, 12
Best, Charles, 15
Betadine, 366
Bhojani, Rehal, Preface
Bialek, W., 93
Billroth, Theodor, 14
Binding pad, 92, 96
Binding site engineering, 82-84, 88-90
Binding sites, active, 86-87
Binnig, Gerd Karl, 56
Biocomputers, 349-351
Bioelectronics, 111, 351
Biofeedback (electric) signals, detection by individual nanorobots, 183
Biofuel cell, 152-154
Biogalvanic energy, 152
Biological power levels, 170, 171
Biological engineering, 27-29
Bioluminescence, 114, 147, 298
Biorobots, 29, 32-35
Biosensors, 93
Biotechnology, 31, 43-46
Biotechnology vs. nanotechnology, 21, 22, 31-35
Bishop, Forrest, Preface
Bishop, W.J., Preface, 3
Bivalvane, 48, 49
Black ass testicles, 4
Black Death, see plague
Blackout, see communications blackout
Blazie Engineering, Inc., 198
Blish, James, 27
Block, Steven M., 93, 110
Blood,
  circulation, 210-211
  composition, Appendix B
Blood flow,
  cell-free plasmatic zone, 302, 305, 312
  inlet length, 302
  parabolic velocity profile, 302-303
  plug and partial-plug, 303, 304
Blood groups, 250-252
Blood perfusion rates, tissues and organs, 214
Blood pressure and pulse, 116
Blood, radial distribution of elements, 301
Blood temperature, 239, 240
Blood transfusions, 14
Blood viscosity, 72, 299-301
Bloodletting or bleeding, see venesection
Blundell, James, 14
Body cell mass (BCM), 247
Boehm, Frank, Preface
Bohlr radius, Appendix A
Bohringer, Karl-Fredrich, 293
Boiling point,
  constants, molal, 374
  function of pressure, 354
  function of solute concentration, 373
  various substances, 374
Boltzmann’s constant, Appendix A
Bond energy and length, Appendix A
Boolean logic/gates, 284-285, 343-344, 349-350, 352
Boron, 143, 152, 166
Bose-Einstein condensates, 341
Bosert, William H., 179
Boswell, Jonathan, Preface
Botulin toxin, LD50, 365
Boundary layer effects, 312
Bova, Ben, 31
Bowling, R. Allen, 275
Bradbury, Robert J., Preface, 21, 29, 35, 179, 180, 196, 297, 298,
  351, 361, 362, 368, 369, 371
Bradwell, Stephen, 8
Braille-like displays, 129, 133, 136, 183, 198
Bramblett, Claud A., Preface
Brandeis University, 64
Brandt, S.A., 337
Bray, Dennis, 193
Breakdown voltage, 109, 141, 291, Appendix A
Brefeldin A, 363
Brenner, Donald W., Preface, 58, 135, 358, Afterword
Bright, Richard, 10
Brin, David, Preface
Broadcast architecture, 65
Bronchial system, human, quantification, 219
Bronchography, human, 219-222
Brookhaven National Laboratory, 71
Broughton, Jeremy Q., 341
Brownian bombardment, effect on nanomachines, 39-40
Brownian computers, 352
Brownian motion/diffusion, 39-40, 72, 124, 180, 235, 245, 302, 305,
  314, 321, 323
Brownian motor, 142, 149, 155
Brownian rotation, 72, 100, 102, 237, 305
Buckling force, 266, 280, 286-287, 329, 331, 358, 370-371
Buckyballs and buckytubes, 51, 53, 58
Bufagin, see arenobufagin
Bulk modulus, Appendix A
Bumpers, metamorphic (nanorobot), 124, 126, 136-137
Buoyancy (nanorobot), 331-332, 358-359
Burke, William, 11
Burney, Fanny, 11
Burton, Alan C., 239
Bush, Vannevar, 343
Butler, J.F., 332
Butterflies, 134, 287, 334
Butterfly molecule, 49

C
CAD/CAM/CASD, 64, 68
Cadherins, 254, 314, 315
Calcium, 79, 193, 194, 204, 260, 367
Calixarenes, 48, 49, 86
Calorimeters, 108
CambridgeSoft, 64
Cambridge University, 193
CAMEO, 64
Cammann, George P., 10
cAMP, see cyclic AMP
Campbell, E.J.M., 19
Cannabis, 4
Canon, 191
Capillaries,
  microcirculation, 211-214
  minimum diameter, 211-213
Capillary forces, 278-280
Carbohydrates, combinatoric diversity, 254
CarboLex, 52
Carbolic acid, 13
Carbon, phases with pressure, 356
Carbon dioxide,
  liquid, 356
  solidification with pressure, 356
Carbon monoxide, 86, 87, 197, 374
Carbon nanotubes, see nanotubes, fullerenes
Carbyne, 140
Carcerands, 48, 86
Carnot efficiency, 142, 152, 170
Carter, Kenneth C., 270
Cartotaxis, 234-235
CAS, 64
CAS Registry, 46
Cassinis (gravity) formula with Bouguer correction, 119
Catalase, 339
Catalytic task space, 44, 85
Catenanes, 50
Cato, 6
Cats, red blood cells, 212
Caveolae, 258
Cavitation, ultrasonic, 109, 161, 273
Cavity Stuffer, 91
Cell, see cytochemical composition, 71, 371
components, quantification, 256
interior hydrodynamic flows, 100, 268, 323
interior map half-life, 268
interior mesh spacing, 267, 271; see also mesh spacing
interior viscosity, 299, 300
mass (typical), 247
maximum number nanorobots inside, 323
mechanical signal transduction, 314, 324-325
schematic view, 255
smallest possible (theoretical), 129, 370
speed of movement, 82, 188, 265, 266, 270, 286, 291, 311, 317,
320, 324, 325
“stab” microinjection, 323
Cell adhesion molecules, 253-254
Cell bioscanning, 111-116
Cell cortex, 265
Cell division, 362, 363
Cell emanations, rf/microwave, 116
Cell junctions, 136-137, 320
Cell membrane,
biomedical composition, 257
elastic moduli, 314-315
lipid bilayer, 255, 257-258, 314, 315
membranolytic limit, 314, 316, 317, 321
penetration, see cytopenetration
self-sealing, 323, 368
snapback time, 315
surface viscosity, 315
Cell mills, 32
Cell poisons, 362-367
Cell repair machines, 30, 32
Cell theory, history of, 14
Cell turnover rate, 90, 200, 235, 258
Cell type identification, 252-255
Cell volume sensing, 114-115
Cells,
detachment force, 136
number in human body, 247
number of types, 252
power density, 171
signal transduction force, 194
types of, Appendix C
Cellular disruption, 368-369
Cellular gate technology, 349-350
Cellular topography, 112
Celsius, 2, 7, 21
Census data, in cyto, 268
Centrifugation, see nanocentrifuge
Cerebrospinal fluid, 229
Cerius2, 64
Chains, Ernst, 15
Chalcogenides, 111, 155
Chamberlain, Fred, Preface
Chamberlain, Linda, Preface
Chancey, C.C., 80
Chang, Thomas M.S., Preface
Changeau, J.P., 26
Channel blockers, 365
Channel destroyers, 365
Chaperones, 263
Charles II of England, 9
Chem3D, 64
Chemical assay, 95
Chemical composition,
bacterial cell, 371
human bile, 225
human bone, 228
human cerebrospinal fluid, 229
human chyme, 245
human duodenal secretion, 226
human extracellular matrix (ECM), 230-231, 319-320
human feces, 227
human flatus, 227
human gastric juice, 224
human blood/serum, Appendix B
human body, 71, 230
human cells, 71, 230-232, 257-266
human lymph, 217-218
human pancreatic juice, 225
human pituitary gland, 230
human pleural fluid, 220
human saliva, 223, 246
human small intestine secretion, 226
human sweat, 273
human synovial fluid, 230
human tooth enamel, 328-329
virus, 369-370
vitamins in human organs, 230, 232
Chemical Computing Group Inc., 64
Chemical energy, human body resources, 147
Chemical force microscopy, 57, 58, 96
Chemical poisoning, 362-367
Chemoelectric cells, 152-153, 154
Chemographicytes, 246
Chemography, 244-246
Chemographicytes, 246
Chemographicytes, 244-246
Chemomapping, 246
Chemomessaging, 191, 193-197
Chemomessenger cable, 185
Chemonavigation, 244-246, 272
Chemosensor, 93-97
Chemotactic sensor, 96, 250, 255
Chemotaxis, 327-328
Chen, Junghuei, 46
Cheselden, William, 9
Chewing, pressure exerted during, 234, 9.5.1
Chien, S., 245, 301, 302, 305
Childbirth, 36
Chinese puzzle box, 294
Chirality, 75, 86
Chiron, 64
Chirurgia (surgery), 6
Chloroform, 12
Chlorophyll, 111, 155
Chlorpromazine (Thorazine), 15, 363
Cholera, 1, 13
Cholesterol, 141, 146, 147, 218, 223, 225, 226, 229, 257,
Appendix B
Chothia, Cyrus, 91
Chromatin (chromosomes), 269, 270, 271, 272
Cytoidentification, 247-255
Cytokines, 193, 194, 195
Cytometrics, 247
Cytonatation, 324-325, 358
Cytopenetration, 321-324, 330
Cytoplasm, 258, 259, 299
Cytoskeleton, 264-266
Cytoplasm, 258-259
Cytotomography, 113-114
Czarn, Andrew, Preface
Daimler-Benz, 41, 98
Damha, Masad J., 45
Danielli, J.F., 28
DARPA, 155, 206
Darsey, Jerry A., Preface
Darwin, Michael G., 29, 31
Data secretion glands, 206
Davy, Sir Humphry, 12
Dead reckoning, 234, 267, 273
deAmor, David W., 248, 260, 264, 265, 268, 286
Death, average age, in Roman times, 2
Death, leading causes in U.S., 1890 and 1990, 9
Debye-Huckel reciprocal length parameter, 84, 278
Debye-Huckel shielding, 83, 109, 113, 373
Decibel notation (acoustics), defined, 117
Dedicated nano-organs,
  biodisposal, 370
  chronometry, 342
  communication, 188, 206
  computation, 353
  energy, 169-170
  library, 188, 238, 353
  navigation, 238
  nuclear power, 157, 170
  sensory, 188
Decensins, 328, 366
Deflection and buckling forces, 286
DeGrado, W.F., 44
De Koninck, Paul, 194
Delaage, Michel, 85
Demarcation, 209, 235
  barographic, 243
  chemical, 244
  thermal, 192, 240
Denaturation, see protein denaturation
Dendrimers, 47, 50, 53, 54, 86, 156
Density,
  various materials, 98, 105, 242, 356, Appendix A
  Dental walking, 328-329
  Denys, Jean-Baptiste, 14
  Dermal displays, 192, 204-206
  Dermatoglyphics, 273
  Dertouzos, Michael, 186
  Descartes, Rene, 24
  Design compilers, 68
  Desmosome, 136, 320
  Detox response (intoxication), 191; see also ethanol
  Dewdney, A.K., 31
  Diagnosis, 21-22, 23-24
Diamond,
  cleavage, 183, 296, Appendix A
  colors, 135, 356
  decomposition, 294-298
  energy storage, 140, 141
  gears, 41
  graphitization, 296
  resistivity, 174, Appendix A
  solvation of, 297
  DiamondCAD, 64
  Diamondoid
    rods and cables, 87, 88, 101, 168, 184, 185
    struts, 51
  Diamondopathy, 298
  Diapedesis (transendothelial migration), 120, 318-319, 327
  Dielectric constant, 84, 109, 110, 141, 151, 155, 166-167, 283,
    Appendix A
    function of temperature, 375
  Dielectric drive, 155
  Dielectric strength, see breakdown voltage
  Diels-Alder reaction, 50, 53, 54
  Difference Engine, 343
  Differential Analyzer, 343
  Diffusion, 72-75
    active intake, 72-74
    elevated rate, due to red cells, 72
    lateral, in plasma membrane, 313
    passive intake, 72, 95-96
    table of coefficients, 73
    translational, 315
  Diffusion cascade sortation, 74-75
  Diffusion coefficients, 73
  Diffusion-limited broadcast, messenger molecules, 179
  Diffusion-limited glucose current, 72, 147
  Diffusion-limited oxygen current, 72, 147
  Diffusion time, glycine, 74
  Dillon, Robert, 307
  Diphtheria, deaths from, 15
  Disassembly for disposal, 293-298
  Disease,
    definition of, 18-20, 36-37
    evolved with man, 1
    of ape ancestors of man, 1
    volitional normative model, 20
  Dislodgement forces, 316, 330
  Dispersion forces, see van der Waals interactions
  Displacement sensor, 97
  Disposability engineering, 294
  Dissociation rate constant, 85
  Ditto, W.L., 351
    computing, 350
    custom sequences, 350
    current (1998) cost, 29, 68
    ordered online, 46, 68
    data storage, 178, 345
    laddering, 360
    mitochondrial, 263
    polyhedra, 45, 46
    superhelical twisting, 168
    UV damage, 369
Dobson, Matthew, 10
Docking,
  of nanorobots, see conjugation
  of nanotubes, 60
Dodecahedron (rhombic, rhombo-hexagonal), 124, 127, 128
DOE, 157
Dogs, red blood cells, 212
Dominance columns, in brain, 192
Dominos, 132
Donaldson, Thomas, Preface, 28-29
Donnan-van’t Hoff formula, 77
Doppelganger proteins, 33
Doppler velocimeter, 98
Dragon Systems, 191
Drain, C.M., 50
Drexler, K. Eric, Foreword, Preface, 26, 27, 30, 31, 35, 36, 39, 43,
  44, 46, 49, 55, 56, 58, 61, 62, 63, 64, 65, 66, 67, 68, 69,
  76, 81, 82, 83, 85, 88, 90, 93, 100, 101, 132, 140, 144,
  145, 155, 166, 148, 153, 165, 174, 175, 178, 184, 192,
  276, 282, 283, 289, 290, 292, 294, 296, 297, 339, 340,
  343, 344, 345, 347, 350, 352, 356, 374, Afterword
Dualist theory of human mind, 24-25
Duan, Yong, 44
Dubois, Eugene, 1
DuCharme, Wesley M., 31
Ducker, W.A., 60
Dworaczyk, Cynthia, Preface
Dye, William L., Preface
Dyer, Mark, Preface, 60
Dynamic Enterprises Ltd., 52
Dynein motor, 116, 147, 263, 286

E

Eachus, Robert I., Preface
Ear, human, 97, 106, 199-200
Earnshaw’s theorem, 151
Earth’s core, 355
Eaton, Philip E., 48
Eavesdropping, 115, 119, 121, 183, 189, 190, 192, 194, 246,
  327-328
Ebers Papyrus, 4
E-CELL, 16
ECG/EEG/EKG, see electrocardiac and electroencephalic
Eckmiller, Rolf, 204
ECM (extracellular matrix), 188, 193, 230-232, 292, 319-320, 360
ECM mesh size, 319
Edelstein, Martin, Preface
Edinburgh Pharmacopoeia of 1803, 23
Edwards, Merrill, 239
Edwin Smith Papyrus, 4-5
Effervescence, 281-282
Eggbeater tool, 292, 372
Eigler, D.M., 57
Einstein, Albert, 57, 58, 72, 315
Einstein-Smoluchowski equation, 315
Einstein-Stokes equation, 72
Elastin, 319
Electrets, 151, 155
Electric field energy, 141-142
Electric field sensors (electrosensors), 109-110, 164-165
Electric fields, comparisons, 109
Electrical insulator, minimum atomic thickness, 42
Electrical interference, 174
Electric vs. mechanical nanosystems, 173-174

Electric Power Research Institute, 160
Electroactive polymers, 154
Electrocardiac (ECG/EKG) and electroencephalic (EEG) signals, detection by individual nanorobots, 183, 192
Electrocution, 109
Electrodynamic suspension, 151
Electromigration, 141, 166
Electron drift speed (mobility), 174, Appendix A
Electronic anesthesia, 197
Electron-volt, Appendix A
Electroporation, 80
Electrosensitive gel sensor, 109
Electrostatic bonds and forces, 83-84, 277
Electrostatic motor/generator, 153-154
Electrostrictive polymers, 154
Electrotaxis, 327
Elementary charge, Appendix A
Elias, Hans, 233
Ellenbogen, J.C., Preface, 348, 349
Emmanuelson, Jerry, 31
Emissivity, Appendix A
Empedocles, 5
Emperipolisis, 267
EMP (electromagnetic pulse), 174, 345
Endocytosis, 258, 261, 262, 323
Endoplasmic reticulum (ER), 260, 263, 266, 267, 268
  rough ER, 197, 250, 260, 269
  smooth ER, 260
Endorphins, 16
Endotheliocytes, 31
Endothelium, 211
Energy,
  nanorobot, see power
  per bit of information, 177, 353
  storage, 139-142
  storage density, various fuels, 141
Engels, Friedrich, 9
Engulf formation, 134, 370
Entropic packing, 308
Enzyme complexes, 259
Enzymes, 86, 108, 349-350
Eosinophils, 325
Epidermal maps, 273
Epinephrine (adrenalin), 194, 196, 246
Equilibrium dissociation constant, 85
Erichsen, Sir John Eric, 14, Afterword
Erasistratus of Chios, 6
Erlangen University, 64
EROS, 64
Erythrocyte, see RBC (red blood cell)
Esophagus, human, 224
Ethane, 87
Ethanol (alcohol), 153, 191, 240, 246, 262, 283, 358, 367, 374
Ether frolics, 12
Etoposide, 363
Ettinger, Robert C.W., 27, 28, 29
Euclid, 6
Euler buckling force, see buckling force
Euler’s theorem, 53
Evans, E.A., 313
Evans, Glen A., 29
Evans, T., 296, 297
Evipan, 15
Examination, of patients, 20-21
Excited states (electronic), 142, 298
Exocytosis, 261, 323
Extremozymes (extremophiles), 298
Extracellular matrix, see ECM
Extrusion, of cell nucleus, see nuclear extrusion
Eyeball, human, 201, 202, 203, 204
Eyring, H., 339

F
Fahraeus, R., 303
Fahy, Gregory M., Preface, 31, 32, 34, 44, 60, 375
Failure (rupture) strength, 33, 265, 266, 295, 313, 314
Fant, K.M., 337
Faraday constant, 81, Appendix A
Fauci, Lisa, 307
Feces, 227
Feinberg, G., 30
Ferric wheel, 48, 49
Ferritin, 262, 351, 366, 368
Feynman Grand Prize, 40, 67
Feynman, Richard P., 26, 30, 39, 40-41, 142, 203, 352, 368, Afterword
Fibroblast, 32, 33, 111, 228, 247, 252, 254, 258, 261, 264, 265, 266, 268, 291, 309, 315, 316, 317, 319, 323, 324, 325, 326, 327, 328, 350, 368
Fiedler, Frederick A., 31
Flight or flight response, 194, 196
Filtration, 77-79
Fingerprints, 273; see also dermatoglyphics
Fingers, metamorphic nanorobotic, 129, 130, 133-134
Finger tapping and typing, 189-191
Fish (swimming), smallest known, 309
Fissile nuclear species, 158
Fission (nuclear power), 158-159
Flagellar ionic motor, 148
Flagellar propulsion, 308, 311
Flashlight beam, 192
Fratu, 118, 227
Fleming, Alexander, 15
Flexible fabric model, 132-133
Flexural rigidity, see rigidity
Flicker frequency, 203
Flight efficiency, insects, 334
“Flight of the Bumblebee” (Rimsky-Korsakov), 190
Flight speed, insects, 332, 334
Flitman, Steven S., Preface
Florey, Howard, 15
Fluconazole, 366
Fluid flow in tubes, 278-281
Fluidic circuits, 284-285
Fluid mixing, 284
Fluorescence, 114, 262, 266
Fluorocarbons and hydrofluorocarbons, 178-179
Fluphenazine, 363
Flying nanorobots, 273-274, 330-335
communication, 207
navigation, 273-274
power and velocity, 332-333
Flywheels (energy storage), 139-140
Focused ultrasound surgery (FUS), 243
Follicular (crinal) map, 273
Food allergy/intolerance, 36
Food viscosity, 223
Foresight Institute, Preface, 40, 67
Forrest, David R., Preface
Foster, John S., 56, 57, 58
Fowler, Berry, Preface
Fowler, Thomas, 343
Frank, E.C., 159
Fraunhofer Institute for Computer Graphics, 64
Fredkin, Edward, 352
Freeman, Tim, Preface
Freeplay Power Group, 145
Freezing point, constants, molal, 374
function of pressure, 355-356
function of solute concentration, 373
various substances, 105, 355-356, 374
Freitas, Barbara Lee, Preface
Freitas, Nancy Ann, Preface
Freitas, Robert A., Jr., Foreword, Preface, 31, 59, 66, Afterword
Freitas, Robert A., Sr., Preface
Freud, Sigmund, 25
Friction and wear, in nanomachines, 40, 49, 51, 55, 61
Frohlich, H., 116
Fronk, K., 245
Fuel cells, 152-153, 154
Fuhr, Gunter, 284
Fujitsu Ltd., 64
Fukuda, Toshio, 309
Fuller, Buckminster, 51
Fullerenes, 16, 51-56, 61, 126, 296, 297, 348, 356
bearings, 55
diameters, 52
gears, 55
motor, 55-56
struts, 56
Fullerene Science and Technology, 51
Fulton, Alice B., 259
Fusion, cell membrane, 261, 322-323
nuclear power
conventional, 159
catalyzed, 159-160
Fusogen, 322

G
Gadolinium-148, 157, 158, 170
Gaethgens, P., 303
Gage, Thomas W., Preface
Gaits, legged nanorobot, 316
Galen of Pergamon, 6, 7, 10, 24
Galileo, 10
Gallabu (barber surgeons), 4
Gallbladder (human organ), 225
Galvanotropic bacteria, 110
Garrett, Steven L., 146
Garrod, Alfred, 10
Gas expansion, temperature effect, 134
Gas leakage, from storage vessels, 356, 358
Gas molecule packing densities, 355
Gasson, Peter C., 128
Gasteiger, Johann, 64
Gastroelectric patterns, 121, 183
Gaussian Inc., 64
Gaussian94, 64
Gaver, Donald, III, 307
Gear trains, 144, 167
General Foods, 356
Gene therapy, 16
Genes, 269, 270
Gen-Etics, 353
Genomes, complete maps, 16
Geographic macrosensing, 119, 120
Geomagnetic field vector, 110, 120
Georgia Tech, 351
Gerber, Ch., 56
Germ theory, history of, 12-14
German Cancer Research Centre, 271
Germanium, 143, 152, 157
Ghadiri, M. Reza, 47, 48, 89
Gibbs free energy, 85, 105-106, 145
Gilbreath, William P., 59, 66
Gilmore, John, Preface
Gimzewski, James K., 40, 49, 89, 343, 346
Girolamo Fracastoro of Verona, 12, 13
Global Positioning System, see GPS
Globus, Al, Preface, 55, 56
Glucose, 141, 147, 149, 196, 245, 261
concentration, 202, 223, 224, 225, 226, 229, 230, Appendix B
engines, 149-152
solubility in water, 150, 282
transporters, 147
Glycocalyx, 210-211, 231, 254, 258, 313, 316
Glycogen, 141, 146, 196, 261
Glycoporphins, 254, 258, 313
Goddard, W.A., III, 62, 63, 64, 65-66
Goldsmith, H.L., 301, 302, 303, 304, 306, 307
Golgi complex, 259, 260-261, 268
Golovchenko, J.A., 57
GPS (Global Positioning System), 183, 188, 237, 238, 342
Gramicidin, 365
Granules (storage, in cells), 261-262
Granulitoxin, 365
Granulocytes, 252
Granulomatous reaction, 169, 179
Granulysin, 328, 362
Graphene, 51-52, 132
Graphite, 51
Graphitization, of diamond, 296
Gravimeter (gravity sensor), 102-103, 119
Gravitational constant, Appendix A
Gravity, acceleration of, Appendix A
Greep, R.O., 221
Greiff, P., 99
Griseofulvin, 363
Guth Laboratories, 153
Guttenstag, Otto E., 2, 20, 22, 24
Guyton, Arthur C., 220, 223, 224, 225, 228, 233, 255
Gyromagnetic ratio, proton, 113, Appendix A
Gyroscopes, nanorobotic, 99, 100
Gyroscopic effects, nanorobots in blood flow, 305-306
Gyrostabilization, of nanorobots, 237, 306

H

Haagensen, C.D., 217, 218
Halacy, D.S., Jr., 28
Haldane, J.B.S., 66
Halg, B., 10.2.1
Hall effect, 111
Halperin, James L., Preface
Hamaker constant, 84, 276, 278
Hammurabic Code, 4, 5
Han, Jie, Preface, 55
Hand, human, bone structure, 189
HandyKey, 190
Hanessian, Steve, 64
Haplotaxis, 319
Hardy, James D., 239
Hare, William, 11
Harpooning, of cells, 368
Harvard Medical School, 204
Harvard University, 64
Harvey, William, 2, Afterword
Haukedalen, Barbara, Preface
Haukedalen, Danny, Preface
Hearing, see audible
Heath, James R., 351
Heat capacity
function of temperature, 374
various materials, 242, Appendix A
Heat of fusion, various materials, 374, Appendix A
Heat of vaporization, various materials, 374, Appendix A
Heat shock proteins, 108, 168
Heat signature, nanorobot, 183
Heedericks, William J., 164
Heidel, Dan, Preface, 43, 44
Heinlein, Robert A., 29-30
HeLa (Henrietta Lacks) cell line, 68, 271, 272, 362
Helical logic, 347-348
Helicopter, 309, 333-334
Helium, solidification with pressure, 356
Hematocrit (Hct), 244, 245, 301, 303
Hemoglobin, Appendix A
Henderson, E.R., 45
Hendrickson, Jim, 64
Henle, Jacob, 14
Henry's law, 281, 282
Heparan sulfate, 109, 231, 253, 319
Hepatocytes (liver cells), 136, 232-233, 253, 261, 264, 269, 271, 319
Hippocractic facies, 6, 21
Hippocratic medicine, 5-6, 7, 8, 22
Hippus movement, 338
Histonatation, 318-321
Histonavigation, 230-234
Hitachi Corp., 346
Hite, B.H., 368
HIV virus, 52-53
HLA, 248-250, 315
HLA specificities, 249
Hocking, Brian, 332
Hoffmann, Roald, 46, 48
Hogg, Tad, Preface
Hoh, Jan H., Preface, 313, 315
Holberton, D.V., 317
Kelvin, Lord (William Thomson), 127
Kerr medium, 111
Kesselman, M.L., 152
Kim, Yun, 57
Kinesin motor, 147, 293, 324, 328
Kinesthesis, 118, 198-199
King, Barry G., 189, 220, 229
King, Penny, Preface
King Asa, 7
King Menelaus, 5
King Ptolemy, 6
Kinnaert, P., 216
Kircher, Athanasius, 13
Kraus, Ronald, 16, 31
Knoke, Tobias A., Preface, 271, 272
Knox, Robert, 11
Ko, W.H., 144
Koch, Robert, 13
Koller, Carl, 12
Kollman, Peter A., 44
Kolostra, Jeffrey D., Preface
Kostruz, A., 315
Kusumi, A., 315
Kwon, Young-Kyun, 348

Laennec, Rene Theophile Hyacinthe, 10
LAHSA, 64
Laing, Richard, 26, 28
Lakkaraju, Sarma, Preface
Lamins, nuclear, see nuclear lamin
Lampert, 241
Lampron, Christopher, 31
Landauer, Rolf, 177, 350, 352
Landes, Ronald G., Preface
Landsteiner, Karl, 14
Langer, Jorg, 271
Laplace equation, Appendix A
LD50, 158, 178, 179, 365, 366
Leave, Geoff U., 51, 56, 63
Leakey, Louis, 1
Lean body mass (LBM), 247
LED, see light emitting devices
Lehn, Jean-Marie, 39, 47, 50, 83
Leiber, Fritz, 27
Lei, Eugene, Preface
Leonardo da Vinci, 35
Leptin, 119, 246
Lesk, Michael, 186
Leukel, Francis, 201
Leukocyte, see white blood cell
Leung, Kevin, Preface
Levine, Philip, 14
Levitation, electrodynamic, 151
Levetin, Lev B., 177, 353
Lewin, Benjamin, 193, 371
Lewis, Eric W., Preface
Lewis, James B., Preface
Libavius, Andreas, 14
Liberman, E.A., 349
Library at Alexandria, 5, 6
Library nodes, 188, 210, 238, 353
Library of Congress, 187
Librium, 15
Lichtman, M.A., 301
Lieber, Charles M., 57, 59
Ligand gating (transport channels), 79
Ligand-receptor dynamics, 86
Ligand-receptor mapping, 90-91
Light-emitting devices (LEDs), 136, 145, 202
Lighthill, J., 308
Lightning, 109, 110
Likharev, Konstantin K., 353
Lincoln Laboratory, 41
Lindqvist, T., 303
Linear accelerator (nuclear), nanoscale, 111, 159
Linear expansion, coefficient of, 108, Appendix A
Lion alcolmeter, 153
Lipid bilayer, see cell membrane lipid bilayer
Lipids, 146, 261
Lipowsky, H.H., 244, 245
Liquid oxygen (LOX), 149
Lister, Joseph, 13
Liver (human organ), 232-234, 246
Load error, 240
Locomotion,
aerial, 330-335
amoeboid, 317-318, 326-327
ECM brachiation, 319-320, 324
ex vivo, 328-335
histological, 318-321
inchworm, 318
in cyto, 324-325
in tissue, 272, 324, 325
intercellular, 320-321
legged, 316-317
nanorobot, 298-335
nanorobot “freeway”, 302, 305, 312
on cell surfaces, see cytoambulation
on dental surface, 328-329
on epidermis, 329-330
on walls and ceilings, 330
saltation (hopping), 330
speed limit, 320
between cells, 320
during cytopenetrant, 327
during sanguinatation, 312
inside cells, 268, 324-325
inside nucleus, 327, 325
through ECM/tissues, 320, 325
swimming, see natation
tank-tread rolling, 317
Loewy, Ariel G., 265
Logajon, James, Preface, 101, 202
Lohn, Jason D., 66
Lomer dislocation, 61, 62
London dispersion forces, see Van der Waals interaction
London Pharmacopoeia of 1809, 23
Long, Alan, 64
Long, Crawford W., 12
Los Alamos National Laboratory, 158, 160
Louis XV of France, 9
Lowell, J., 165
Lower, Richard, 14
Luciano, Dorothy S., 221, 264
Lung (human organ), 220
Lymph, chemical composition, 218
Lymphatic system and lymph, 214-219
Lymphatic system, human, quantification, 215
Lymphocytes, 215, 217, 218, 247, 253, 263, 267, 270, 316, 318,
323, 325, 326, 327
Lysosomes, 262
Lysozymes, 84, 321, 362
Ma, Shugen, 287
MacDonald, Noel C., 60, 61
Macnab, R., 96
MacNicol, D.D., 47
Macrophage, 158, 217, 229, 247, 254, 258, 261, 323, 325, 326,
327, 328, 360, 367
Macrosensing, 116-122, 189
acoustic, 116-118
body weight, 119
electric/magnetic, 119-120, 183
gеographical, 119, 120
kinesthetic, 118-119
neural, 121
optical, 120-121
orientational, 119
temporal, 338
Magnetic cooling, 375
Magnetic field energy, 141
Magnetic field sensors (magnetsensors), 110-111
Magnetic fields, comparisons, 110
Magnetic moment, Appendix A
Magnetic resonance elastography (MRE), 244
Magnetic resonance force microscopy (MRFM), 57, 102, 104
Magnetic Stereotaxis System, 285
Magnetic storms, 120
Magnetocaloric refrigerators, 375
Magnetocencephalograph, 111
Magnetostriiction, 111, 155
Magnetotactic bacteria, 110, 307
Maillard reaction, 149, 366
Major Histocompatibility Complex, see MHC
Malpighi, Marcello, 26
Manipulators,
arrays, 293
control/sensors, 133-134, 292-293
hyper-redundant (snake, tentacle), 289
metamorphic, 290-291, 317-318
metamorphic fingers, 129-130, 133-134
nanorobot, 285-298
pneumatic, 287-289
telescoping, 65, 289-290
tool tips and grippers, 291-292
Margination, 306
Margolus, Norman, 353
Marianas Trench, 355
Marieb, Elaine N., 265
Market Intelligence Research Corp., 41
Mariol, 6
Martin, Charles R., 80
Mason, S.G., 301, 304, 306
Mass, elementary particles, Appendix A
planets and sun, Appendix A
typical tissue cell, 247
Mass, Frank, 181
Massometer, 103-104
Matthes, David, Preface
Matthews, B.W., 350
Maurotoxin, 365
McCammon, J.A., 86
McClure, W.E., 26
McConkey, Edwin H., 259
McKendree, Thomas, Preface, 67
McLellan, William, 40
McNeill, William H., 3
MDL Information Systems Inc., 64
Mechanical power transmission, 167-169
Mechanical information transmission, 184-185
Mechanical rectification (energy), 144
Mechanical signal transduction (cells), 194, 314, 324-325
Mechanical strain, in cells, 313, 314, 328
Mechanosynthesis, 58, 59
Medicine, definition of, 17, 18
Melting point, see freezing point
Membrane-skeleton fence model, 315-316
Memory tape, hydrofluorocarbon, 178, 185, 210
MEMS (MicroElectroMechanical Systems), 41-43, 285, 293, 345
Menstrual cycle, 240, 244, 299, 338
Merck, 12
Merkle, Ralph C., Preface, 31, 59, 60, 61, 62, 63, 65, 66, 67, 68,
132, 140, 153, 174, 175, 185, 290, 297, 347, Afterword
Merrill, Corbett, 31
Mesh sizes (biological), 265, 266, 267, 271, 319, 324
Messenger molecules, 178-179, 191
Messenger pill, 191
Metamorphic bumpers, see bumpers
Metamorphic fingers, 129, 130, 133-134
Metamorphic surfaces, 128-136
MHC (Major Histocompatibility Complex), 135, 248-250
Michaelis-Menten enzyme, 110
Michaelson, S.M., 163
Michl, Josef, 50-51
Microbiorobotics, 29
Microbiotagraphics, 246-247
Microelectromechanical systems, see MEMS
Microfilament (actin), 264-265, 313, 324, 326
Micro Machine Project (Japan), 43
Micromanipulators, 42, 43
Micro Networks, 341
Microtransponders, 165, 183, 235-237
Microtubules, 266, 286, 324
Microvilli, 226
Microwave hearing, 121, 342
Military Health Service Systems (MHSS) 2020, 27
Millard, Nellie D., 189, 220, 229
Mills, Russell, 44
Minsky, Marvin, 25, 30-31, 285, 343
Minton, Allen P., 259
Mirkin, C.A., 45
Misol, Kurt, 48-49
MIT, 41, 190, 204, 343, 349
Mitchell, Duncan, 241
Mitchell, P., 26
Mitchell, Peter Chalmers, 226
Mitochondria, 148, 153, 171, 246, 262-264, 268
Mitosis, 269-270, 271, 362, 363
MITRE Corp., 345, 348
Miura, H., 42
Mo, Y., 345
Mohorovic discontinuity, 355
Molar volume of ideal gas, Appendix A
Molecular abacus, 58
Molecular assembler, 39, 40, 64-69
Molecular Assembly Sequence Software, 64
Molecular bearings, 55, 61
Molecular building blocks, 50
Molecular CAD systems, 64
Molecular conveyor, 63
Molecular Design Institute, 58
Molecular dynamics simulations, 55-56, 61-64
Molecular fine-motion controller, 63-64
Molecular gear systems, 26, 48-49, 55, 61-63
Molecular imprinting receptor model, 88
Molecular LEGO, 50
Molecular manufacturing, 39-69
Molecular mass, Appendix A
Molecular Meccano, 50
Molecular mechanical components, 61-64
Molecular mills, 82
Molecular Modelling Toolkit, 64
Molecular nanoelectronics, 345-349
Molecular nanotechnology, 25-27
Molecular number density, Appendix A
Molecular Operating Environment, 64
Molecular pumps, 63, 80-82
Molecular receptor engineering, 82-92
Molecular recognition, 82-86
Molecular rotation, 72, 77-78
Molecular Simulations Inc., 64
Molecular sorting rotor, see sorting rotor
Molecular struts, 51, 56
Molecular Tinkertoys, 47, 50, 55
Molecular tongs, 49
Molecular volume, Appendix A
Molecular wires, 345-346
Mole quantities, of nanorobots, 175
Monkeywrenching, 368
Montagu, Mary Wortley, 9
Montemagno, Carlo D., 45
Moore, Jeffrey S., 50
Moore's law, 69
MOPAC97, 64
Moravec, Hans, Preface, 69
Morcellation and mincing, 293-294
More, Max, Preface, 31
Morowitz, H.J., 28, 66
Morrison, Philip, 186
Morton, William Thomas Green, 12
Mosaic receptor model, 88-89
Moubarak, Klaus, 88
Mosquitoes, 107, 274, 334
Motomiya, Mineo, 307
Motorola, 98, 349
Mouth, human, 220
MRFM, see magnetic resonance force microscopy
MRI (magnetic resonance imaging) scanning, 110, 113-114, 192
MTOCs, 266, 267
Muller, D.E., 337
Multi-loop subcompartment (MLS) model, 271-272
Munkel, Christian, 271
Muon Catalyzed Fusion, 160
Muons, 159-160
Murphy, Edmond A., 2
Murphy, John B., 14
Murray, Joseph E., 15
Murray, Robert K., 255
Murray's law, 284
Musgrave, Charles B., 58
Myosin (actomyosin) motor, 147, 266, 326

N
Nagoya University, 204
Nakajima, Naomasa, 142
Nakka, Richard, Preface
Nanapheresis, 358-359
Nanobacteria, 37, 321
Nanoballoons, 331-332
Nanobiotics, 32, 209
NanoCAD, 64
Nanocentrifuge, 76-77, 100, 104
Nanocomputer, 342-353
biochemical, 349-350
bioelectronic, 351
biomechanical, 350-351
electronic, 345-349
mechanical, 65, 343-345
memory tape, 65, 185, 345
quantum, 352
reversible, 352
Nanocrit (Nct), 302, 303, 312
Nanocrit, maximum in vivo, 312
Nanoelectromechanical systems (NEMS), 45
Nanoflight, see flying nanorobots
Nanogen, 16
Nanograpmmeter, 102-103
Nanoguitar, 42
Nanogysroscope, 100
Nanoimprint lithography (NIL), 42
Nanomanipulator, The, 58
Nanomedicine, definition of, 2, 18, Glossary
Nanopendulum, 100-102
Nanorobots,
communication, 177-207
computers, 342-353
“freeway” (blood vessel plasmatic zone), 302, 305, 312
geometrical factors, 124, 125
locomotion, 298-335
manipulators, 285-298
maximum size, if bloodborne, 124-125, 211-212
maximum speed in cyto, 268, 324
maximum speed in nucleo, 272, 325
navigation, 209-274
power, 139-176
power scaling, 172-173
second skin, 129
sensors, 93-122
shape, 123-128
waste heat, 242-243
Nanosensors, 93-122
acceleration, 98-100
angular displacement, 100-102
cellular bioscanning, 111-116
chemical, 93-97
electric fields, 109-110
flow rate, 98
force, 102-104
linear displacement, 97
magnetic fields, 110-111
nuclear particles, 111, 157-158
optical, 111
pressure, 105-107
temperature, 107-109
velocity, 97-98
Nanosieve, 77-79
Nanosieve pore clogging, 78
Nanotechnology, 25-27, 31
Nanotubes, 47, 52-53, 55, 61, 97, 279, 280, 296, 348, 355
 diameters, 52
 pressurized gas storage, 354-355, 358
 sword and sheath failure, 60
 Nanowires, 165-166, 345-346
 Napoleon, 10, 11
 NASA, 55, 56, 146, 157, 198, 307
 NASA/Ames Research Center, 55, 56
 Natation, 309-312
 in blood, see sanguinatation
 through cells, see cytonatation
 through tissues, see histonatation
 maximum speed, 312
 National Library of Medicine, 209
 NaturallySpeaking (Dragon Systems), 191
 Naturophilia, 12, 35-37
 Naval Research Laboratory (NRL), 55, 351
 Navicytes, 235-237
 Navigation,
 airborne, 273-274
 barographic, 243-244
 cartotaxis, 234-235
 chemical “area code”, 246, 254, 318
 chemographic, 244-246
 cytographic, 247-273
 dead reckoning, 234, 267
 epidermal, 273
 exodermal, 273-274
 ex vivo, 273-274
 grid (navicyte mobile network), 235-237
 in cyto, see cytonavigation
 in nucleo, see nucleonavigation
 intraorgan, 232-234
 microbiotigraphic, 246-247
 nanorobot, 209-274
 thermographic, 238-243
 vascular bifurcation detection, 237
 Neandertal, 1, 36
 Near-field anomalous radiative transfer, 143, 152, 374
 Near-field scanning optical microscope, see NSOM
 Needles, falling in fluid, 308
 Nemes, T., 30
 NEMS, see nano-electromechanical systems
 Nested annuli, 132

 Networks,
 for nanorobot communication, 186-188
 for nanorobot navigation, 235-237
 for nanorobot power, 169, 170
 Neumister, 60
 Neurographics, 196
 Neuroinformatics, 196
 Neuron, 115-116, 121, 192, 196-197, 198, 200-201, 203-204, 247, 253, 261
 as rf antenna, 342
 discharge cycle, 115, 196
 Neuropeptides, 193, 195, 197, 201, 244
 Neurosensing, 115-116, 121, 196, 197
 Neurotech (Paris), 29
 Neurotransmitters, 193, 195, 196-197
 Neutrophils, 218, 247, 253, 254, 318, 327
 New York University, 45
 NFPA, 179
 Nicotine, 79, 240
 Niemeyer, C.M., 45
 Nippondenso Co. micro-car, 42-43
 NIST, 341, 352
 Nitinol, 143
 Nitric oxide, 87, 195, 197, 363
 Nitrous oxide, as anesthetic, 12
 NLS, see nuclear localization sequence
 NMR, see nuclear magnetic resonance
 Nodes of Ranvier, 115, 342
 No-fly zones, nanorobot, 274, 334-335
 Noid, Donald W., 142
 NOR, see nuclear organizer region
 Nordenstrom, B., 119-120
 NRL, see Naval Research Laboratory
 NSF, 58
 NSOM (near-field scanning optical microscope), 114, 348
 Nuclear battery, 156, 158
 Nuclear catalysis (fusion energy), 159-160
 Nuclear cortex, 269
 Nuclear envelope, 269
 Nuclear envelope, penetration by nanorobots, 324
 Nuclear extrusion, 362, 363-364
 Nuclear fission (energy), 158-159
 Nuclear fusion (energy), 159-160
 Nuclear lamina, 269, 360
 Nuclear lamins, 265-266, 269, 360
 Nuclear localization sequence (NLS), 269
 Nuclear magnetic resonance (NMR) imaging, 113-114
 Nuclear matrix, 271-272
 Nuclear membrane penetration, 324
 Nuclear organizer region (NOR), 269, 270, 271
 Nuclear particle accelerator, 111, 157, 159
 Nuclear pore complex, 269
 Nuclear power/energy, 142, 156-160, 170
 Nuclear power organ, 157, 170
 Nuclear radius (atomic), Rutherford’s formula, 156
 Nucleography, 268-273
 Nucleolus, 270-271
 Nucleon, 270-273
 Nucleonavigation, 272-273
 Nucleoplasm, 269-270
 Nucleosol, 269-270
 Nucleosomes, 270
 Nucleus (cell), 268-273
 cells lacking, 269
 cells with multiple, 269, 271
Null Convention Logic, 337
Nyquist noise (thermal), 142

Oak Ridge National Laboratory (ORNL), 55, 142
Octahedron (nonregular, regular, truncated), 124, 127, 128
Octopus, 129
Ocular Vergence and Accommodation Sensor (OVAS), 191
Odlert, H.S., 25
Oligomycin, 363
Olliver, Vik, Preface
Olson, R.M., 243
Olympiadiane, 50
Olympic games, 5
Open-loop control, 292
Optacon, 198
Optical attenuation in body tissue, 120-121, 202
Optical Coherence Tomography (OCT), 114
Optical power, 121, 155-156, 164, 166-167, 202
Optical fiber, 166-167, 184
Optical refrigerators, 155, 375
Optical tweezers, 164, 284, 291
Optocution, 369
Optofluidic pumping, 284
Orchid Biocomputer, 285
Organ mills, 32
Organography (organs), 230-234
Organs, human, mass and volume, 231
Origami, 130, 285
ORNL, see Oak Ridge National Laboratory
OSHA, 179
Osler, Sir William, 22
Osmotic pressure, 77
Osteography, human, 227-230
Oster, George F., 287
Ostman, Charles, 31
Ouabain, 197
Outmessaging, 197-207

Packing density, 125, 127
Pagodane, 48, 49
Pain, see threshold of pain
Pallet transport, 82
Panacea, 5
Paracrine mediators, 193
Paramecium, 291, 310, 314
Parasol configuration, 130-132
Pare, Ambroise, 8, 12
Paris, John Ayrton, 24
Park, Michael, Preface
Particle accelerator (nuclear), 111, 157, 159
Particles, injected in vivo, 211-212, 214
Partition, of conjugated nanorobots, 321
Pascaline, 343
Paschen effect, 291
Passens, Andreas, Cover Art
PAS Systems, 153
Pasteur, Louis, 13
Pease, Roland, 51
Peltier effect cooling, 155, 375
Penicillin, 15
Penrose, L.S., 66
Peptide nucleic acids (PNAs), 33
Perelson, Alan S., 287
Pergamit, Gayle, 26, 67
Permeability (magnetic) constant, Appendix A
Permittivity (electric) constant, Appendix A
Peroxisome, 16, 262
Personal Area Network (PAN/IBM), 207
Personal status monitors, 206
Peseshet, 4
Peterson, Chris, Preface, 26, 30, 67
Petrie, T.R.W., 159
Pet Trac, 353
Pfizer, 15
Phage exclusion, 361
Phagocytic flagging, 361-362
Phalloidin, 327
Pharmacytes, 32, 367
Pharmacogenes, 32, 367
Phanerozoic flagging, 244
Photoelectric cells, 155-156
Phonomodulation, 284
Phoshostriction, 155
Photosynthesis (chlorophyll), 109, 111, 155
Physical constants - Appendix A
Physician-patient relationship, 24
Physios, 16
Pienza, Kenneth J., 271
Piezoelectric fields and materials, 109, 120, 144, 145, 154, 155
Piezoelectric stress macro-sensing, 120
Piezoluminescence, 145
Piggybacking, 328
Pin cushion receptor model, 90
Pinneo, L.R., 152
Pipe flow, 280-281
Plague, 7-8
Planck's constant, Appendix A
Platelets, 247, 252-253, 269, 301, 306, 307, 310, 312, 326
alloantigen systems, 253
Platinum, for dense nanorobot components, 98
Plicamycin (mithramycin), 362
Pohl, Frederik, Preface
Poiseuille (Hagen-Poiseuille law) flow, 78, 167, 185, 280-281, 285, 302, 303
Poisons, cellular, 362-367
Poison noise, 339
Poison's ratio, 340, 358
Polyribosome, 259
Pope Honorius III, 7
Pope Innocent VIII, 14
Porins, 364-365
Porter, Roy, Preface, 3, 16, 36
Positional uncertainty, due to thermal noise, 39
Postrel, Virginia, Preface, 35, 37
Potassium, 79, 80, 197, 247
Pot, Percival, 9
Power,
acoustic, 145-146, 160-163, 167, 181
analysis in design, 174-175
chemical, 146-153
conversion/transduction, 142-160
electrical vs. mechanical, 173-174
extradermal acoustic sources, 161-162, 205
output, human, 170, 171
pill, 157
pressurized fluids, 140
solar, 121, 155-156, 205
transmission, 160-170
Power Braille, 198
Prandtl, L., 334
Praxagoras of Cos, 6
Pressure sensors, 105-107
Pressure storage and ballasting, 353-359
Pressure vessels and hulls, 353-354, 358
Pressure vessels, leakage and flammability, 356, 358
Pressures, comparative, 355
Pressurized fluids, energy content, 140
Price, Charles, 28
Price Centrum (at T-10), 235
Princeton Simulations, 64
Protein Data Bank, 71
Protein denaturation, 356, 357, 375
Protein fold design, 44
Protein fold families, 91
Protein folding, thermal dependence, 108
Protein surface area (ellipsoidal), 87
Proton exchange membrane (PEM), 153
Proton pump, 148, 155
Pulmonary arterial network, quantification, 213
Pumps (fluid), 282-285
Purcell, Edward M., 73, 95, 96, 305, 309, 311
Purdue University, 347
Pyroelectric effect, 144

Q
Q (resonator or circuit), 104, 164, 340
Quantum computers, 352
Quantum effects, on nanomachines, 39
Quantum game theory, 352
Quartic, H., 190
Quate, Calvin F., 56, 60
Queen Anne (England), 9
Queen Mary (England), 9
Queen Victoria (England), 12
Quorum sensing, 181, 246

R
Radar microchips, 191
Radiation damage, of nanomachines, 40
Radiation dosage (LD50), 158
Radiation shielding, alpha particles, 156, 158
Radiative heat transfer, classical, 143, 152, 157, 175
Radiative heat transfer, near-field anomalous, 143, 152
Radio pills, 189
Radio signals, see TV/radio
Radioactive thermal generators (RTGs), 157
Radiofrequency antenna, neurons used as, 342
Radiofrequency dipole antenna, electric vs. magnetic, 182
Radiofrequency perception by humans, see microwave hearing
Radiofrequency power,
  attenuation, 164, 165
  safe exposure limits, 163-164
  transmission, 156, 162-165
Radiofrequency receivers, 110, 164-165, 182-183
Radiolarians, 52, 53, 126
Radionuclides,
  as power source, 156-159, 170
  power density, 157
Radius,
  particles, Appendix A
  planetary, Appendix A
Ramstrom, Olof, 88
Random-walk giant loop (RWGL) model, 271, 272
Rank, Otto, 25
Raoult’s law, 76
Ras superfamily, 326-327
Rattlesnakes, 107, 137
Rayleigh, Lord, 167
RBC (red blood cell),
  generally, 72, 129, 130, 212-213, 229, 247, 250-252, 254, 258,
  265, 269, 276, 278, 299-301, 303, 305, 306, 307, 308,
  312, 313, 315, 316, 326
osmotic lysis, 314
REAACC, 64
Reaction rate, temperature coefficient, 108
RealMol/CAVE/NAMD, 64
Rebek, Julius, 48, 63
Receptor array, 93-94
Receptor-based transport, 80-82
Receptor design, 86-87, 88-90
Receptor sensor, 96-97
Reciprocity theorem (electromagnetic field theory), 182
Red blood cell, see RBC
Refractive index, 169, Appendix A
Refrigerators, micron-scale, 374-375
Regulation, 373
Rehn, Ludwig, 14
Reifman, Edward M., Preface, 31, 329
Replicators, 59, 65-67
  information content, 65
  medical nanorobots, 35, 67
  unit growth, 67
  unit replication, 66, 67
Requicha, A., 59
Re sistivity, Appendix A; and see conductivity, electrical
Resonance frequency,
  acoustic, 161, 354
  mechanical, 106, 161
NMR, 104, 113-114
Respirocytes, 32, 66, 209, 342, Afterword
Rest mass energy, Appendix A
Reticuloplasmin, 260
Retina, human, 203
Retinal implant devices, 203-204
Reversible computers, 352
Reynolds, Glenn H., 31
Reynolds, L.W., 152
Reynolds number, 117, 217, 280-281, 283, 302, 303-305, 308, 324, 331
rf, see radiofrequency
Rheology, nanorobotic biofluids, 299-303
Rho subfamily, 326-327
Ribosomes, 43-44, 64, 65, 259, 260
Rice University, 16, 58
Ricin toxin, LD50, 366
Riesz, Edward A., Preface, 42
Rigidity, flexural, 266, 311; see also stiffness and bending
Riva-Rocci, Scipione, 10
RNA polymerase, 148
Roberts, Markus, Preface
Roberts, R.S., 19
Roberts, Dennis W., 31
Rohrer, Heinrich, 56
Rose-Innes, A.C, 165
Rotanes, 48
Rotaxanes, 49-50, 53, 54, 347
Roukes, M.L., 165
Royer, H., 368
RTG, see radioactive thermal generator
Ruby, see sapphire and ruby
Rudman, D., 16
Russell, Eric Frank, 30
SAIC, 351
Sako, Y., 315
Sakur-Tetrode equations, 85
Salivarex, 69
Salsbury, Patrick, Preface, 133
Salt bridge, 83
Sandia National Laboratory, 41, 142, 144, 339, 343
Sanguinatation, 124, 303-312, 316
Santoli, Salvatore, Preface
Sapphire and ruby, 135, 170, 173, 295, 296, 297, 298, 331, 332, 340, 356, 358, 372, 374, Appendix A
Sapphire, melting point, 170, 173, 296, 372
softening point, 296, 372
Sapphire-diamond junctions, 135-136, 358
Sapphirophagy, 298
Sarcoplasmic reticulum (SR), 193, 260
Saturation velocity (electronic), Appendix A
Scadding, J.G., 19
Scanning Probe Microscope, see SPM
Scanning probe technology, 56-61
Scanning Tunneling Microscope, see STM
Schaffer, Tilman E., Preface
Schall, J. David, Preface
Schiff, Marcelo, 353
Schleiden, Matthias, 14
Schmidt, Christoph E., 265
Schmidt, R.E., 190
Schneider, Conrad, Preface
Schrodinger, Erwin, 30
Schulman, Howard, 194
Schultz, Peter G., 45
Schwan, Martin A.K., 164
Schwann, Theodor, 14
Schweizer, E.K., 57
Scientific medicine, 3
Scorpion venom, 365
Screw drive, nanorobotic, 310-311, 322
Sculpi, 64
Sedimentation coefficients, 76
Seebeck effect, 155, 375
Seeman, Nadrian C., Preface, 45-46, 49, 350, 351
Selectins, 254, 317, 318, 328
Self-assembly, 45, 47, 66
Self-limiting (self-curing) illness, 22
Self-replicating machine systems, see replicators
Self-replication, see replicators
SEM (scanning electron microscope), 60
Semaphore, presentation, 135, 322, 323, 325
Seminario, Jorge M., 348-349, 352
Semmelweis, Ignaz Philipp, 13
Sensor-based motion control (groping), 292-293
Sensor (chemical) cycle time, 94
Sensors, nanorobot, 93-122
Shape, nanorobot, 123-128
Shapiro, Ehud, 351
Shear modulus, 168, Appendix A
Shearwell Data, 353
Sheehan, Paul E., Preface, 57, 59
Sheffield, Charles, 36
Shen, Zilan, 136
Sherman, James H., 221, 264
Sherwood number, 73
Shi, J., 45
Shinkai, S., 49
Shock, Brian, Preface
Shor, P.W., 352
Shortt clock, 339
Shoulders, K.R., 30
Showers, Maryjane, 189, 220, 229
Shu-He, Wang, 116-117
Shuttle pump, 91-92
Sickle cell, 20
Sidles, John A., Preface, 139
Sieving, see nanosieve
Signaling molecules, human, major classes, 195
Signal/noise ratio (SNR), 97
Sign language, Native American, 190
Silicon, 41, 42, 157
Silicon Designs, 98
Silicon Graphics, 58
Silver, bacteriocidal properties, 366
Silver, Rollo, 285, 343
Silverman, B., 343
Silverman, S.R., 199
Simon, Herbert, 35
Simpson, Sir James Young, 12
Synovial fluid, 230
SYNTREE, 64
Syphilis, 23
Systems Planning Corp., 41
Systron-Donner Analog Computer, 349
Szent-Gyorgyi, A., 119

T4 bacteriophage, 280, 321
Tachometer, 102
Tactile receptors, human skin, 198
Talking, acoustic power, see audible power intensity
Talking, overheard by nanorobots, 118, 274; see also audible power intensity
Talking tattoo, 204
Talking wristwatch, 192, 204
Tanaka, Hirohisa, 288
Taniguchi, Norio, 25, 26
Tatara, Yoichi, 148
Tattoos, 192, 204, 244, 273
Taxol, 327, 363
Taylor, G.R., 28
T cells, 29, 135, 232, 248, 253, 316, 327, 328, 360
number of MHC complexes needed to activate, 135
Telescoping manipulator arm, 289-290
Telesensory, Inc., 198
Television signals, see TV/radio
Telomerase, 246
Telomerases, 362
Temperature-altered biological properties, 375
Temperature-altered physical properties, 372, 373
Temperature effects, on medical nanorobots, 372-375
Temperature of human body, 239; see also thermography
Temperature sensors, 107-109
Tensegrity, in cells, 264, 265, 266
Tensile strength of liquids, 280
Terminal velocity, 331, 334
Terra, Richard P., Preface
Tessellating surfaces, 125-127
Tethers,
communication, 184-185
power transmission, 165-170
utility and biocompatibility, 169, 184, 188
Tetraethylammonium, 197
Tetrodotoxin, 197, 365
Texas Instruments, 206
Thackrah, Charles Turner, 9
Thermal conductivity, Appendix A; see also conductivity, thermal
Thermal emissivity, Appendix A
Thermal equilibration time, 107
Thermal expansion coefficients, 108, 143, Appendix A
Thermoacoustic refrigeration, 146, 375
Thermogenic limits,
individual nanorobots, 170-173
nanorobot aggregates, 172-173
Thermographicytes, 240, 241, 242
Thermography, 238-243
Thermology, skin, 239
Thermometer, clinical, 10
Thioridazine, 363
Thouless electron pumping, 347
Threshold of pain,
acoustic, 161, 162, 172
radiant heat, 164, 172
Tillyard, R.J., 332
Tinkertoy computer, 343
Tissue mills, 32
Tissues, 230-232
Toffoli, Tommaso, 352
Tomanek, David, 57
Tomographic receptor model, 89-90
Tool tips, 58-59, 291-292
Tooth surfaces, 328-329
Topometrix, 58
Tortora, Gerard J., 213, 215, 216, 217
Torh-Fejel, Tihamar, Preface
Tour, James M., Preface, 345, 346, 347, 348-349, 352
Towing force, 316, 317, 325, 326
Toxins and venoms, 365, 366, 370
Toyota Motor Corp., 42
Trabesinger, Werner, Preface
Transcript domains (cell nucleus), 271-272
Transendothelial migration, see diapedesis
Transit times, 211
Transporter pumps, 80-81
Transvenue outmessaging, 207
Trapezohedron, 128
Travis, John, 324
Trephination, 3
Triangulane, 48
Trinity Software Inc., 64
Tripós Inc., 64
Trombone, 132
Troyk, Philip R., 164
Truncated octahedron, see octahedron
Ullage sensors, 106-107
Ullmann, Emerich, 15
Ultrasonic hearing, 162, 200
Ultraviolet, 111, 136, 161, 167, 184, 359, 360, 369
Umami, 200
Units of measurement, Appendix A
TV/radio reception,
by dedicated communication organ, 120, 188
by individual nanorobots, 120, 182-183
Tvibell, Ty S., Preface
Twiddler, 190
Tyvand, P.A., 373

Ubiquitin, 262, 364
Ubility sensors, 106-107
Uilleann, Emerich, 15
Ultrasonic hearing, 162, 200
Ultraviolet, 111, 136, 161, 167, 184, 359, 360, 369
Umami, 200
Units of measurement, Appendix A
Universal assembler, 64
Universal gas constant, Appendix A
Universitat Bonn, 204
University of California, Berkeley (UCB), 41
University of California, Los Angeles (UCLA), 16
Vacuum-isolated heat engine, 143
Vacuum pumping, 358
Vacuum pumping, wall material sublimation, 358
Valdes, Francisco, Preface
Valinomycin, 365
Valium, 15
Vancomycin, 364
Van de Graaff generator, 154, 166
Vander, Arthur J., 203, 221, 264
Van der Waals equation,
gas constants, 355
gas pressure, 134, 354-355
Van der Waals interaction (forces), 84, 87, 90, 101, 275-277, 278, 291
Van Hove, Michel A., 373
Van Leeuwenhoek, Anthoni, 13
Van Nedervelde, Philippe, Preface
Van Vugt, A.E., 27
Vasculography, human, 210-219
Vaults, 262
Veins, 210-211
Vein valve architecture, 211
Velocity of sound, see speed of sound
Vena cava, human, 210, 211, 212
Venerable Bede, 7
Veinesection, 5, 7, 8
Veratridine, 196
Vesalius, Andreas, 10-11
Vesicles, 261-262
Vestibular mechanism, frequency response, 102
Vetter, Steven C., Preface, 57
Vierordt, Karl, 10
Vinblastine, 327, 363
Vincristine, 327, 362, 363, 368
Viny sulphone, 363
Vion Pharmaceuticals, 29
Vircow, Rudolph, 14
Virtual Cell, 16
Virtual Reality Dataglove, 58
Virucide,
biochemical, 366
mechanical, 356, 357, 369-370
Viruses, 369
Viruses, enveloped, 254, 322, 369
Viscosity, 299, 300, Appendix A
Viscous anchoring, 311-312
Visible Human Project, 209-210
Vitamin parts (in replicator theory), 59

Vitamins,
concentrations in human organs, 232
nutritional, 15, 227, 230, 232
Vogel, Steven, 34, 35, 325
Vogel, Viola, 45, 351
Vogel Scientific Software Inc., 64
VoiceCommands/VoicePro (Kurzweil), 191
Voice Type Dictation/Personal Dictation (IPDS), 191
Volitional normative model of disease, 20
Volkenstein, Mikhail V., 30
Voltage gating (transport channels), 79-80, 194
Vomeronasal organ, 191, 219
Von Ehr, James, Preface, 65, 67
Von Foester, H., 25
Von Hippel, A.R., 25
Von Kolliker, Albert, 14
Von Neumann, John, 30, 65, 66
Von Pettenkofer, Max, 13

W
Waksman, Selman A., 15
Wamsley, Michelle, Preface
Ware, William, Preface, 182
Warwick, Kevin, 353
Washington University, 16
Weiss, L., 221
Wells, Horace, 12
Wenger, C. Bruce, 239
Wenz, Gerhard, 50
Westinghouse Science and Technology Center, 41
Wetting, of surfaces, 278, 279
Wever, Rutger, 239
Weyl, H., 127
White blood cells (leukocytes), 120, 194, 229, 247, 253, 254, 301, 308, 312, 317, 318, 325, 326, 327, 328
White, J., 28, 29
Whitesides, George M., 47
Whitmore, R.L., 303
Wien's relation, 152
Will, Peter, 293
Williams, Robert, 127, 128
Wilson, Edward O., 179
Wilson, James A., 200
Wilson, J.F., 288
Winslow effect, 284
Wiseman, Richard, 8
Wolf packs, 2
Wong, S.S., 58
Wood, James Brent, Preface
Working stress, 99, 168
Worldwide data storage, 186
Worth, Christian P., Preface
Wound sealing, by insects, 1-2
Wowk, Brian, 26-27, 31
Wren, Sir Christopher, 14
Wunderlich, Carl, 10
WWV/WWVB, 342; see also Coordinated Universal Time

X
X-ray holography, 114

Y
Yale University, 29, 64
Yamamoto, G., 26, 48
Yao, J. Jason, 61
Year, tropical, Appendix A
Yoctocalorimetry, 107
Yodoxin, 366
Young's modulus,
  generally, 139, 234, 286, 294, 358, 372
  various materials, 52, 295, 329, 340, 341, 350, Appendix A
Young, Thomas, 9

Z
Zarate, Lucia, 209
Zeolites, 47, 89, 128
Zimmermann, M., 280
Zingsheim, H.P., 25
Zippocytes (dermal zippers), 32
Zweifach, B.W., 244, 245
Zwitterions, 84, 111
Zymogen, 224, 261
Zyvex LLC, 58, 60, 64, 65, 67