

NEUROSCIENCE

Third Edition

NEUROSCIENCE

THIRD EDITION

Edited by

DALE PURVES

GEORGE J. AUGUSTINE

DAVID FITZPATRICK

WILLIAM C. HALL

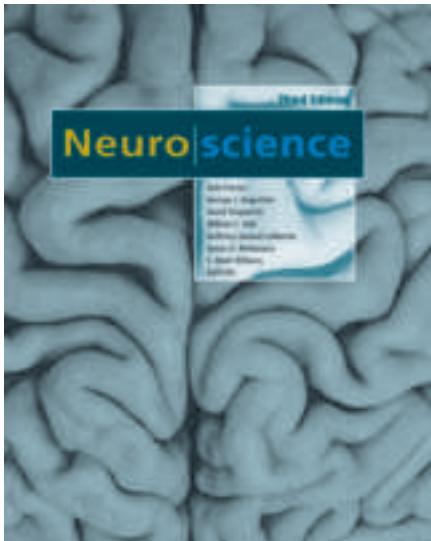
ANTHONY-SAMUEL LAMANTIA

JAMES O. McNAMARA

S. MARK WILLIAMS



Sinauer Associates, Inc. • Publishers
Sunderland, Massachusetts U.S.A.



THE COVER

Dorsal view of the human brain.
(Courtesy of S. Mark Williams.)

NEUROSCIENCE: Third Edition

Copyright © 2004 by Sinauer Associates, Inc. All rights reserved.
This book may not be reproduced in whole or in part without permission.

Address inquiries and orders to
Sinauer Associates, Inc.
23 Plumtree Road
Sunderland, MA 01375 U.S.A.

www.sinauer.com
FAX: 413-549-1118
orders@sinauer.com
publish@sinauer.com

Library of Congress Cataloging-in-Publication Data

Neuroscience / edited by Dale Purves ... [et al.].— 3rd ed.

p. ; cm.

Includes bibliographical references and index.

ISBN 0-87893-725-0 (casebound : alk. paper)

1. Neurosciences.

[DNLM: 1. Nervous System Physiology. 2. Neurochemistry.

WL 102 N50588 2004] I. Purves, Dale.

QP355.2.N487 2004

612.8—dc22

2004003973

Printed in U.S.A.

5 4 3 2 1

Contributors

George J. Augustine, Ph.D.
Dona M. Chikaraishi, Ph.D.
Michael D. Ehlers, M.D., Ph.D.
Gillian Einstein, Ph.D.
David Fitzpatrick, Ph.D.
William C. Hall, Ph.D.
Erich Jarvis, Ph.D.
Lawrence C. Katz, Ph.D.
Julie Kauer, Ph.D.
Anthony-Samuel LaMantia, Ph.D.
James O. McNamara, M.D.
Richard D. Mooney, Ph.D.
Miguel A. L. Nicolelis, M.D., Ph.D.
Dale Purves, M.D.
Peter H. Reinhart, Ph.D.
Sidney A. Simon, Ph.D.
J. H. Pate Skene, Ph.D.
James Voyvodic, Ph.D.
Leonard E. White, Ph.D.
S. Mark Williams, Ph.D.

UNIT EDITORS

UNIT I: George J. Augustine
UNIT II: David Fitzpatrick
UNIT III: William C. Hall
UNIT IV: Anthony-Samuel LaMantia
UNIT V: Dale Purves

Contents in Brief

1. Studying the Nervous Systems of Humans and Other Animals 1

UNIT I NEURAL SIGNALING

2. Electrical Signals of Nerve Cells 31
3. Voltage-Dependent Membrane Permeability 47
4. Channels and Transporters 69
5. Synaptic Transmission 93
6. Neurotransmitters, Receptors, and Their Effects 129
7. Molecular Signaling within Neurons 165

UNIT II SENSATION AND SENSORY PROCESSING

8. The Somatic Sensory System 189
9. Pain 209
10. Vision: The Eye 229
11. Central Visual Pathways 259
12. The Auditory System 283
13. The Vestibular System 315
14. The Chemical Senses 337

UNIT III MOVEMENT AND ITS CENTRAL CONTROL

15. Lower Motor Neuron Circuits and Motor Control 371
16. Upper Motor Neuron Control of the Brainstem and Spinal Cord 393
17. Modulation of Movement by the Basal Ganglia 417
18. Modulation of Movement by the Cerebellum 435
19. Eye Movements and Sensory Motor Integration 453
20. The Visceral Motor System 469

UNIT IV THE CHANGING BRAIN

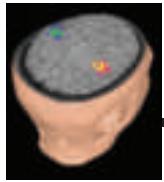
21. Early Brain Development 501
22. Construction of Neural Circuits 521
23. Modification of Brain Circuits as a Result of Experience 557
24. Plasticity of Mature Synapses and Circuits 575

UNIT V COMPLEX BRAIN FUNCTIONS

25. The Association Cortices 613
26. Language and Speech 637
27. Sleep and Wakefulness 659
28. Emotions 687
29. Sex, Sexuality, and the Brain 711
30. Memory 733

APPENDIX A THE BRAINSTEM AND CRANIAL NERVES 755

APPENDIX B VASCULAR SUPPLY, THE MENINGES, AND THE VENTRICULAR SYSTEM 763



Contents

Preface *xvi*

Acknowledgments *xvii*

Supplements to Accompany NEUROSCIENCE *xviii*

Chapter 1 Studying the Nervous Systems of Humans and Other Animals 1

Overview 1

Genetics, Genomics, and the Brain 1

The Cellular Components of the Nervous System 2

Neurons 4

Neuroglial Cells 8

Cellular Diversity in the Nervous System 9

Neural Circuits 11

Overall Organization of the Human Nervous System 14

Neuroanatomical Terminology 16

The Subdivisions of the Central Nervous System 18

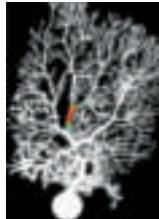
Organizational Principles of Neural Systems 20

Functional Analysis of Neural Systems 23

Analyzing Complex Behavior 24

Box A Brain Imaging Techniques 25

Summary 26



Unit I NEURAL SIGNALING

Chapter 2 Electrical Signals of Nerve Cells 31

Overview 31

Electrical Potentials across Nerve Cell Membranes 31

How Ionic Movements Produce Electrical Signals 34

The Forces That Create Membrane Potentials 36

Electrochemical Equilibrium in an Environment with More Than One Permeant Ion 38

The Ionic Basis of the Resting Membrane Potential 40

Box A The Remarkable Giant Nerve Cells of Squid 41

The Ionic Basis of Action Potentials 43

Box B Action Potential Form and Nomenclature 44

Summary 45

Chapter 3 Voltage-Dependent Membrane Permeability 47

Overview 47

Ionic Currents Across Nerve Cell Membranes 47

Box A The Voltage Clamp Method 48

Two Types of Voltage-Dependent Ionic Current 49

Two Voltage-Dependent Membrane Conductances 52

Reconstruction of the Action Potential 54

Long-Distance Signaling by Means of Action Potentials 56

Box B Threshold 57

Box C Passive Membrane Properties 60

The Refractory Period 61

Increased Conduction Velocity as a Result of Myelination 63

Summary 65

Box D Multiple Sclerosis 66

Chapter 4 Channels and Transporters 69

Overview	69	
Ion Channels Underlying Action Potentials		69
Box A The Patch Clamp Method		70
The Diversity of Ion Channels		73
Box B Expression of Ion Channels in <i>Xenopus</i> Oocytes		75
Voltage-Gated Ion Channels		76
Ligand-Gated Ion Channels		78
Stretch- and Heat-Activated Channels		78
The Molecular Structure of Ion Channels		79
Box C Toxins That Poison Ion Channels		82
Box D Diseases Caused by Altered Ion Channels		84
Active Transporters Create and Maintain Ion Gradients		86
Functional Properties of the Na^+/K^+ Pump		87
The Molecular Structure of the Na^+/K^+ Pump		89
Summary		90

Chapter 5 Synaptic Transmission 93

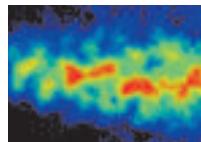
Overview	93	
Electrical Synapses		93
Signal Transmission at Chemical Synapses		96
Properties of Neurotransmitters		96
Box A Criteria That Define a Neurotransmitter		99
Quantal Release of Neurotransmitters		102
Release of Transmitters from Synaptic Vesicles		103
Local Recycling of Synaptic Vesicles		105
The Role of Calcium in Transmitter Secretion		107
Box B Diseases That Affect the Presynaptic Terminal		108
Molecular Mechanisms of Transmitter Secretion		110
Neurotransmitter Receptors		113
Box C Toxins That Affect Transmitter Release		115
Postsynaptic Membrane Permeability Changes during Synaptic Transmission		116
Excitatory and Inhibitory Postsynaptic Potentials		121
Summation of Synaptic Potentials		123
Two Families of Postsynaptic Receptors		124
Summary		126

Chapter 6 Neurotransmitters and Their Receptors 129

Overview	129	
Categories of Neurotransmitters		129
Acetylcholine		129
Box A Addiction		134
Box B Neurotoxins that Act on Postsynaptic Receptors		136
Glutamate		137
Box C Myasthenia Gravis: An Autoimmune Disease of Neuromuscular Synapses		140
GABA and Glycine		143
Box D Excitotoxicity Following Acute Brain Injury		145
The Biogenic Amines		147
Box E Biogenic Amine Neurotransmitters and Psychiatric Disorders		148
ATP and Other Purines		152
Peptide Neurotransmitters		153
Unconventional Neurotransmitters		157
Box F Marijuana and the Brain		160
Summary		161

Chapter 7 Molecular Signaling within Neurons 165

Overview	165	
Strategies of Molecular Signaling		165
The Activation of Signaling Pathways		167
Receptor Types		168
G-Proteins and Their Molecular Targets		170
Second Messengers		172
Second Messenger Targets: Protein Kinases and Phosphatases		175
Nuclear Signaling		178
Examples of Neuronal Signal Transduction		181
Summary		184



Unit II SENSATION AND SENSORY PROCESSING

Chapter 8 The Somatic Sensory System 189

- Overview 189
 Cutaneous and Subcutaneous Somatic Sensory Receptors 189
 Mechanoreceptors Specialized to Receive Tactile Information 192
 Differences in Mechanosensory Discrimination across the Body Surface 193
Box A Receptive Fields and Sensory Maps in the Cricket 195
Box B Dynamic Aspects of Somatic Sensory Receptive Fields 196
 Mechanoreceptors Specialized for Proprioception 197
 Active Tactile Exploration 199
 The Major Afferent Pathway for Mechanosensory Information: The Dorsal Column–Medial Lemniscus System 199
 The Trigeminal Portion of the Mechanosensory System 202
Box C Dermatomes 202
 The Somatic Sensory Components of the Thalamus 203
 The Somatic Sensory Cortex 203
 Higher-Order Cortical Representations 206
Box D Patterns of Organization within the Sensory Cortices: Brain Modules 207
 Summary 208

Chapter 9 Pain 209

- Overview 209
 Nociceptors 209
 Transduction of Nociceptive Signals 211
Box A Capsaicin 212
 Central Pain Pathways 213
Box B Referred Pain 215
Box C A Dorsal Column Pathway for Visceral Pain 218
 Sensitization 220
Box D Phantom Limbs and Phantom Pain 222
 Descending Control of Pain Perception 224
 The Placebo Effect 224
 The Physiological Basis of Pain Modulation 225
 Summary 227

Chapter 10 Vision: The Eye 229

- Overview 229
 Anatomy of the Eye 229
 The Formation of Images on the Retina 231
Box A Myopia and Other Refractive Errors 232
 The Retina 234
 Phototransduction 236
Box B Retinitis Pigmentosa 239
 Functional Specialization of the Rod and Cone Systems 240
Box C Macular Degeneration 243
 Anatomical Distribution of Rods and Cones 244
 Cones and Color Vision 245
Box D The Importance of Context in Color Perception 247
 Retinal Circuits for Detecting Luminance Change 249
Box E The Perception of Light Intensity 250
 Contribution of Retinal Circuits to Light Adaptation 254
 Summary 257

Chapter 11 Central Visual Pathways 259

- Overview 259
 Central Projections of Retinal Ganglion Cells 259
Box A The Blind Spot 262
 The Retinotopic Representation of the Visual Field 263
 Visual Field Deficits 267
 The Functional Organization of the Striate Cortex 269
 The Columnar Organization of the Striate Cortex 271
Box B Random Dot Stereograms and Related Amusements 272
 Division of Labor within the Primary Visual Pathway 275
Box C Optical Imaging of Functional Domains in the Visual Cortex 276
 The Functional Organization of Extrastriate Visual Areas 278
 Summary 281

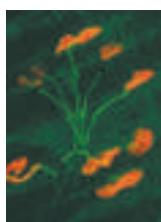
Chapter 12 The Auditory System 283

- Overview 283
 Sound 283
 The Audible Spectrum 284

A Synopsis of Auditory Function	285
Box A Four Causes of Acquired Hearing Loss	285
Box B Music	286
The External Ear	287
The Middle Ear	289
The Inner Ear	289
Box C Sensorineural Hearing Loss and Cochlear Implants	290
Box D The Sweet Sound of Distortion	295
Hair Cells and the Mechanoelectrical Transduction of Sound Waves	294
Two Kinds of Hair Cells in the Cochlea	300
Tuning and Timing in the Auditory Nerve	301
How Information from the Cochlea Reaches Targets in the Brainstem	303
Integrating Information from the Two Ears	303
Monaural Pathways from the Cochlear Nucleus to the Lateral Lemniscus	307
Integration in the Inferior Colliculus	307
The Auditory Thalamus	308
The Auditory Cortex	309
Box E Representing Complex Sounds in the Brains of Bats and Humans	310
Summary	313

Chapter 13 The Vestibular System 315

Overview	315
The Vestibular Labyrinth	315
Vestibular Hair Cells	316
The Otolith Organs: The Utricle and Saccule	317
Box A A Primer on Vestibular Navigation	318
Box B Adaptation and Tuning of Vestibular Hair Cells	320
How Otolith Neurons Sense Linear Forces	322
The Semicircular Canals	324



Unit III MOVEMENT AND ITS CENTRAL CONTROL

Chapter 15 Lower Motor Neuron Circuits and Motor Control 371

Overview	371
Neural Centers Responsible for Movement	371

How Semicircular Canal Neurons Sense Angular Accelerations	325
Box C Throwing Cold Water on the Vestibular System	326
Central Pathways for Stabilizing Gaze, Head, and Posture	328
Vestibular Pathways to the Thalamus and Cortex	331
Box D Mauthner Cells in Fish	332
Summary	333

Chapter 14 The Chemical Senses 337

Overview	337
The Organization of the Olfactory System	337
Olfactory Perception in Humans	339
Physiological and Behavioral Responses to Odorants	341
The Olfactory Epithelium and Olfactory Receptor Neurons	342

Box A Olfaction, Pheromones, and Behavior in the Hawk Moth 344

The Transduction of Olfactory Signals	345
Odorant Receptors	346
Olfactory Coding	348
The Olfactory Bulb	350

Box B Temporal "Coding" of Olfactory Information in Insects 350

Central Projections of the Olfactory Bulb	353
The Organization of the Taste System	354
Taste Perception in Humans	356
Idiosyncratic Responses to Tastants	357
The Organization of the Peripheral Taste System	359
Taste Receptors and the Transduction of Taste Signals	360
Neural Coding in the Taste System	362
Trigeminal Chemoreception	363
Summary	366

Motor Neuron–Muscle Relationships	373
The Motor Unit	375
The Regulation of Muscle Force	377
The Spinal Cord Circuitry Underlying Muscle Stretch Reflexes	379

The Influence of Sensory Activity on Motor Behavior	381
Other Sensory Feedback That Affects Motor Performance	382
Box A Locomotion in the Leech and the Lamprey	
384	
Flexion Reflex Pathways	387
Spinal Cord Circuitry and Locomotion	387
Box B The Autonomy of Central Pattern Generators: Evidence from the Lobster Stomatogastric Ganglion	
388	
The Lower Motor Neuron Syndrome	389
Box C Amyotrophic Lateral Sclerosis	
391	
Summary	391

Chapter 16 Upper Motor Neuron Control of the Brainstem and Spinal Cord 393

Overview	393
Descending Control of Spinal Cord Circuitry: General Information	393
Motor Control Centers in the Brainstem: Upper Motor Neurons That Maintain Balance and Posture	397
Box A The Reticular Formation	
398	
The Corticospinal and Corticobulbar Pathways: Upper Motor Neurons That Initiate Complex Voluntary Movements	402
Box B Descending Projections to Cranial Nerve Motor Nuclei and Their Importance in Diagnosing the Cause of Motor Deficits	
404	
Functional Organization of the Primary Motor Cortex	405
Box C What Do Motor Maps Represent?	
408	
The Premotor Cortex	411
Box D Sensory Motor Talents and Cortical Space	
410	
Damage to Descending Motor Pathways: The Upper Motor Neuron Syndrome	412
Box E Muscle Tone	
414	
Summary	415

Chapter 17 Modulation of Movement by the Basal Ganglia 417

Overview	417
Projections to the Basal Ganglia	417
Projections from the Basal Ganglia to Other Brain Regions	422
Evidence from Studies of Eye Movements	423

Circuits within the Basal Ganglia System	424
Box A Huntington's Disease	
426	
Box B Parkinson's Disease: An Opportunity for Novel Therapeutic Approaches	
429	
Box C Basal Ganglia Loops and Non-Motor Brain Functions	
432	
Summary	433

Chapter 18 Modulation of Movement by the Cerebellum 435

Overview	435
Organization of the Cerebellum	435
Projections to the Cerebellum	438
Projections from the Cerebellum	440
Circuits within the Cerebellum	441
Box A Prion Diseases	
444	
Cerebellar Circuitry and the Coordination of Ongoing Movement	445
Futher Consequences of Cerebellar Lesions	448
Summary	449
Box B Genetic Analysis of Cerebellar Function	
450	

Chapter 19 Eye Movements and Sensory Motor Integration 453

Overview	453
What Eye Movements Accomplish	453
The Actions and Innervation of Extraocular Muscles	454
Box A The Perception of Stabilized Retinal Images	
456	
Types of Eye Movements and Their Functions	457
Neural Control of Saccadic Eye Movements	458
Box B Sensory Motor Integration in the Superior Colliculus	
462	
Neural Control of Smooth Pursuit Movements	466
Neural Control of Vergence Movements	466
Summary	467

Chapter 20 The Visceral Motor System 469

Overview	469
Early Studies of the Visceral Motor System	469
Distinctive Features of the Visceral Motor System	470
The Sympathetic Division of the Visceral Motor System	471
The Parasympathetic Division of the Visceral Motor System	476
The Enteric Nervous System	479
Sensory Components of the Visceral Motor System	480

Central Control of Visceral Motor Functions	483
Box A The Hypothalamus	484
Neurotransmission in the Visceral Motor System	487
Box B Horner's Syndrome	488
Box C Obesity and the Brain	490

Visceral Motor Reflex Functions	491
Autonomic Regulation of Cardiovascular Function	491
Autonomic Regulation of the Bladder	493
Autonomic Regulation of Sexual Function	496
Summary	498



Unit IV THE CHANGING BRAIN

Chapter 21 Early Brain Development **501**

Overview	501
The Initial Formation of the Nervous System: Gastrulation and Neurulation	501
The Molecular Basis of Neural Induction	503
Box A Stem Cells: Promise and Perils	504
Box B Retinoic Acid: Teratogen and Inductive Signal	506
Formation of the Major Brain Subdivisions	510
Box C Homeotic Genes and Human Brain Development	513
Box D Rhombomeres	514
Genetic Abnormalities and Altered Human Brain Development	515
The Initial Differentiation of Neurons and Glia	516
Box E Neurogenesis and Neuronal Birthdating	517
The Generation of Neuronal Diversity	518
Neuronal Migration	520
Box F Mixing It Up: Long-Distance Neuronal Migration	524
Summary	525

Chapter 22 Construction of Neural Circuits **527**

Overview	527
The Axonal Growth Cone	527
Non-Diffusible Signals for Axon Guidance	528
Box A Choosing Sides: Axon Guidance at the Optic Chiasm	530
Diffusible Signals for Axon Guidance: Chemoattraction and Repulsion	534
The Formation of Topographic Maps	537
Selective Synapse Formation	539

Box B Molecular Signals That Promote Synapse Formation **542**

Trophic Interactions and the Ultimate Size of Neuronal Populations	543
Further Competitive Interactions in the Formation of Neuronal Connections	545
Molecular Basis of Trophic Interactions	547
Box C Why Do Neurons Have Dendrites?	548
Box D The Discovery of BDNF and the Neurotrophin Family	552
Neurotrophin Signaling	553
Summary	554

Chapter 23 Modification of Brain Circuits as a Result of Experience **557**

Overview	557
Critical Periods	557
Box A Built-In Behaviors	558
The Development of Language: Example of a Human Critical Period	559
Box B Birdsong	560
Critical Periods in Visual System Development	562
Effects of Visual Deprivation on Ocular Dominance	563
Box C Transneuronal Labeling with Radioactive Amino Acids	564
Visual Deprivation and Amblyopia in Humans	568
Mechanisms by which Neuronal Activity Affects the Development of Neural Circuits	569
Cellular and Molecular Correlates of Activity- Dependent Plasticity during Critical Periods	572
Evidence for Critical Periods in Other Sensory Systems	572
Summary	573

Chapter 24 Plasticity of Mature Synapses and Circuits 575

Overview 575

Synaptic Plasticity Underlies Behavioral Modification in Invertebrates 575

Box A Genetics of Learning and Memory in the Fruit Fly 581

Short-Term Synaptic Plasticity in the Mammalian Nervous System 582

Long-Term Synaptic Plasticity in the Mammalian Nervous System 583

Long-Term Potentiation of Hippocampal Synapses 584

Molecular Mechanisms Underlying LTP 587

Box B Dendritic Spines 590

Long-Term Synaptic Depression 592

Box C Silent Synapses 594

Changes in Gene Expression Cause Enduring Changes in Synaptic Function during LTP and LTD 597

Plasticity in the Adult Cerebral Cortex 599

Box D Epilepsy: The Effect of Pathological Activity on Neural Circuitry 600

Recovery from Neural Injury 602

Generation of Neurons in the Adult Brain 605

Box E Why Aren't We More Like Fish and Frogs? 606

Summary 609



Unit V COMPLEX BRAIN FUNCTIONS

Chapter 25 The Association Cortices 613

Overview 613

The Association Cortices 613

An Overview of Cortical Structure 614

Specific Features of the Association Cortices 615

Box A A More Detailed Look at Cortical Lamination 617

Lesions of the Parietal Association Cortex: Deficits of Attention 619

Lesions of the Temporal Association Cortex: Deficits of Recognition 622

Lesions of the Frontal Association Cortex: Deficits of Planning 623

Box B Psychosurgery 625

"Attention Neurons" in the Monkey Parietal Cortex 626

"Recognition Neurons" in the Monkey Temporal Cortex 627

"Planning Neurons" in the Monkey Frontal Cortex 630

Box C Neuropsychological Testing 632

Box D Brain Size and Intelligence 634

Summary 635

Chapter 26 Language and Speech 637

Overview 637

Language Is Both Localized and Lateralized 637

Aphasias 638

Box A Speech 640

Box B Do Other Animals Have Language? 642

Box C Words and Meaning 645

A Dramatic Confirmation of Language Lateralization 646

Anatomical Differences between the Right and Left Hemispheres 648

Mapping Language Functions 649

Box D Language and Handedness 650

The Role of the Right Hemisphere in Language 654

Sign Language 655

Summary 656

Chapter 27 Sleep and Wakefulness 659

Overview 659

Why Do Humans (and Many Other Animals) Sleep? 659

Box A Styles of Sleep in Different Species 661

The Circadian Cycle of Sleep and Wakefulness	662
Stages of Sleep	665
Box B Molecular Mechanisms of Biological Clocks	666
Box C Electroencephalography	668
Physiological Changes in Sleep States	671
The Possible Functions of REM Sleep and Dreaming	671
Neural Circuits Governing Sleep	674
Box D Consciousness	675
Thalamocortical Interactions	679
Sleep Disorders	681
Box E Drugs and Sleep	682
Summary	684

Chapter 28 Emotions 687

Overview	687
Physiological Changes Associated with Emotion	687
The Integration of Emotional Behavior	688
Box A Facial Expressions: Pyramidal and Extrapyramidal Contributions	690
The Limbic System	693
Box B The Anatomy of the Amygdala	696
The Importance of the Amygdala	697
Box C The Reasoning Behind an Important Discovery	698
The Relationship between Neocortex and Amygdala	701
Box D Fear and the Human Amygdala: A Case Study	702
Box E Affective Disorders	704
Cortical Lateralization of Emotional Functions	705
Emotion, Reason, and Social Behavior	707
Summary	708

Chapter 29 Sex, Sexuality, and the Brain 711

Overview	711
Sexually Dimorphic Behavior	711
What Is Sex?	712
Box A The Development of Male and Female Phenotypes	714
Hormonal Influences on Sexual Dimorphism	715
Box B The Case of Bruce/Brenda	716
The Effect of Sex Hormones on Neural Circuitry	718

Box C The Actions of Sex Hormones	718
Other Central Nervous System Dimorphisms Specifically Related to Reproductive Behaviors	720
Brain Dimorphisms Related to Cognitive Function	728
Hormone-Sensitive Brain Circuits in Adult Animals	729
Summary	731

Chapter 30 Memory 733

Overview	733
Qualitative Categories of Human Memory	733
Temporal Categories of Memory	734
Box A Phylogenetic Memory	735
The Importance of Association in Information Storage	736
Forgetting	738
Box B Savant Syndrome	739
Brain Systems Underlying Declarative Memory Formation	741
Box C Clinical Cases That Reveal the Anatomical Substrate for Declarative Memories	742
Brain Systems Underlying Long-Term Storage of Declarative Memory	746
Brain Systems Underlying Nondeclarative Learning and Memory	748
Memory and Aging	749
Box D Alzheimer's Disease	750
Summary	753

Appendix A The Brainstem and Cranial Nerves 755

Appendix B Vascular Supply, the Meninges, and the Ventricular System 763

The Blood Supply of the Brain and Spinal Cord	763
The Blood-Brain Barrier	766
Box A Stroke	767
The Meninges	768
The Ventricular System	770

Glossary

Illustration Source References

Index

Preface

Whether judged in molecular, cellular, systemic, behavioral, or cognitive terms, the human nervous system is a stupendous piece of biological machinery. Given its accomplishments—all the artifacts of human culture, for instance—there is good reason for wanting to understand how the brain and the rest of the nervous system works. The debilitating and costly effects of neurological and psychiatric disease add a further sense of urgency to this quest. The aim of this book is to highlight the intellectual challenges and excitement—as well as the uncertainties—of what many see as the last great frontier of biological science. The information presented should serve as a starting point for undergraduates, medical students, graduate students in the neurosciences, and others who want to understand how the human nervous system operates. Like any other great challenge, neuroscience should be, and is, full of debate, dissension, and considerable fun. All these ingredients have gone into the construction of the third edition of this book; we hope they will be conveyed in equal measure to readers at all levels.

Acknowledgments

We are grateful to numerous colleagues who provided helpful contributions, criticisms and suggestions to this and previous editions. We particularly wish to thank Ralph Adolphs, David Amaral, Eva Anton, Gary Banker, Bob Barlow, Marlene Behrmann, Ursula Bellugi, Dan Blazer, Bob Burke, Roberto Cabeza, Nell Cant, Jim Cavanaugh, John Chapin, Milt Charlton, Michael Davis, Rob Deaner, Bob Desimone, Allison Doupe, Sasha du Lac, Jen Eilers, Anne Fausto-Sterling, Howard Fields, Elizabeth Finch, Nancy Forger, Jannon Fuchs, Michela Gallagher, Dana Garcia, Steve George, the late Patricia Goldman-Rakic, Mike Haglund, Zach Hall, Kristen Harris, Bill Henson, John Heuser, Jonathan Horton, Ron Hoy, Alan Humphrey, Jon Kaas, Jagmeet Kanwal, Herb Killackey, Len Kitzes, Arthur Lander, Story Landis, Simon LeVay, Darrell Lewis, Jeff Lichtman, Alan Light, Steve Lisberger, Donald Lo, Arthur Loewy, Ron Mangun, Eve Marder, Robert McCarley, Greg McCarthy, Jim McIlwain, Chris Muly, Vic Nadler, Ron Oppenheim, Larysa Pevny, Michael Platt, Franck Polleux, Scott Pomeroy, Rodney Radtke, Louis Reichardt, Marnie Riddle, Jamie Roitman, Steve Roper, John Rubenstein, Ben Rubin, David Rubin, Josh Sanes, Cliff Saper, Lynn Selement, Carla Shatz, Bill Snider, Larry Squire, John Staddon, Peter Strick, Warren Strittmatter, Joe Takahashi, Richard Weinberg, Jonathan Weiner, Christina Williams, Joel Winston, and Fulton Wong. It is understood, of course, that any errors are in no way attributable to our critics and advisors.

We also thank the students at Duke University Medical School as well as many other students and colleagues who provided suggestions for improvement of the last edition. Finally, we owe special thanks to Robert Reynolds and Nate O'Keefe, who labored long and hard to put the third edition together, and to Andy Sinauer, Graig Donini, Carol Wigg, Christopher Small, Janice Holabird, and the rest of the staff at Sinauer Associates for their outstanding work and high standards.

Supplements to Accompany NEUROSCIENCE Third Edition



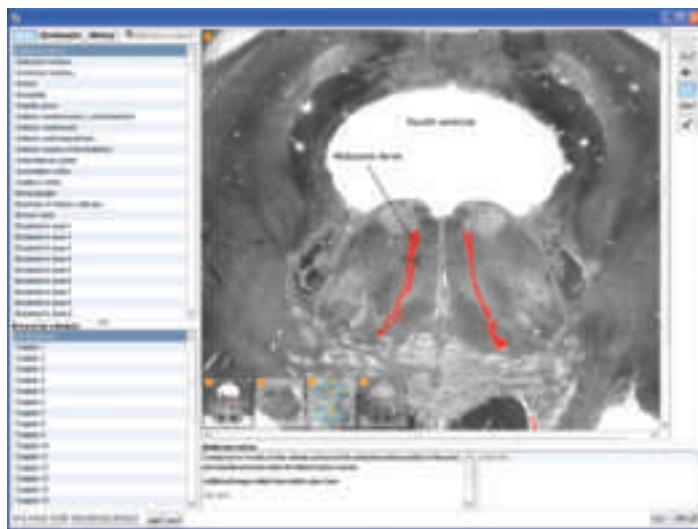
For the Student

Sylvius for Neuroscience:

A Visual Glossary of Human Neuroanatomy (CD-ROM)

S. Mark Williams, Leonard E. White, and Andrew C. Mace

Sylvius for Neuroscience: A Visual Glossary of Human Neuroanatomy, included in every copy of the textbook, is an interactive CD reference guide to the structure of the human nervous system. By entering a corresponding page number from the textbook, students can quickly search the CD for any neuroanatomical structure or term and view corresponding images and animations. Descriptive information is provided with all images and animations. Additionally, students can take notes on the content and share these with other *Sylvius* users. *Sylvius* is an essential study aid for learning basic human neuroanatomy.



Sylvius for Neuroscience features:

- Over 400 neuroanatomical structures and terms.
- High-resolution images.
- Animations of pathways and 3-D reconstructions.
- Definitions and descriptions.
- Audio pronunciations.
- A searchable glossary.
- Categories of anatomical structures and terms (e.g., cranial nerves, spinal cord tracts, lobes, cortical areas, etc.), that can be easily browsed. In addition, structures can be browsed by textbook chapter.

- Images and text relevant to the textbook: Icons in the textbook indicate specific content on the CD. By entering a textbook page number, students can automatically load the relevant images and text.
- A history feature that allows the student to quickly reload recently viewed structures.
- A bookmark feature that adds bookmarks to structures of interest; bookmarks are automatically stored on the student's computer.
- A notes feature that allows students to type notes for any selected structure; notes are automatically saved on the student's computer and can be shared among students and instructors (i.e., imported and exported).
- A self-quiz mode that allows for testing on structure identification and functional information.
- A print feature that formats images and text for printed output.
- An image zoom tool.

For the Instructor

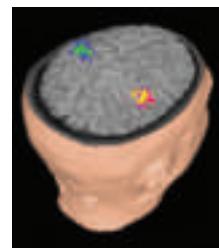
Instructor's Resource CD (ISBN 0-87893-750-1)

This expanded resource includes all the figures and tables from the textbook in JPEG format, reformatted and relabeled for optimal readability. Also included are ready-to-use PowerPoint® presentations of all figures and tables. In addition, new for the Third Edition, the Instructor's Resource CD includes a set of short-answer study questions for each chapter in Microsoft® Word® format.

Overhead Transparencies (ISBN 0-87893-751-X)

This set includes 100 illustrations (approximately 150 transparencies), selected from throughout the textbook for teaching purposes. These are relabeled and optimized for projection in classrooms.

Chapter 1



Studying the Nervous Systems of Humans and Other Animals

Overview

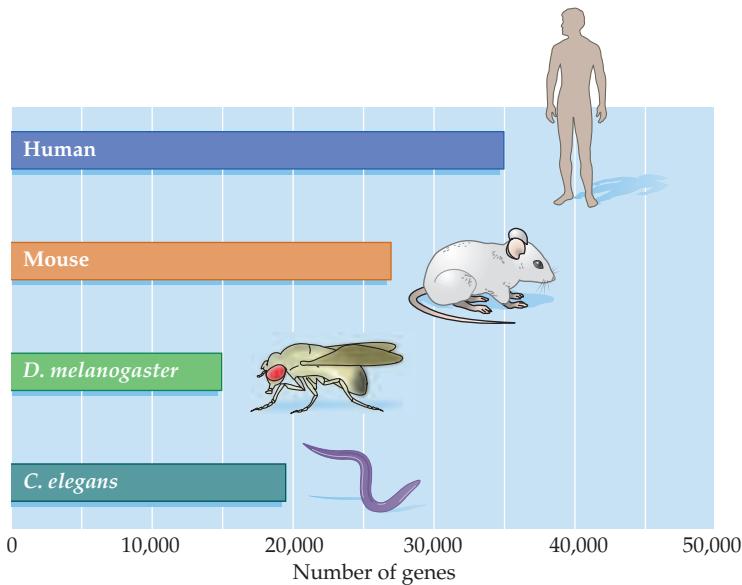
Neuroscience encompasses a broad range of questions about how nervous systems are organized, and how they function to generate behavior. These questions can be explored using the analytical tools of genetics, molecular and cell biology, systems anatomy and physiology, behavioral biology, and psychology. The major challenge for a student of neuroscience is to integrate the diverse knowledge derived from these various levels of analysis into a more or less coherent understanding of brain structure and function (one has to qualify this statement because so many questions remain unanswered). Many of the issues that have been explored successfully concern how the principal cells of any nervous system—neurons and glia—perform their basic functions in anatomical, electrophysiological, and molecular terms. The varieties of neurons and supporting glial cells that have been identified are assembled into ensembles called neural circuits, and these circuits are the primary components of neural systems that process specific types of information. Neural systems comprise neurons and circuits in a number of discrete anatomical locations in the brain. These systems subserve one of three general functions. Sensory systems represent information about the state of the organism and its environment, motor systems organize and generate actions; and associational systems link the sensory and motor sides of the nervous system, providing the basis for “higher-order” functions such as perception, attention, cognition, emotions, rational thinking, and other complex brain functions that lie at the core of understanding human beings, their history and their future.

Genetics, Genomics, and the Brain

The recently completed sequencing of the genome in humans, mice, the fruit fly *Drosophila melanogaster*, and the nematode worm *Caenorhabditis elegans* is perhaps the logical starting point for studying the brain and the rest of the nervous system; after all, this inherited information is also the starting point of each individual organism. The relative ease of obtaining, analyzing, and correlating gene sequences with neurobiological observations has facilitated a wealth of new insights into the basic biology of the nervous system. In parallel with studies of normal nervous systems, the genetic analysis of human pedigrees with various brain diseases has led to a widespread sense that it will soon be possible to understand and treat disorders long considered beyond the reach of science and medicine.

A gene consists of DNA sequences called **exons** that are transcribed into a messenger RNA and subsequently a protein. The set of exons that defines

Figure 1.1 Estimates of the number of genes in the human genome, as well as in the genomes of the mouse, the fruit fly *Drosophila melanogaster*, and the nematode worm *Caenorhabditis elegans*.



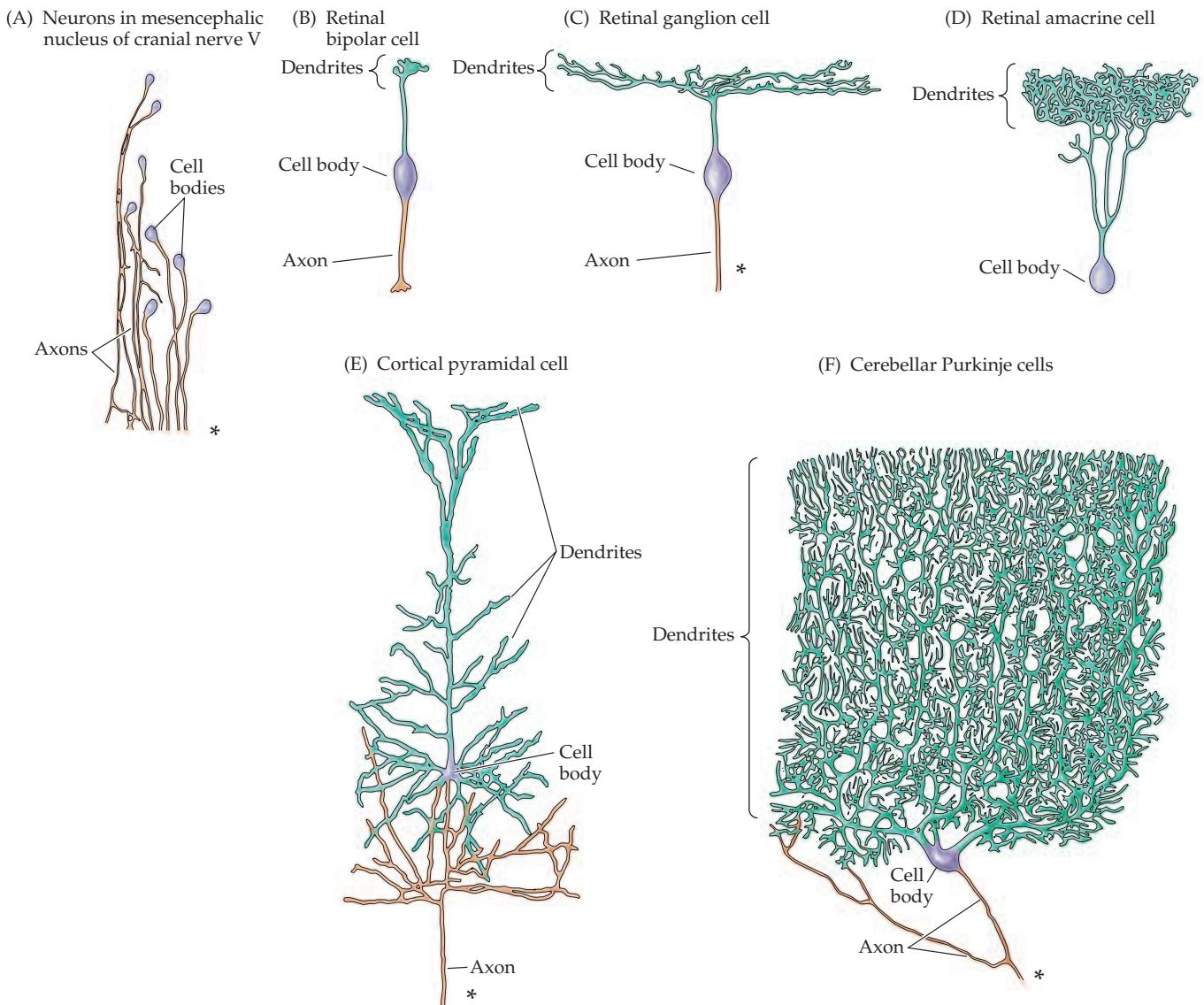
the transcript of any gene is flanked by upstream (or 5') and downstream (or 3') regulatory sequences that control gene expression. In addition, sequences between exons—called **introns**—further influence transcription. Of the approximately 35,000 genes in the human genome, a majority are expressed in the developing and adult brain; the same is true in mice, flies, and worms—the species commonly used in modern genetics (and increasingly in neuroscience) (Figure 1.1). Nevertheless, very few genes are *uniquely* expressed in neurons, indicating that nerve cells share most of the basic structural and functional properties of other cells. Accordingly, most “brain-specific” genetic information must reside in the remainder of nucleic acid sequences—regulatory sequences and introns—that control the timing, quantity, variability and cellular specificity of gene expression.

One of the most promising dividends of sequencing the human genome has been the realization that one or a few genes, when altered (mutated), can begin to explain some aspects of neurological and psychiatric diseases. Before the “postgenomic era” (which began following completion of the sequencing of the human genome), many of the most devastating brain diseases remained largely mysterious because there was little sense of how or why the normal biology of the nervous system was compromised. The identification of genes correlated with disorders such as Huntington’s disease, Parkinson’s disease, Alzheimer’s disease, major depression, and schizophrenia has provided a promising start to understanding these pathological processes in a much deeper way (and thus devising rational therapies).

Genetic and genomic information alone do not completely explain how the brain normally works or how disease processes disrupt its function. To achieve these goals it is equally essential to understand the cell biology, anatomy, and physiology of the brain in health as well as disease.

The Cellular Components of the Nervous System

Early in the nineteenth century, the cell was recognized as the fundamental unit of all living organisms. It was not until well into the twentieth century, however, that neuroscientists agreed that nervous tissue, like all other organs, is made up of these fundamental units. The major reason was that the first generation of “modern” neurobiologists in the nineteenth century had difficulty resolving the unitary nature of nerve cells with the microscopes and cell staining techniques that were then available. This inade-



quacy was exacerbated by the extraordinarily complex shapes and extensive branches of individual nerve cells, which further obscured their resemblance to the geometrically simpler cells of other tissues (Figures 1.2–1.4). As a result, some biologists of that era concluded that each nerve cell was connected to its neighbors by protoplasmic links, forming a continuous nerve cell network, or *reticulum*. The “reticular theory” of nerve cell communication, which was championed by the Italian neuropathologist Camillo Golgi (for whom the Golgi apparatus in cells is named), eventually fell from favor and was replaced by what came to be known as the “neuron doctrine.” The major proponents of this new perspective were the Spanish neuroanatomist Santiago Ramón y Cajal and the British physiologist Charles Sherrington.

The contrasting views represented by Golgi and Cajal occasioned a spirited debate in the early twentieth century that set the course of modern neuroscience. Based on light microscopic examination of nervous tissue stained with silver salts according to a method pioneered by Golgi, Cajal argued persuasively that nerve cells are discrete entities, and that they communicate

Figure 1.2 Examples of the rich variety of nerve cell morphologies found in the human nervous system. Tracings are from actual nerve cells stained by impregnation with silver salts (the so-called Golgi technique, the method used in the classical studies of Golgi and Cajal). Asterisks indicate that the axon runs on much farther than shown. Note that some cells, like the retinal bipolar cell, have a very short axon, and that others, like the retinal amacrine cell, have no axon at all. The drawings are not all at the same scale.

with one another by means of specialized contacts that Sherrington called “synapses.” The work that framed this debate was recognized by the award of the Nobel Prize for Physiology or Medicine in 1906 to both Golgi and Cajal (the joint award suggests some ongoing concern about just who was correct, despite Cajal’s overwhelming evidence). The subsequent work of Sherrington and others demonstrating the transfer of electrical signals at synaptic junctions between nerve cells provided strong support of the “neuron doctrine,” but challenges to the autonomy of individual neurons remained. It was not until the advent of electron microscopy in the 1950s that any lingering doubts about the discreteness of neurons were resolved. The high-magnification, high-resolution pictures that could be obtained with the electron microscope clearly established that nerve cells are functionally independent units; such pictures also identified the specialized cellular junctions that Sherrington had named synapses (see Figures 1.3 and 1.4).

The histological studies of Cajal, Golgi, and a host of successors led to the further consensus that the cells of the nervous system can be divided into two broad categories: **nerve cells** (or **neurons**), and supporting cells called **neuroglia** (or simply **glia**; see Figure 1.5). Nerve cells are specialized for electrical signaling over long distances, and understanding this process represents one of the more dramatic success stories in modern biology (and the subject of Unit I of this book). Supporting cells, in contrast, are not capable of electrical signaling; nevertheless, they have several essential functions in the developing and adult brain.

Neurons

Neurons and glia share the complement of organelles found in all cells, including the endoplasmic reticulum and Golgi apparatus, mitochondria, and a variety of vesicular structures. In neurons, however, these organelles are often more prominent in distinct regions of the cell. In addition to the distribution of organelles and subcellular components, neurons and glia are in some measure different from other cells in the specialized fibrillar or tubular proteins that constitute the cytoskeleton (Figures 1.3 and 1.4). Although many of these proteins—isoforms of actin, tubulin, and myosin, as well as several others—are found in other cells, their distinctive organization in neurons is critical for the stability and function of neuronal processes and synaptic junctions. The filaments, tubules, vesicular motors, and scaffolding proteins of neurons orchestrate the growth of axons and dendrites; the trafficking and appropriate positioning of membrane components, organelles, and vesicles; and the active processes of exocytosis and endocytosis that underlie synaptic communication. Understanding the ways in which these molecular components are used to insure the proper development and function of neurons and glia remains a primary focus of modern neurobiology.

The basic cellular organization of neurons resembles that of other cells; however, they are clearly distinguished by specialization for intercellular communication. This attribute is apparent in their overall morphology, in the specific organization of their membrane components for electrical signaling, and in the structural and functional intricacies of the synaptic contacts between neurons (see Figures 1.3 and 1.4). The most obvious sign of neuronal specialization for communication via electrical signaling is the extensive branching of neurons. The most salient aspect of this branching for typical nerve cells is the elaborate arborization of **dendrites** that arise from the neuronal cell body (also called *dendritic branches* or *dendritic processes*). Dendrites are the primary target for synaptic input from other neurons and are

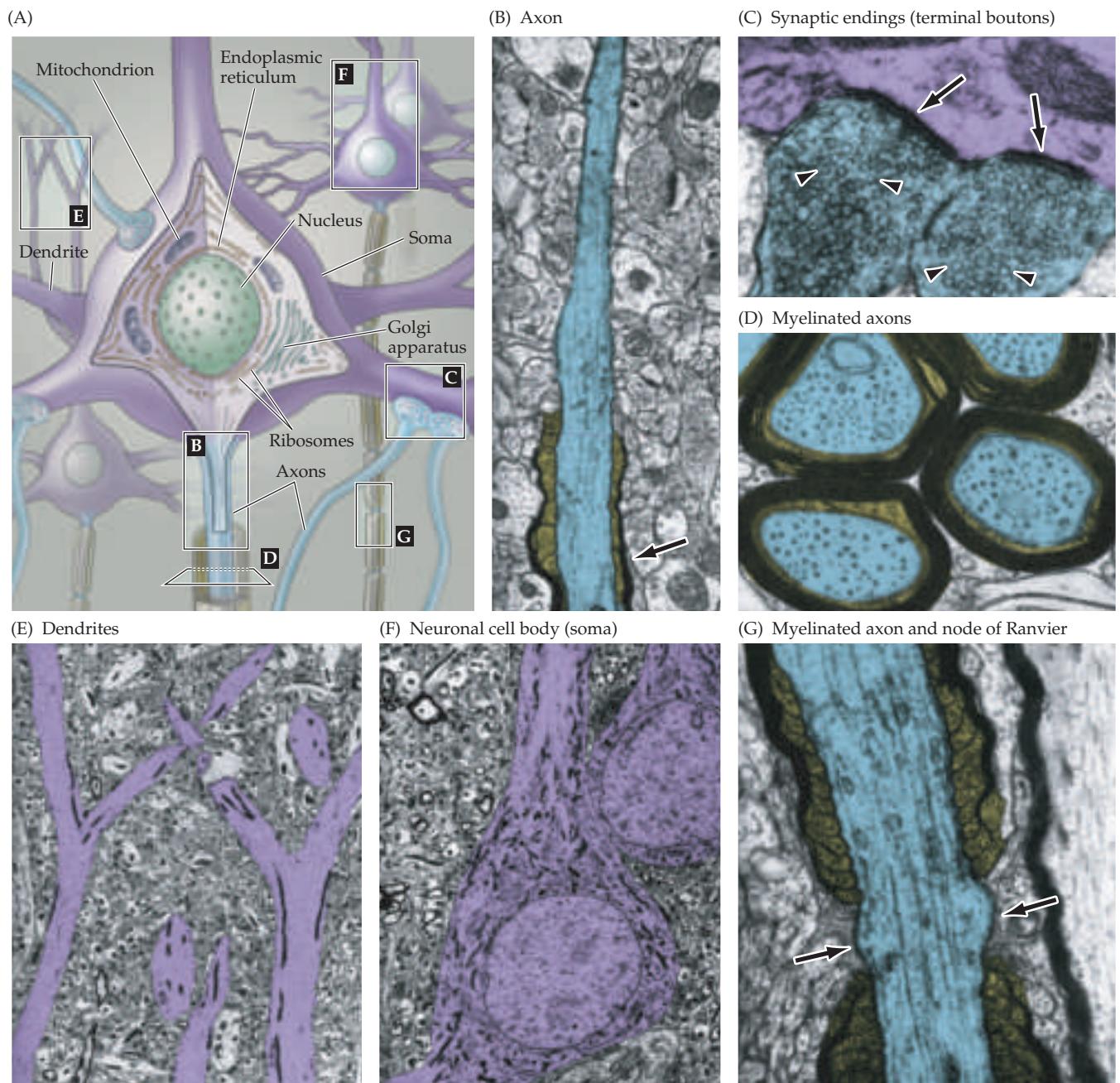
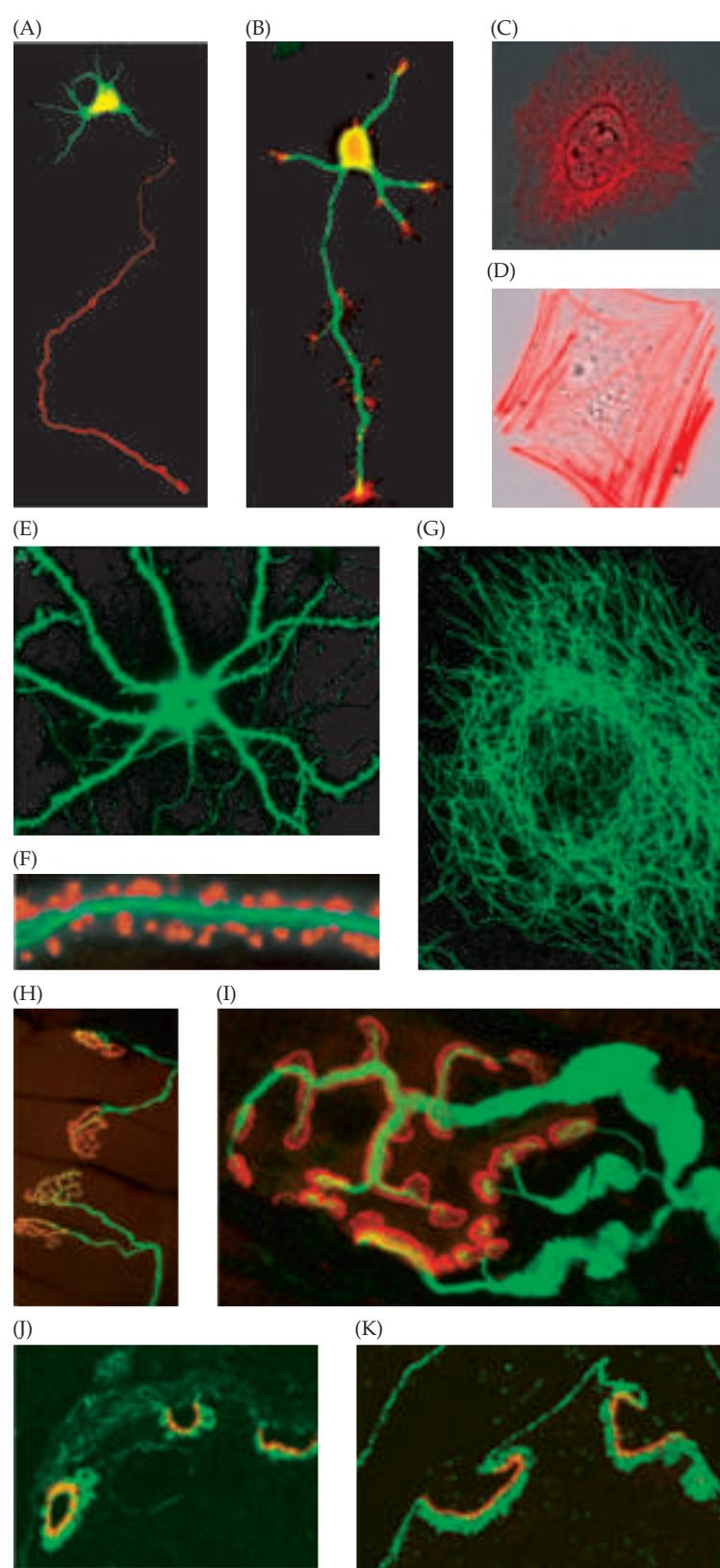


Figure 1.3 The major light and electron microscopical features of neurons. (A) Diagram of nerve cells and their component parts. (B) Axon initial segment (blue) entering a myelin sheath (gold). (C) Terminal boutons (blue) loaded with synaptic vesicles (arrowheads) forming synapses (arrows) with a dendrite (purple). (D) Transverse section of axons (blue) ensheathed by the processes of oligodendrocytes (gold). (E) Apical dendrites (purple) of cortical pyramidal cells. (F) Nerve cell bodies (purple) occupied by large round nuclei. (G) Portion of a myelinated axon (blue) illustrating the intervals between adjacent segments of myelin (gold) referred to as nodes of Ranvier (arrows). (Micrographs from Peters et al., 1991.)

Figure 1.4 Distinctive arrangement of cytoskeletal elements in neurons. (A) The cell body, axons, and dendrites are distinguished by the distribution of tubulin (green throughout cell) versus other cytoskeletal elements—in this case, Tau (red), a microtubule-binding protein found only in axons. (B) The strikingly distinct localization of actin (red) to the growing tips of axonal and dendritic processes is shown here in cultured neuron taken from the hippocampus. (C) In contrast, in a cultured epithelial cell, actin (red) is distributed in fibrils that occupy most of the cell body. (D) In astroglial cells in culture, actin (red) is also seen in fibrillar bundles. (E) Tubulin (green) is seen throughout the cell body and dendrites of neurons. (F) Although tubulin is a major component of dendrites, extending into spines, the head of the spine is enriched in actin (red). (G) The tubulin component of the cytoskeleton in non-neuronal cells is arrayed in filamentous networks. (H–K) Synapses have a distinct arrangement of cytoskeletal elements, receptors, and scaffold proteins. (H) Two axons (green; tubulin) from motor neurons are seen issuing two branches each to four muscle fibers. The red shows the clustering of postsynaptic receptors (in this case for the neurotransmitter acetylcholine). (I) A higher power view of a single motor neuron synapse shows the relationship between the axon (green) and the postsynaptic receptors (red). (J) The extracellular space between the axon and its target muscle is shown in green. (K) The clustering of scaffolding proteins (in this case, dystrophin) that localize receptors and link them to other cytoskeletal elements is shown in green. (A courtesy of Y. N. Jan; B courtesy of E. Dent and F. Gertler; C courtesy of D. Arneman and C. Otey; D courtesy of A. Gonzales and R. Cheney; E from Sheng, 2003; F from Matus, 2000; G courtesy of T. Salmon et al.; H–K courtesy of R. Sealock.)



also distinguished by their high content of ribosomes as well as specific cytoskeletal proteins that reflect their function in receiving and integrating information from other neurons. The spectrum of neuronal geometries ranges from a small minority of cells that lack dendrites altogether to neurons with dendritic arborizations that rival the complexity of a mature tree (see Figure 1.2). The number of inputs that a particular neuron receives depends on the complexity of its dendritic arbor: nerve cells that lack dendrites are innervated by (thus, receive electrical signals from) just one or a few other nerve cells, whereas those with increasingly elaborate dendrites are innervated by a commensurately larger number of other neurons.

The synaptic contacts made on dendrites (and, less frequently, on neuronal cell bodies) comprise a special elaboration of the secretory apparatus found in most polarized epithelial cells. Typically, the **presynaptic terminal** is immediately adjacent to a **postsynaptic specialization** of the target cell (see Figure 1.3). For the majority of synapses, there is no physical continuity between these pre- and postsynaptic elements. Instead, pre- and postsynaptic components communicate via secretion of molecules from the presynaptic terminal that bind to receptors in the postsynaptic specialization. These molecules must traverse an interval of extracellular space between pre- and postsynaptic elements called the **synaptic cleft**. The synaptic cleft, however, is not simply a space to be traversed; rather, it is the site of extracellular proteins that influence the diffusion, binding, and degradation of molecules secreted by the presynaptic terminal (see Figure 1.4). The number of synaptic inputs received by each nerve cell in the human nervous system varies from 1 to about 100,000. This range reflects a fundamental purpose of nerve cells, namely to integrate information from other neurons. The number of synaptic contacts from different presynaptic neurons onto any particular cell is therefore an especially important determinant of neuronal function.

The information conveyed by synapses on the neuronal dendrites is integrated and “read out” at the origin of the **axon**, the portion of the nerve cell specialized for signal conduction to the next site of synaptic interaction (see Figures 1.2 and 1.3). The axon is a unique extension from the neuronal cell body that may travel a few hundred micrometers (μm ; usually called microns) or much farther, depending on the type of neuron and the size of the species. Moreover, the axon also has a distinct cytoskeleton whose elements are central for its functional integrity (see Figure 1.4). Many nerve cells in the human brain (as well as that of other species) have axons no more than a few millimeters long, and a few have no axons at all.

Relatively short axons are a feature of **local circuit neurons** or **interneurons** throughout the brain. The axons of projection neurons, however, extend to distant targets. For example, the axons that run from the human spinal cord to the foot are about a meter long. The electrical event that carries signals over such distances is called the **action potential**, which is a self-regenerating wave of electrical activity that propagates from its point of initiation at the cell body (called the **axon hillock**) to the terminus of the axon where synaptic contacts are made. The target cells of neurons include other nerve cells in the brain, spinal cord, and autonomic ganglia, and the cells of muscles and glands throughout the body.

The chemical and electrical process by which the information encoded by action potentials is passed on at synaptic contacts to the next cell in a pathway is called **synaptic transmission**. Presynaptic terminals (also called *synaptic endings*, *axon terminals*, or *terminal boutons*) and their postsynaptic specializations are typically **chemical synapses**, the most abundant type of

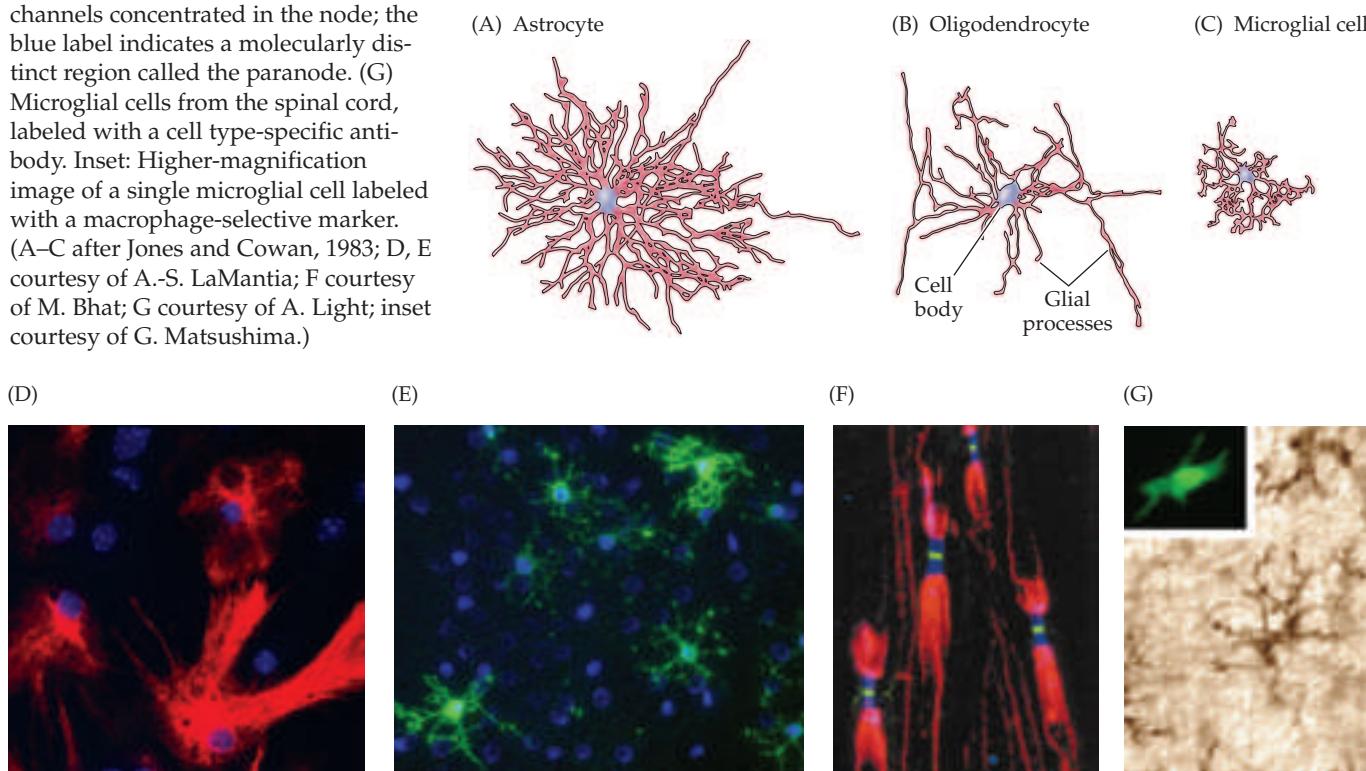
synapse in the nervous system. Another type, the electrical synapse, is far more rare (see Chapter 5). The secretory organelles in the presynaptic terminal of chemical synapses are **synaptic vesicles** (see Figure 1.3), which are generally spherical structures filled with **neurotransmitter** molecules. The positioning of synaptic vesicles at the presynaptic membrane and their fusion to initiate neurotransmitter release is regulated by a number of proteins either within or associated with the vesicle. The neurotransmitters released from synaptic vesicles modify the electrical properties of the target cell by binding to **neurotransmitter receptors** (Figure 1.4), which are localized primarily at the postsynaptic specialization.

The intricate and concerted activity of neurotransmitters, receptors, related cytoskeletal elements, and signal transduction molecules are thus the basis for nerve cells communicating with one another, and with effector cells in muscles and glands.

Figure 1.5 Varieties of neuroglial cells. Tracings of an astrocyte (A), an oligodendrocyte (B), and a microglial cell (C) visualized using the Golgi method. The images are at approximately the same scale. (D) Astrocytes in tissue culture, labeled (red) with an antibody against an astrocyte-specific protein. (E) Oligodendroglial cells in tissue culture labeled with an antibody against an oligodendroglial-specific protein. (F) Peripheral axon are ensheathed by myelin (labeled red) except at a distinct region called the node of Ranvier. The green label indicates ion channels concentrated in the node; the blue label indicates a molecularly distinct region called the paranode. (G) Microglial cells from the spinal cord, labeled with a cell type-specific antibody. Inset: Higher-magnification image of a single microglial cell labeled with a macrophage-selective marker. (A–C after Jones and Cowan, 1983; D, E courtesy of A.-S. LaMantia; F courtesy of M. Bhat; G courtesy of A. Light; inset courtesy of G. Matsushima.)

Neuroglial Cells

Neuroglial cells—also referred to as glial cells or simply glia—are quite different from nerve cells. Glia are more numerous than neurons in the brain, outnumbering them by a ratio of perhaps 3 to 1. The major distinction is that glia do not participate directly in synaptic interactions and electrical signaling, although their supportive functions help define synaptic contacts and maintain the signaling abilities of neurons. Although glial cells also have complex processes extending from their cell bodies, these are generally less prominent than neuronal branches, and do not serve the same purposes as axons and dendrites (Figure 1.5).



The term *glia* (from the Greek word meaning “glue”) reflects the nineteenth-century presumption that these cells held the nervous system together in some way. The word has survived, despite the lack of any evidence that binding nerve cells together is among the many functions of glial cells. Glial roles that *are* well-established include maintaining the ionic milieu of nerve cells, modulating the rate of nerve signal propagation, modulating synaptic action by controlling the uptake of neurotransmitters at or near the synaptic cleft, providing a scaffold for some aspects of neural development, and aiding in (or impeding, in some instances) recovery from neural injury.

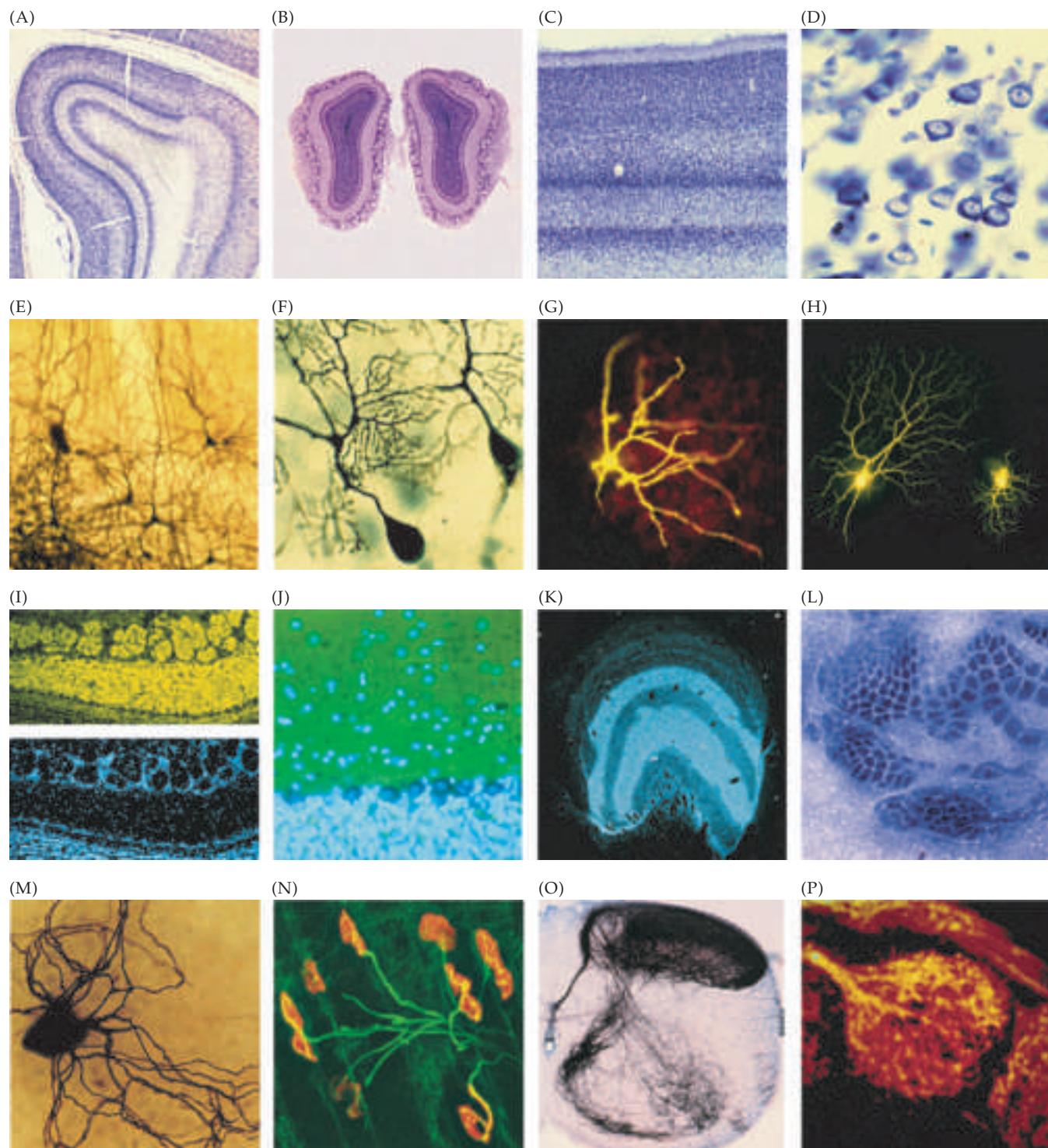
There are three types of glial cells in the mature central nervous system: astrocytes, oligodendrocytes, and microglial cells (see Figure 1.5). **Astrocytes**, which are restricted to the brain and spinal cord, have elaborate local processes that give these cells a starlike appearance (hence the prefix “astro”). A major function of astrocytes is to maintain, in a variety of ways, an appropriate chemical environment for neuronal signaling. **Oligodendrocytes**, which are also restricted to the central nervous system, lay down a laminated, lipid-rich wrapping called **myelin** around some, but not all, axons. Myelin has important effects on the speed of the transmission of electrical signals (see Chapter 3). In the peripheral nervous system, the cells that elaborate myelin are called **Schwann cells**.

Finally, **microglial cells** are derived primarily from hematopoietic precursor cells (although some may be derived directly from neural precursor cells). They share many properties with macrophages found in other tissues, and are primarily scavenger cells that remove cellular debris from sites of injury or normal cell turnover. In addition, microglia, like their macrophage counterparts, secrete signaling molecules—particularly a wide range of cytokines that are also produced by cells of the immune system—that can modulate local inflammation and influence cell survival or death. Indeed, some neurobiologists prefer to categorize microglia as a type of macrophage. Following brain damage, the number of microglia at the site of injury increases dramatically. Some of these cells proliferate from microglia resident in the brain, while others come from macrophages that migrate to the injured area and enter the brain via local disruptions in the cerebral vasculature.

Cellular Diversity in the Nervous System

Although the cellular constituents of the human nervous system are in many ways similar to those of other organs, they are unusual in their extraordinary numbers: the human brain is estimated to contain 100 billion neurons and several times as many supporting cells. More importantly, the nervous system has a greater range of distinct cell types—whether categorized by morphology, molecular identity, or physiological activity—than any other organ system (a fact that presumably explains why so many different genes are expressed in the nervous system; see above). The cellular diversity of any nervous system—including our own—undoubtedly underlies the capacity of the system to form increasingly complicated networks to mediate increasingly sophisticated behaviors.

For much of the twentieth century, neuroscientists relied on the same set of techniques developed by Cajal and Golgi to describe and categorize the diversity of cell types in the nervous system. From the late 1970s onward, however, new technologies made possible by the advances in cell and molecular biology provided investigators with many additional tools to discern the properties of neurons (Figure 1.6). Whereas general cell staining methods



showed mainly differences in cell size and distribution, antibody stains and probes for messenger RNA added greatly to the appreciation of distinctive types of neurons and glia in various regions of the nervous system. At the same time, new tract tracing methods using a wide variety of tracing substances allowed the interconnections among specific groups of neurons to be

◀ **Figure 1.6** Structural diversity in the nervous system demonstrated with cellular and molecular markers. *First row:* Cellular organization of different brain regions demonstrated with Nissl stains, which label nerve and glial cell bodies. (A) The cerebral cortex at the boundary between the primary and secondary visual areas. (B) The olfactory bulbs. (C) Differences in cell density in cerebral cortical layers. (D) Individual Nissl-stained neurons and glia at higher magnification. *Second row:* Classical and modern approaches to seeing individual neurons and their processes. (E) Golgi-labeled cortical pyramidal cells. (F) Golgi-labeled cerebellar Purkinje cells. (G) Cortical interneuron labeled by intracellular injection of a fluorescent dye. (H) Retinal neurons labeled by intracellular injection of fluorescent dye. *Third row:* Cellular and molecular approaches to seeing neural connections and systems. (I) At top, an antibody that detects synaptic proteins in the olfactory bulb; at bottom, a fluorescent label shows the location of cell bodies. (J) Synaptic zones and the location of Purkinje cell bodies in the cerebellar cortex labeled with synapse-specific antibodies (green) and a cell body marker (blue). (K) The projection from one eye to the lateral geniculate nucleus in the thalamus, traced with radioactive amino acids (the bright label shows the axon terminals from the eye in distinct layers of the nucleus). (L) The map of the body surface of a rat in the somatic sensory cortex, shown with a marker that distinguishes zones of higher synapse density and metabolic activity. *Fourth row:* Peripheral neurons and their projections. (M) An autonomic neuron labeled by intracellular injection of an enzyme marker. (N) Motor axons (green) and neuromuscular synapses (orange) in transgenic mice genetically engineered to express fluorescent proteins. (O) The projection of dorsal root ganglia to the spinal cord, demonstrated by an enzymatic tracer. (P) Axons of olfactory receptor neurons from the nose labeled in the olfactory bulb with a vital fluorescent dye. (G courtesy of L. C. Katz; H courtesy of C. J. Shatz; N,O courtesy of W. Snider and J. Lichtman; all others courtesy of A.-S. LaMantia and D. Purves.)

explored much more fully. Tracers can be introduced into either living or fixed tissue, and are transported along nerve cell processes to reveal their origin and termination. More recently, genetic and neuroanatomical methods have been combined to visualize the expression of fluorescent or other tracer molecules under the control of regulatory sequences of neural genes. This approach, which shows individual cells in fixed or living tissue in remarkable detail, allows nerve cells to be identified by both their transcriptional state and their structure. Finally, ways of determining the molecular identity and morphology of nerve cells can be combined with measurements of their physiological activity, thus illuminating structure–function relationships. Examples of these various approaches are shown in Figure 1.6.

Neural Circuits

Neurons never function in isolation; they are organized into ensembles or **neural circuits** that process specific kinds of information and provide the foundation of sensation, perception and behavior. The synaptic connections that define such circuits are typically made in a dense tangle of dendrites, axons terminals, and glial cell processes that together constitute what is called **neuropil** (the suffix *-pil* comes from the Greek word *pilos*, meaning “felt”; see Figure 1.3). The neuropil is thus the region between nerve cell bodies where most synaptic connectivity occurs.

Although the arrangement of neural circuits varies greatly according to the function being served, some features are characteristic of all such ensembles. Preeminent is the direction of information flow in any particular circuit, which is obviously essential to understanding its purpose. Nerve cells that

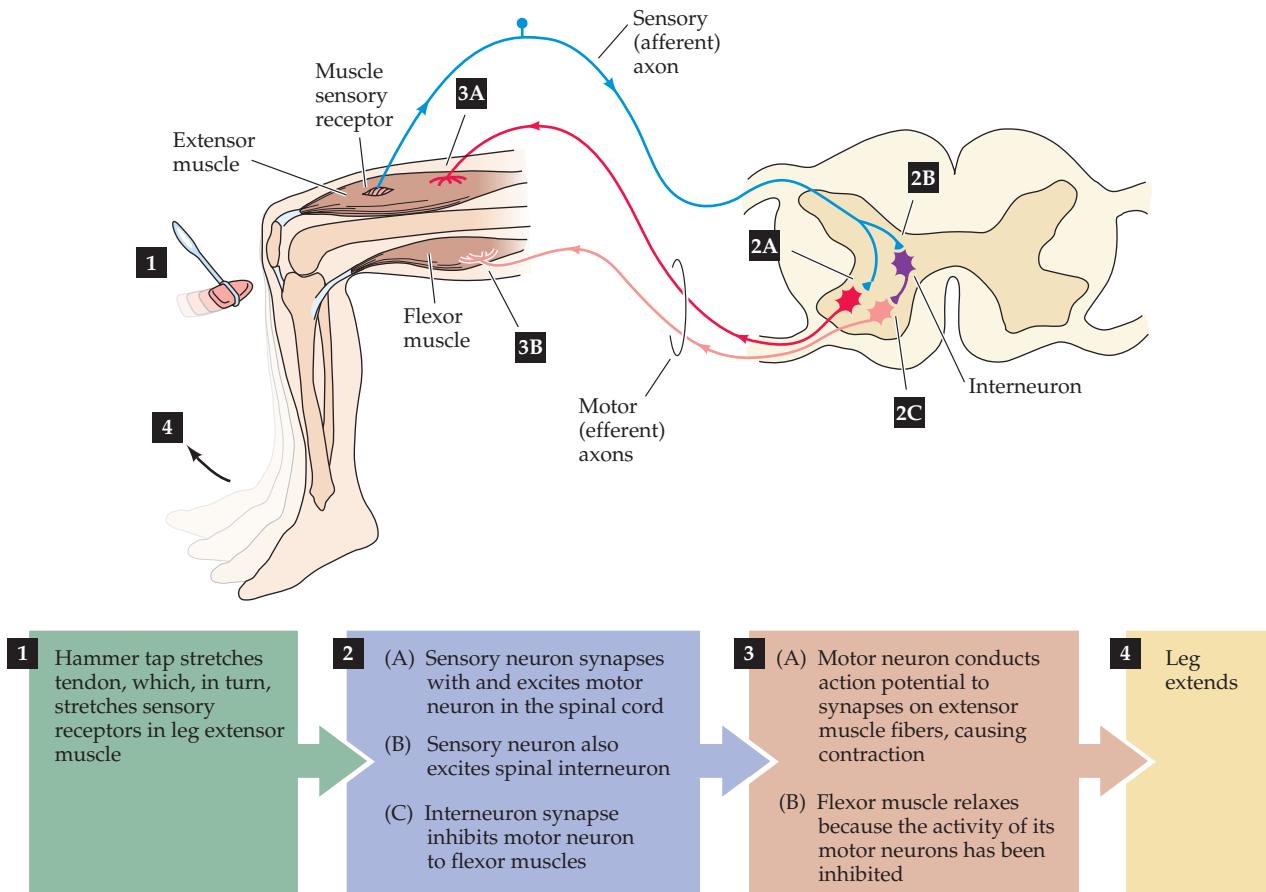


Figure 1.7 A simple reflex circuit, the knee-jerk response (more formally, the myotatic reflex), illustrates several points about the functional organization of neural circuits. Stimulation of peripheral sensors (a muscle stretch receptor in this case) initiates receptor potentials that trigger action potentials that travel centrally along the *afferent* axons of the sensory neurons. This information stimulates spinal motor neurons by means of synaptic contacts. The action potentials triggered by the synaptic potential in motor neurons travel peripherally in *efferent* axons, giving rise to muscle contraction and a behavioral response. One of the purposes of this particular reflex is to help maintain an upright posture in the face of unexpected changes.

carry information *toward* the brain or spinal cord (or farther centrally within the spinal cord and brain) are called **afferent neurons**; nerve cells that carry information *away* from the brain or spinal cord (or away from the circuit in question) are called **efferent neurons**. **Interneurons** or **local circuit neurons** only participate in the local aspects of a circuit, based on the short distances over which their axons extend. These three functional classes—afferent neurons, efferent neurons, and interneurons—are the basic constituents of all neural circuits.

A simple example of a neural circuit is the ensemble of cells that subserves the **myotatic spinal reflex** (the “knee-jerk” reflex; Figure 1.7). The afferent neurons of the reflex are **sensory neurons** whose cell bodies lie the **dorsal root ganglia** and whose peripheral axons terminate in sensory endings in skeletal muscles (the ganglia that serve this same of function for much of the head and neck are called **cranial nerve ganglia**; see Appendix A). The central axons of these afferent sensory neurons enter the the spinal cord where they terminate on a variety of central neurons concerned with the regulation of muscle tone, most obviously the **motor neurons** that determine the activity of the related muscles. These neurons constitute the efferent neurons as well as interneurons of the circuit. One group of these efferent neurons in the ventral horn of the spinal cord projects to the flexor muscles in the limb, and the other to extensor muscles. Spinal cord interneurons are the third element of this circuit. The interneurons receive synaptic contacts from sensory afferent neurons and make synapses on the efferent motor neurons that project to the

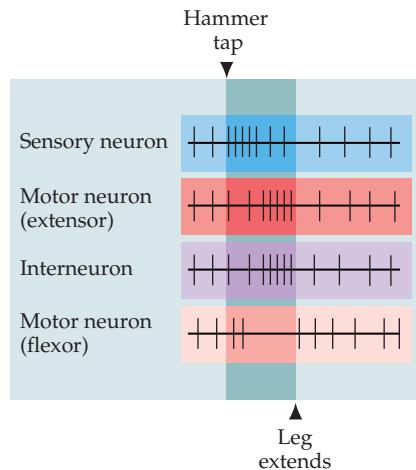
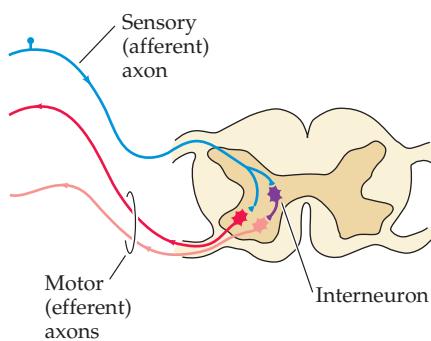


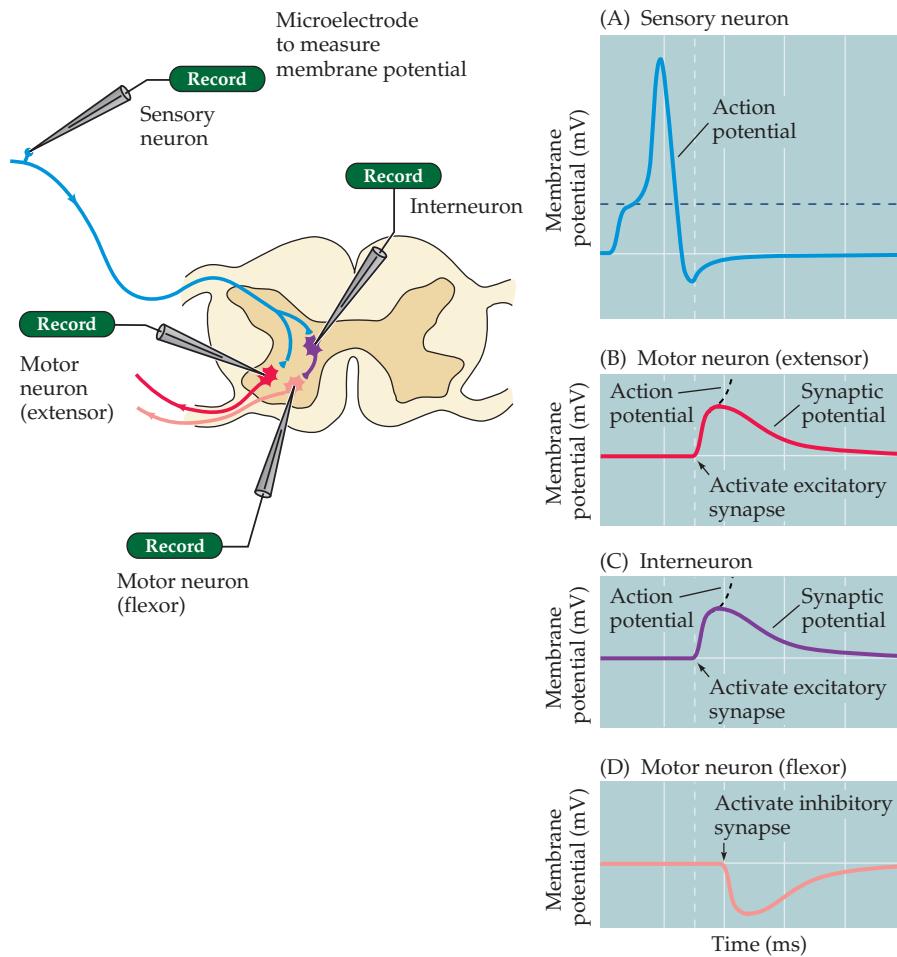
Figure 1.8 Relative frequency of action potentials (indicated by individual vertical lines) in different components of the myotatic reflex as the reflex pathway is activated. Notice the modulatory effect of the interneuron.

flexor muscles; therefore they are capable of modulating the input–output linkage. The excitatory synaptic connections between the sensory afferents and the extensor efferent motor neurons cause the extensor muscles to contract; at the same time, the interneurons activated by the afferents are inhibitory, and their activation diminishes electrical activity in flexor efferent motor neurons and causes the flexor muscles to become less active (Figure 1.8). The result is a complementary activation and inactivation of the synergist and antagonist muscles that control the position of the leg.

A more detailed picture of the events underlying the myotatic or any other circuit can be obtained by electrophysiological recording (Figure 1.9). There are two basic approaches to measuring the electrical activity of a nerve cell: **extracellular recording** (also referred to as single-unit recording), where an electrode is placed *near* the nerve cell of interest to detect its activity; and **intracellular recording**, where the electrode is placed *inside* the cell. Extracellular recordings primarily detect **action potentials**, the all-or-nothing changes in the potential across nerve cell membranes that convey information from one point to another in the nervous system. This sort of recording is particularly useful for detecting temporal patterns of action potential activity and relating those patterns to stimulation by other inputs, or to specific behavioral events. Intracellular recordings can detect the smaller, graded potential changes that trigger action potentials, and thus allow a more detailed analysis of communication between neurons within a circuit. These graded triggering potentials can arise at either sensory receptors or synapses and are called **receptor potentials** or **synaptic potentials**, respectively.

For the myotatic circuit, electrical activity can be measured both extracellularly and intracellularly, thus defining the functional relationships between neurons in the circuit. The pattern of action potential activity can be measured for each element of the circuit (afferents, efferents, and interneurons) before, during, and after a stimulus (see Figure 1.8). By comparing the onset, duration, and frequency of action potential activity in each cell, a functional picture of the circuit emerges. As a result of the stimulus, the sensory neuron is triggered to fire at higher frequency (i.e., more action potentials per unit time). This increase triggers a higher frequency of action potentials in both the extensor motor neurons and the interneurons. Concurrently, the inhibitory synapses made by the interneurons onto the flexor motor neurons cause the frequency of action potentials in these cells to decline. Using intracellular recording, it is possible to observe directly the potential changes underlying the synaptic connections of the myotatic reflex circuit (see Figure 1.9).

Figure 1.9 Intracellularly recorded responses underlying the myotatic reflex. (A) Action potential measured in a sensory neuron. (B) Postsynaptic triggering potential recorded in an extensor motor neuron. (C) Postsynaptic triggering potential in an interneuron. (D) Postsynaptic inhibitory potential in a flexor motor neuron. Such intracellular recordings are the basis for understanding the cellular mechanisms of action potential generation, and the sensory receptor and synaptic potentials that trigger these conducted signals.

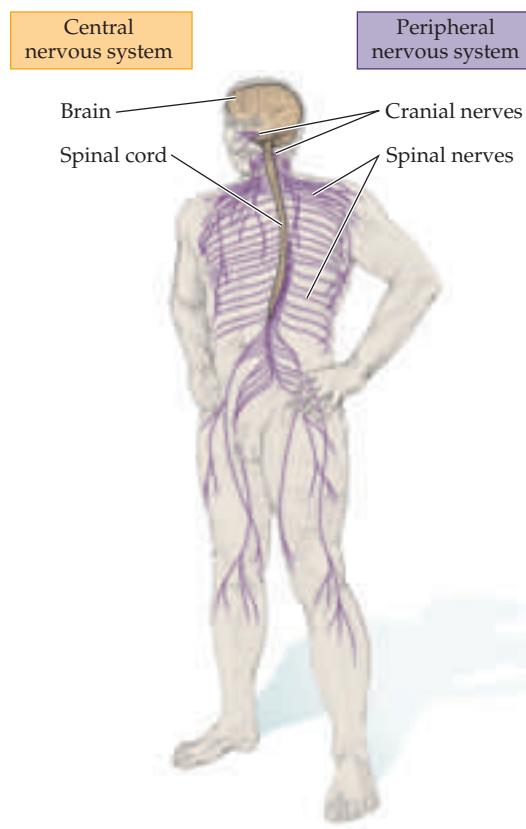


Overall Organization of the Human Nervous System

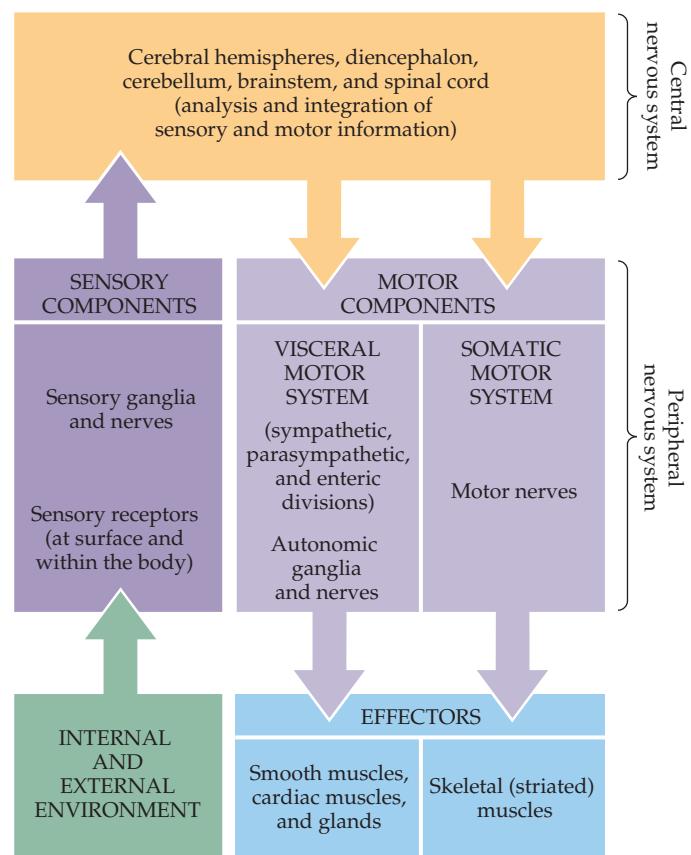
When considered together, circuits that process similar types of information comprise **neural systems** that serve broader behavioral purposes. The most general functional distinction divides such collections into **sensory systems** that acquire and process information from the environment (e.g., the visual system or the auditory system, see Unit II), and **motor systems** that respond to such information by generating movements and other behavior (see Unit III). There are, however, large numbers of cells and circuits that lie between these relatively well-defined input and output systems. These are collectively referred to as **associational systems**, and they mediate the most complex and least well-characterized brain functions (see Unit V).

In addition to these broad functional distinctions, neuroscientists and neurologists have conventionally divided the vertebrate nervous system anatomically into central and peripheral components (Figure 1.10). The **central nervous system**, typically referred to as the **CNS**, comprises the **brain** (cerebral hemispheres, diencephalon, cerebellum, and brainstem) and the **spinal cord** (see Appendix A for more information about the gross anatomical features of the CNS). The **peripheral nervous system (PNS)** includes the sensory neurons that link sensory receptors on the body surface or deeper within it with relevant processing circuits in the central nervous system. The motor portion of the peripheral nervous system in turn consists of two components. The motor axons that connect the brain and spinal cord to skeletal

(A)



(B)



muscles make up the **somatic motor division** of the peripheral nervous system, whereas the cells and axons that innervate smooth muscles, cardiac muscle, and glands make up the **visceral or autonomic motor division**.

Those nerve cell bodies that reside in the peripheral nervous system are located in **ganglia**, which are simply local accumulations of nerve cell bodies (and supporting cells). Peripheral axons are gathered into bundles called **nerves**, many of which are enveloped by the glial cells of the peripheral nervous system called **Schwann cells**. In the central nervous system, nerve cells are arranged in two different ways. **Nuclei** are local accumulations of neurons having roughly similar connections and functions; such collections are found throughout the cerebrum, brainstem and spinal cord. In contrast, **cortex** (plural, *cortices*) describes sheet-like arrays of nerve cells (again, consult Appendix A for additional information and illustrations). The cortices of the cerebral hemispheres and of the cerebellum provide the clearest example of this organizational principle.

Axons in the central nervous system are gathered into **tracts** that are more or less analogous to nerves in the periphery. Tracts that cross the midline of the brain are referred to as **commissures**. Two gross histological terms distinguish regions rich in neuronal cell bodies versus regions rich in axons. **Gray matter** refers to any accumulation of cell bodies and neuropil in the brain and spinal cord (e.g., nuclei or cortices), whereas **white matter**, named for its relatively light appearance resulting from the lipid content of myelin, refers to axon tracts and commissures.

Figure 1.10 The major components of the nervous system and their functional relationships. (A) The CNS (brain and spinal cord) and PNS (spinal and cranial nerves). (B) Diagram of the major components of the central and peripheral nervous systems and their functional relationships. Stimuli from the environment convey information to processing circuits within the brain and spinal cord, which in turn interpret their significance and send signals to peripheral effectors that move the body and adjust the workings of its internal organs.

The organization of the visceral motor division of the peripheral nervous system is a bit more complicated (see Chapter 20). Visceral motor neurons in the brainstem and spinal cord, the so-called preganglionic neurons, form synapses with peripheral motor neurons that lie in the **autonomic ganglia**. The motor neurons in autonomic ganglia innervate smooth muscle, glands, and cardiac muscle, thus controlling most involuntary (visceral) behavior. In the **sympathetic division** of the autonomic motor system, the ganglia lie along or in front of the vertebral column and send their axons to a variety of peripheral targets. In the **parasympathetic division**, the ganglia are found within the organs they innervate. Another component of the visceral motor system, called the **enteric system**, is made up of small ganglia as well as individual neurons scattered throughout the wall of the gut. These neurons influence gastric motility and secretion.

Neuroanatomical Terminology

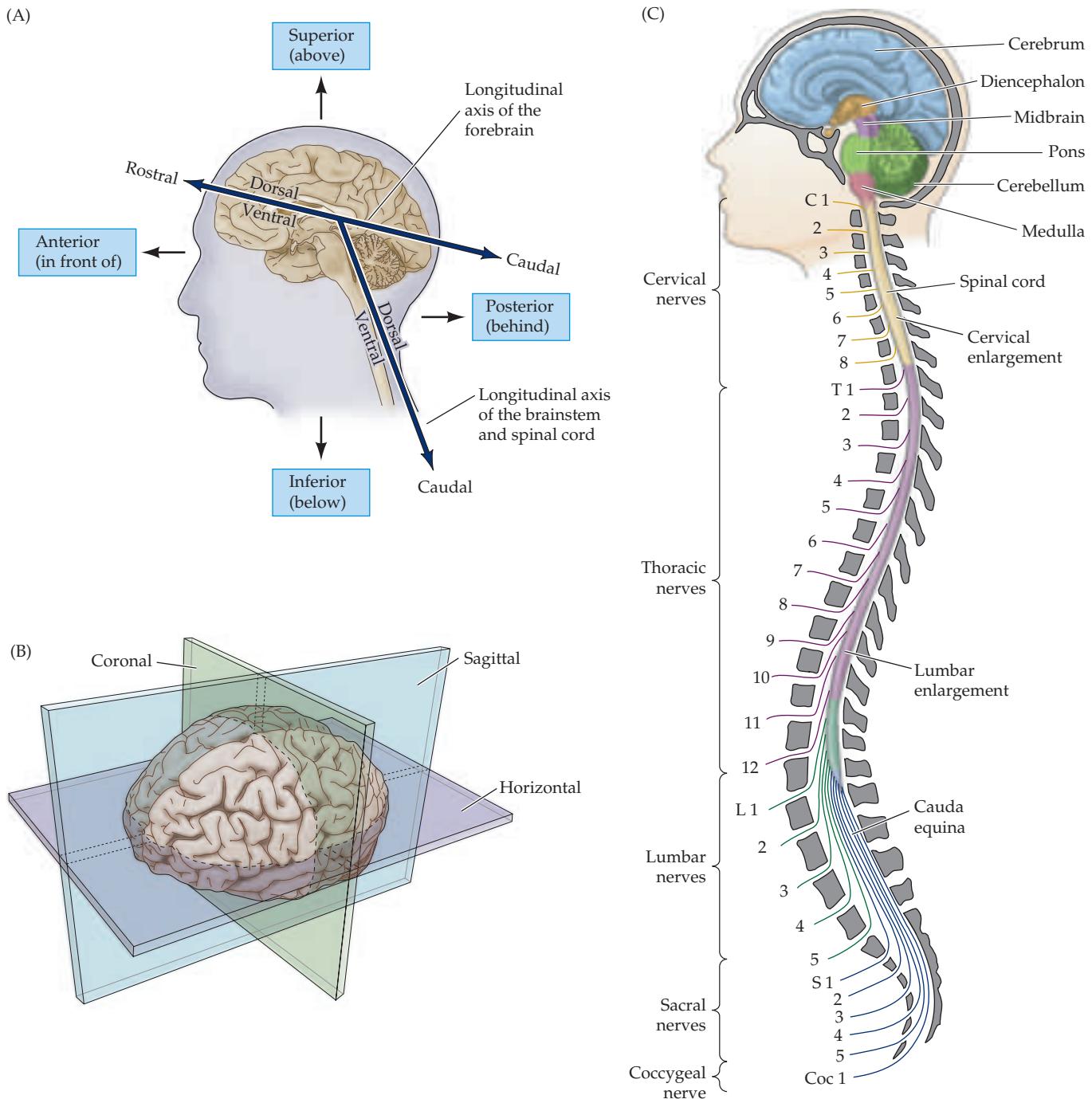
Describing the organization of any neural system requires a rudimentary understanding of anatomical terminology. The terms used to specify location in the central nervous system are the same as those used for the gross anatomical description of the rest of the body (Figure 1.11). Thus, *anterior* and *posterior* indicate front and back (head and tail); *rostral* and *caudal*, toward the head and tail; *dorsal* and *ventral*, top and bottom (back and belly); and *medial* and *lateral*, at the midline or to the side. Nevertheless, the comparison between these coordinates in the body versus the brain can be confusing. For the entire body these anatomical terms refer to the long axis, which is straight. The long axis of the central nervous system, however, has a bend in it. In humans and other bipeds, a compensatory tilting of the rostral-caudal axis for the brain is necessary to properly compare body axes to brain axes. Once this adjustment has been made, the other axes for the brain can be easily assigned.

The proper assignment of the anatomical axes then dictates the standard planes for histological sections or live images (see Box A) used to study the internal anatomy of the brain (see Figure 1.11B). **Horizontal sections** (also referred to as **axial** or **transverse** sections) are taken parallel to the rostral-caudal axis of the brain; thus, in an individual standing upright, such sections are parallel to the ground. Sections taken in the plane dividing the two hemispheres are **sagittal**, and can be further categorized as **midsagittal** and **parasagittal**, according to whether the section is near the midline (midsagittal)



Figure 1.11 A flexure in the long axis of the nervous system arose as humans evolved upright posture, leading to an approximately 120° angle between the long axis of the brainstem and that of the forebrain. The consequences of this flexure for anatomical terminology are indicated in (A). The terms *anterior*, *posterior*, *superior*, and *inferior* refer to the long axis of the body, which is straight. Therefore, these terms indicate the same direction for both the forebrain and the brainstem. In contrast, the terms *dorsal*, *ventral*, *rostral*, and *caudal* refer to the long axis of the central nervous system. The dorsal direction is toward the back for the brainstem and spinal cord, but toward the top of the head for the forebrain. The opposite direction is ventral. The rostral direction is toward the top of the head for the brainstem and spinal cord, but toward the face for the forebrain. The opposite direction is caudal. (B) The major planes of section used in cutting or imaging the brain. (C) The subdivisions and components of the central nervous system. (Note that the position of the brackets on the left side of the figure refers to the vertebrae, not the spinal segments.)

or more lateral (parasagittal). Sections in the plane of the face are called **coronal** or **frontal**. Different terms are usually used to refer to sections of the spinal cord. The plane of section orthogonal to the long axis of the cord is called **transverse**, whereas sections parallel to the long axis of the cord are called **longitudinal**. In a transverse section through the human spinal cord, the dorsal and ventral axes and the anterior and posterior axes indicate the same directions (see Figure 1.11). Tedious though this terminology may be, it



is essential for understanding the basic subdivisions of the nervous system (Figure 1.11C).

The Subdivisions of the Central Nervous System

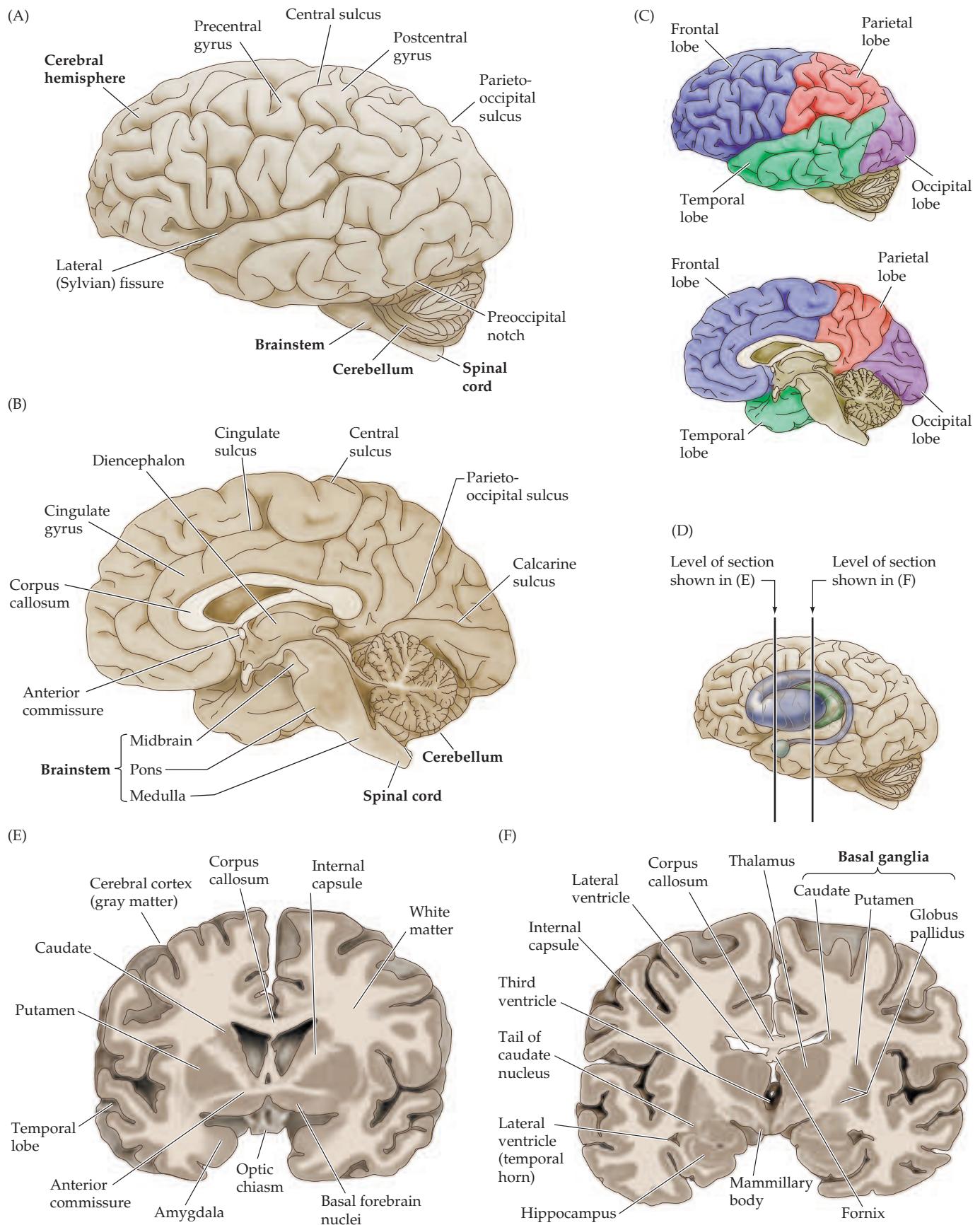
The central nervous system (defined as the brain and spinal cord) is usually considered to have seven basic parts: the **spinal cord**, the **medulla**, the **pons**, the **cerebellum**, the **midbrain**, the **diencephalon**, and the **cerebral hemispheres** (see Figures 1.10 and 1.11C). Running through all of these subdivisions are fluid-filled spaces called **ventricles** (a detailed account of the ventricular system can be found in Appendix B). These ventricles are the remnants of the continuous lumen initially enclosed by the neural plate as it rounded to become the neural tube during early development (see Chapter 21). Variations in the shape and size of the mature ventricular space are characteristic of each adult brain region. The medulla, pons, and midbrain are collectively called the **brainstem** and they surround the **4th ventricle** (medulla and pons) and **cerebral aqueduct** (midbrain). The diencephalon and cerebral hemispheres are collectively called the **forebrain**, and they enclose the **3rd and lateral ventricles**, respectively. Within the brainstem are the **cranial nerve nuclei** that either receive input from the **cranial sensory ganglia** mentioned earlier via the **cranial sensory nerves**, or give rise to axons that constitute the **cranial motor nerves** (see Appendix A).

The brainstem is also a conduit for several major tracts in the central nervous system that relay sensory information from the spinal cord and brainstem to the forebrain, or relay motor commands from forebrain back to motor neurons in the brainstem and spinal cord. Accordingly, detailed knowledge of the consequences of damage to the brainstem provides neurologists and other clinicians an essential tool in the localization and diagnosis of brain injury. The brainstem contains numerous additional nuclei that are involved in a myriad of important functions including the control of heart rate, respiration, blood pressure, and level of consciousness. Finally, one of the most prominent features of the brainstem is the **cerebellum**, which extends over much of its dorsal aspect. The cerebellum is essential for the coordination and planning of movements (see Chapter 18) as well as learning motor tasks and storing that information (see Chapter 30).

There are several anatomical subdivisions of the forebrain. The most obvious anatomical structures are the prominent **cerebral hemispheres** (Figure 1.12). In humans, the cerebral hemispheres (the outermost portions of which are continuous, highly folded sheets of cortex) are proportionally larger than in any other mammal, and are characterized by the **gyri** (singular, *gyrus*) or crests of folded cortical tissue, and **sulci** (singular, *sulcus*) the grooves that divide gyri from one another (as pictured on the cover of this book, for example). Although gyral and sulcal patterns vary from individual to individual, there are some fairly consistent landmarks that help divide the hemispheres into four **lobes**. The names of the lobes are derived from the cranial bones that overlie them: **occipital**, **temporal**, **parietal**, and **frontal**. A key feature of the surface anatomy of the cerebrum is the **central sulcus** located



Figure 1.12 Gross anatomy of the forebrain (A) Cerebral hemisphere surface anatomy, showing the four lobes of the brain and the major sulci and gyri. The ventricular system and basal ganglia can also be seen in this phantom view. (B) Mid-sagittal view showing the location of the hippocampus, amygdala, thalamus and hypothalamus.



roughly halfway between the rostral and caudal poles of the hemispheres (Figure 1.12A). This prominent sulcus divides the frontal lobe at the rostral end of the hemisphere from the more caudal parietal lobe. Prominent on either side of the central sulcus are the pre- and postcentral gyri. These gyri are also functionally significant in that the precentral gyrus contains the primary motor cortex important for the control of movement, and the postcentral gyrus contains the primary somatic sensory cortex which is important for the bodily senses (see below).

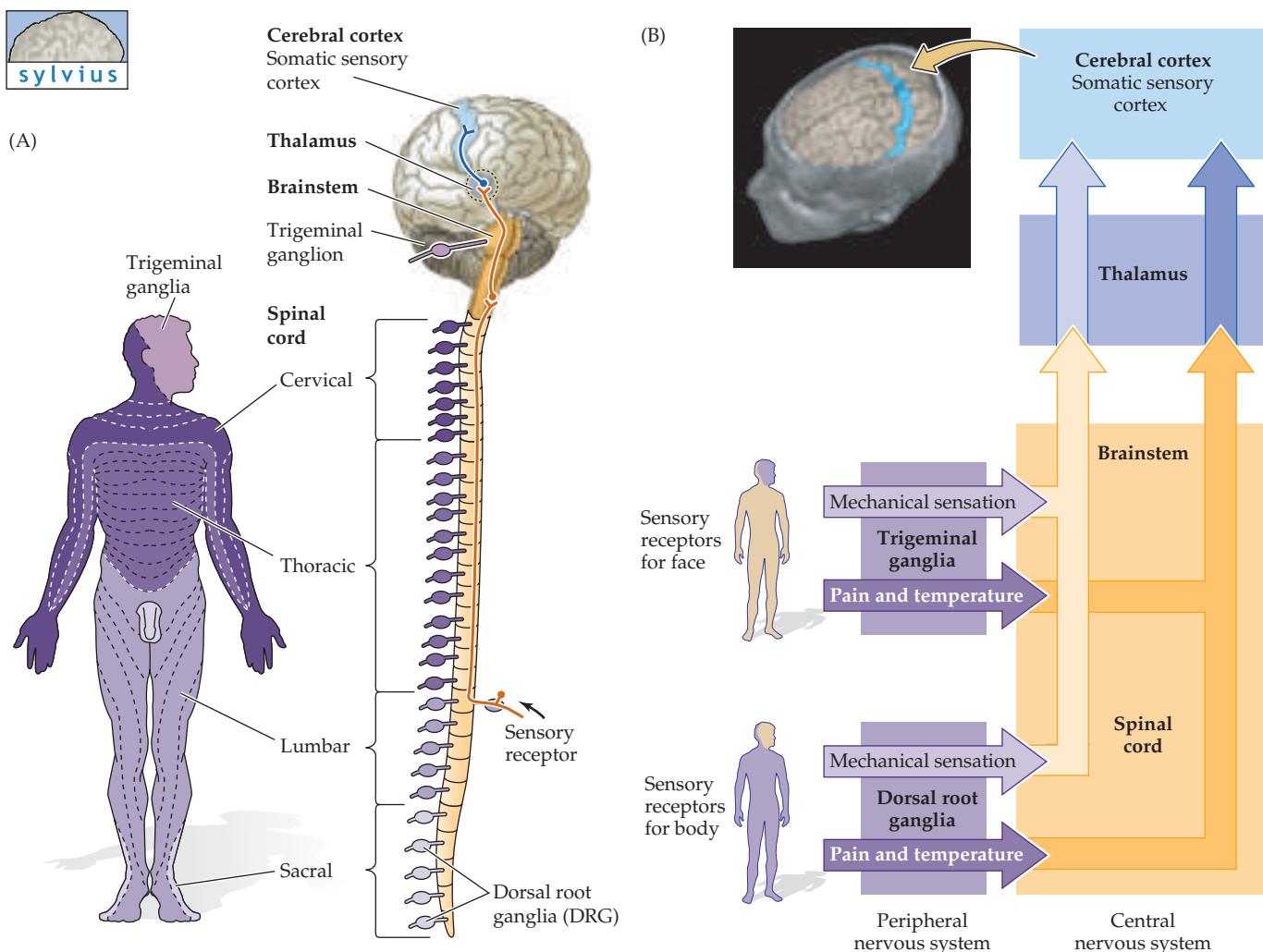
The remaining subdivisions of the forebrain lie deeper in the cerebral hemispheres (Figure 1.12B). The most prominent of these is the collection of deep structures involved in motor and cognitive processes collectively referred to as the **basal ganglia**. Other particularly important structures are the **hippocampus** and **amygdala** in the temporal lobes (these are vital substrates for memory and emotional behavior, respectively), and the **olfactory bulbs** (the central stations for processing chemosensory information arising from receptor neurons in the nasal cavity) on the anterior-inferior aspect of the frontal lobes. Finally, the **thalamus** lies in the diencephalon and is a critical relay for sensory information (although it has many other functions as well); the **hypothalamus**, which as the name implies lies below the thalamus, is the central organizing structure for the regulation of the body's many homeostatic functions (e.g., feeding, drinking, thermoregulation).

This rudimentary description of some prominent anatomical landmarks provides a framework for understanding how neurons resident in a number of widely distributed and distinct brain structures communicate with one another to define **neural systems** dedicated to encoding, processing and relaying specific sorts of information about aspects of the organism's environment, and then initiating and coordinating appropriate behavioral responses.

Organizational Principles of Neural Systems

These complex perceptual and motor capacities of the brain reflect the integrated function of various neural systems. The processing of somatic sensory information (arising from receptors in the skin, subcutaneous tissues, and the musculoskeletal system that respond to physical deformation at the body surface or displacement of muscles and joints) provides a convenient example. These widely distributed structures that participate in generating somatic sensations are referred to as the **somatic sensory system** (Figure 1.13). The components in the peripheral nervous system include the receptors distributed throughout the skin as well as in muscles and tendons, the related neurons in dorsal root ganglia, and neurons in some cranial ganglia. The central nervous system components include neurons in the spinal cord, as well as the long tracts of their axons that originate in the spinal cord, travel through the brainstem, and ultimately terminate in distinct **relay nuclei** in the thalamus in the diencephalon. The still-higher targets of the thalamic neurons are the cortical areas around the postcentral gyrus that are collectively referred to as the **somatic sensory cortex**. Thus, the somatic sensory system includes specific populations of neurons in practically every subdivision of the nervous system.

Two further principles of neural system organization are evident in the somatic sensory system: **topographic organization** and the prevalence of **parallel pathways** (see Figure 1.13). As the name implies, topography refers to a mapping function—in this case a map of the body surface that can be discerned within the various structures that constitute the somatic sensory



system. Thus, adjacent areas on the body surface are mapped to adjacent regions in nuclei, in white matter tracts, and in the thalamic and cortical targets of the system. Beginning in the periphery, the cells in each dorsal root ganglion define a discrete **dermatome** (the area of the skin innervated by the processes of cells from a single dorsal root). In the spinal cord, from caudal to rostral, the dermatomes are represented in corresponding regions of the spinal cord from sacral (back) to lumbar (legs) to thoracic (chest) and cervical (arms and shoulders) (see Figures 1.13 and 1.11C). This so-called **somatotopy** is maintained in the somatic sensory tracts in spinal cord and brainstem that convey information to the relevant forebrain structures of the somatic sensory system (Figure 1.14).

Parallel pathways refer to the segregation of nerve cell axons that process the distinct stimulus attributes that comprise a particular sensory, motor, or cognitive modality. For somatic sensation, the stimulus attributes relayed via parallel pathways are pain, temperature, touch, pressure, and proprioception (the sense of joint or limb position). From the dorsal root ganglia, through

Figure 1.13 The anatomical and functional organization of the somatic sensory system. Central nervous system components of the somatic sensory system are found in the spinal cord, brainstem, thalamus, and cerebral cortex. (A) Somatosensory information from the body surface is mapped onto dorsal root ganglia (DRG), schematically depicted here as attachments to the spinal cord. The various shades of purple indicate correspondence between regions of the body and the DRG that relay information from the body surface to the central nervous system. Information from the head and neck is relayed to the CNS via the trigeminal ganglia. (B) Somatosensory information travels from the peripheral sensory receptors via parallel pathways for mechanical sensation and for the sensation of pain and temperature. These parallel pathways relay through the spinal cord and brainstem, ultimately sending sensory information to the thalamus, from which it is relayed to the somatic sensory cortex in the postcentral gyrus (indicated in blue in the image of the whole brain; MRI courtesy of L. E. White, J. Vovoydic, and S. M. Williams).

the spinal cord and brainstem, and on to the somatic sensory cortex, these submodalities are kept largely segregated. Thus anatomically, biochemically, and physiologically distinct neurons transduce, encode, and relay pain, temperature, and mechanical information. Although this information is subsequently integrated to provide unitary perception of the relevant stimuli, neurons and circuits in the somatic sensory system are clearly specialized to process discrete aspects of somatic sensation.

This basic outline of the organization of the somatic system is representative of the principles pertinent to understanding any neural system. It will in every case be pertinent to consider the anatomical distribution of neural circuits dedicated to a particular function, how the function is represented or “mapped” onto the neural elements within the system, and how distinct stimulus attributes are segregated within subsets of neurons that comprise the system. Such details provide a framework for understanding how activity within the system provides a representation of relevant stimulus, the required motor response, and higher order cognitive correlates.

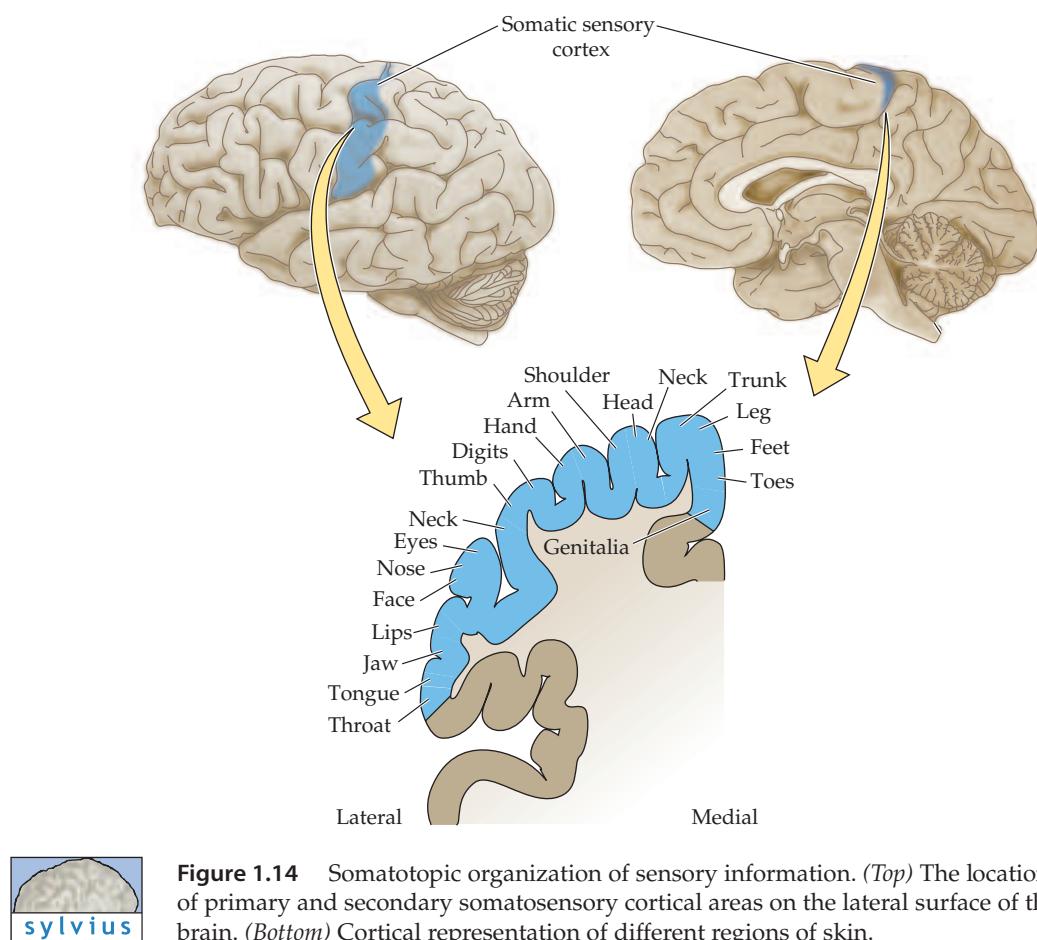


Figure 1.14 Somatotopic organization of sensory information. (Top) The locations of primary and secondary somatosensory cortical areas on the lateral surface of the brain. (Bottom) Cortical representation of different regions of skin.

Functional Analysis of Neural Systems

A wide range of physiological methods is now available to evaluate the electrical (and metabolic) activity of the neuronal circuits that make up a neural system. Two approaches, however, have been particularly useful in defining how neural systems represent information. The most widely used method is **single-cell**, or **single-unit electrophysiological recording** with microelectrodes (see above; this method often records from several nearby cells in addition to the one selected, providing further useful information). The use of microelectrodes to record action potential activity provides a cell-by-cell analysis of the organization topographic maps (Figure 1.15), and can give specific insight into the type of stimulus to which the neuron is “tuned” (i.e., the stimulus that elicits a maximal change in action potential activity from the baseline state). Single-unit analysis is often used to define a neuron’s **receptive field**—the region in sensory space (e.g., the body surface, or a specialized structure such as the retina) within which a specific stimulus elicits the greatest action potential response. This approach to understanding neural systems was introduced by Stephen Kuffler and Vernon Mountcastle in the early 1950s and has now been used by several generations of neuroscientists to evaluate the relationship between stimuli and neuronal responses in both sensory and motor systems. Electrical recording techniques

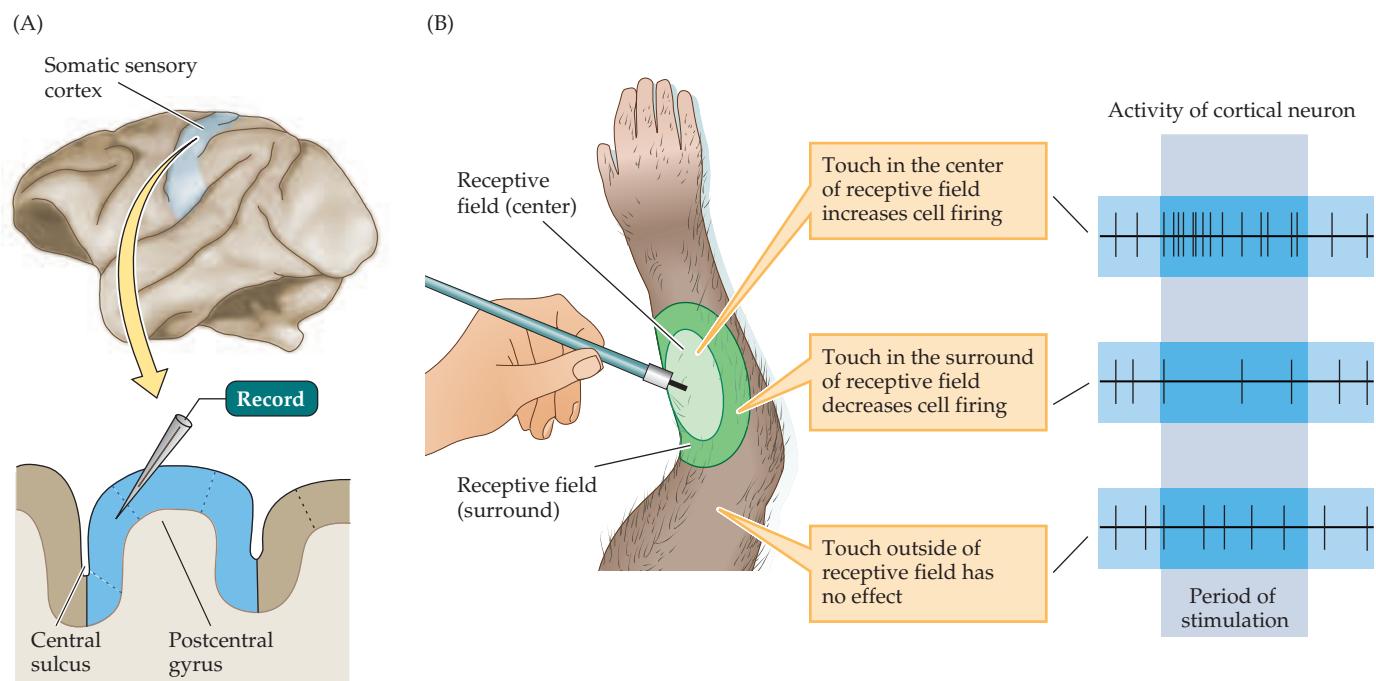


Figure 1.15 Single-unit electrophysiological recording from cortical pyramidal neuron, showing the firing pattern in response to a specific peripheral stimulus. (A) Typical experimental set-up. (B) Defining neuronal receptive fields.

at the single-cell level have now been extended and refined to include single and simultaneous multiple cell analysis in animals performing complex cognitive tasks, intracellular recordings in intact animals, and the use of patch electrodes to detect and monitor the activity of the individual membrane molecules that ultimately underlie neural signaling (see Unit I).

The second major area in which remarkable technical advances have been made is **functional brain imaging** in human subjects (and to a lesser extent animals), which has revolutionized the functional understanding of neural systems over the last two decades (Box A). Unlike electrical methods of recording neural activity, which are invasive in the sense of having to expose the brain and insert electrodes into it, functional imaging is noninvasive and thus applicable to both patients and normal human subjects. Moreover, functional imaging allows the simultaneous evaluation of multiple brain structures (which is possible but obviously difficult with electrical recording methods). The tasks that can be evaluated with functional imaging permit a far more ambitious and integrative approach to studying the operations of a neural system.

Over the last 20 years, these noninvasive methods have allowed neuroscientists to evaluate the representation of an enormous number of complex human behaviors, and at the same time have provided diagnostic tools that are used more and more routinely. Many of the resulting observations have confirmed inferences about functional localization and the organization of neural systems that were originally based on the study of neurological patients who exhibited altered behavior after stroke or other forms of brain injury. Others findings, however, have given new insights into the way neural systems function in the human brain.

Analyzing Complex Behavior

Many of the most widely heralded advances in modern neuroscience have involved reducing the complexity of the brain to more readily analyzed components—i.e., genes, molecules, or cells. Nevertheless, the brain functions as a whole, and the study of more complex (and, some might argue, more interesting) brain functions such as perception, language, emotion, memory, and consciousness remain a central challenge for contemporary neuroscientists. In recognition of this challenge, over the last 20 years or so a field called **cognitive neuroscience** has emerged that is specifically devoted to understanding these issues (see Unit V). This evolution has also rejuvenated the field of neuroethology (which is devoted to observing complex behaviors of animals in their native environments—for example, social communication in birds and non-human primates), and has encouraged the development of tasks to better evaluate the genesis of complex behaviors in human subjects. When used in combination with functional imaging, well designed behavioral tasks can facilitate identification of brain networks devoted to specific complex functions, including language skills, mathematical and musical ability, emotional responses, aesthetic judgments, and abstract thinking. Carefully constructed behavioral tasks can also be used to study the pathology of complex brain diseases that compromise cognition, such Alzheimer's disease, schizophrenia, and depression.

In short, new or revitalized efforts to study higher brain functions with increasingly powerful techniques offer ways of beginning to understand even the most complex aspects of human behavior.

Box A

Brain Imaging Techniques

In the 1970s, **computerized tomography**, or **CT**, opened a new era in noninvasive imaging by introducing the use of computer processing technology to help probe the living brain. Prior to CT, the only brain imaging technique available was standard X-ray film, which has poor soft tissue contrast and involves relatively high radiation exposure.

The CT approach uses a narrow X-ray beam and a row of very sensitive detectors placed on opposite sides of the head to probe just a small portion of tissue at a time with limited radiation exposure (see Figure A). In order to make an image, the X-ray tube and detectors rotate around the head to collect radiodensity information from every orientation around a narrow slice. Computer processing techniques then calculate the radiodensity of each point within the slice plane, producing a tomographic image (*tomo* means “cut” or “slice”). If the patient is slowly moved through the scanner while the X-ray tube rotates in this way, a three-

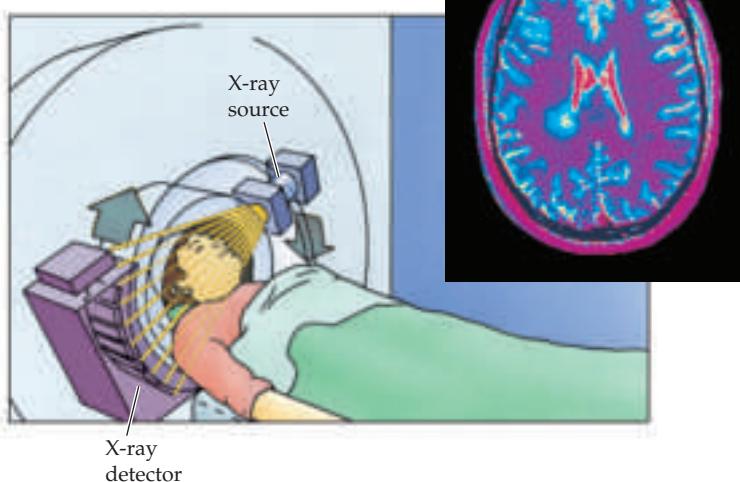
dimensional radiodensity matrix can be created, allowing images to be computed for any plane through the brain. CT scans can readily distinguish gray matter and white matter, differentiate the ventricles quite well, and show many other brain structures with a spatial resolution of several millimeters.

Brain imaging took another large step forward in the 1980s with the development of **magnetic resonance imaging** (**MRI**). MRI is based on the fact that the nuclei of some atoms act as spinning magnets, and that if they are placed in a strong magnetic field they will line up with the field and spin at a frequency that is dependent on the field strength. If they then receive a brief radiofrequency pulse tuned to their spinning frequency they are knocked out of alignment with the field, and subsequently emit energy in an oscillatory fashion as they gradually realign themselves with the field. The strength of the emitted signal depends on how many nuclei are involved in this

process. To get spatial information in MRI, the magnetic field is distorted slightly by imposing magnetic gradients along three different spatial axes so that only nuclei at certain locations are tuned to the detector’s frequency at any given time. Almost all MRI scanners use detectors tuned to the radio frequencies of spinning hydrogen nuclei in water molecules, and thus create images based on the distribution of water in different tissues. Careful manipulation of magnetic field gradients and radiofrequency pulses make it possible to construct extraordinarily detailed images of the brain at any location and orientation with sub-millimeter resolution.

The strong magnetic field and radiofrequency pulses used in MRI scanning are harmless, making this technique completely noninvasive (although metal objects in or near a scanner are a safety concern) (see Figure B). MRI is also extremely versatile because, by changing the scanning parameters, images based on a wide variety of different contrast mechanisms can be generated. For example, conventional MR images take advantage of the fact that hydrogen in different types of tissue (e.g., gray matter, white matter, cerebrospinal fluid) have slightly different realignment rates, meaning that soft tissue contrast can be manipulated simply by adjusting when the realigning hydrogen signal is measured. Different parameter settings can also be used to generate images in which gray and white matter are invisible but in which the brain vasculature stands out in sharp detail. Safety and versatility have made MRI the technique of choice for imaging brain structure in most applications.

Imaging functional variations in the living brain has also become possible with the recent development of techniques for detecting small, localized



(A) In computerized tomography, the X-ray source and detectors are moved around the patient’s head. The inset shows a horizontal CT section of a normal adult brain.

(continued)

Box A (continued)

Brain Imaging Techniques

changes in metabolism or cerebral blood flow. To conserve energy, the brain regulates its blood flow such that active neurons with relatively high metabolic demands receive more blood than relatively inactive neurons. Detecting and mapping these local changes in cerebral blood flow forms the basis for three widely used functional brain imaging techniques: **positron emission tomography (PET)**, **single-photon emission computerized tomography (SPECT)**, and **functional magnetic resonance imaging (fMRI)**.

In PET scanning, unstable positron-emitting isotopes are incorporated into different reagents (including water, precursor molecules of specific neurotransmitters, or glucose) and injected into the bloodstream. Labeled oxygen and glucose quickly accumulate in more metabolically active areas, and labeled transmitter probes are taken up selectively by appropriate regions. As the unstable isotope decays, it results in the emission of two positrons moving in opposite directions. Gamma ray detectors placed around the head register a “hit” only when two detectors 180° apart react simultaneously. Images of tissue isotope density can then be generated (much the way CT images are calculated) showing the location of active regions with a spa-



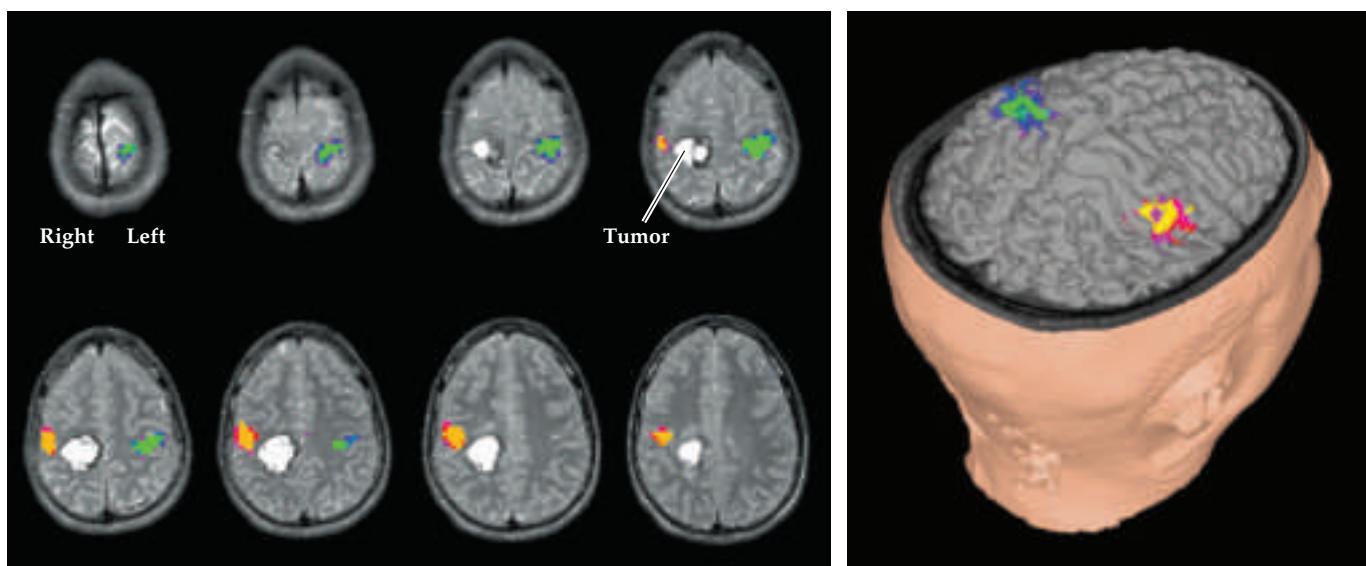
(B) In MRI scanning, the head is placed in the center of a large magnet. A radiofrequency antenna coil is placed around the head for exciting and recording the magnetic resonance signal. For fMRI, stimuli can be presented using virtual reality video goggles and stereo headphones while inside the scanner.

tial resolution of about 4 mm. Depending on the probe injected, PET imaging can be used to visualize activity-dependent changes in blood flow, tissue metabolism, or biochemical activity. SPECT imaging is similar to PET in that it involves injection or inhalation of a radiolabeled compound (for example, ^{133}Xe or ^{123}I -labeled iodoamphetamine), which produce photons that are detected by a gamma camera moving rapidly around the head.

Functional MRI, a variant of MRI, currently offers the best approach for visualizing brain function based on local metabolism. fMRI is predicated on the fact that hemoglobin in blood slightly distorts the magnetic resonance properties of hydrogen nuclei in its vicinity, and

Summary

The brain can be studied by methods that range from genetics and molecular biology to behavioral testing of normal human subjects. In addition to an ever-increasing store of knowledge about the anatomical organization of the nervous system, many of the brightest successes of modern neuroscience have come from understanding nerve cells as the basic structural and functional unit of the nervous system. Studies of the distinct cellular architecture and molecular components of neurons and glia have revealed much about



(C) MRI images of an adult patient with a brain tumor, with fMRI activity during a hand motion task superimposed (left hand activity is shown in yellow, right hand activity in green). At right is a three-dimensional surface reconstructed view of the same data.

the amount of magnetic distortion changes depending on whether the hemoglobin has oxygen bound to it. When a brain area is activated by a specific task it begins to use more oxygen and within seconds the brain microvasculature responds by increasing the flow of oxygen-rich blood to the active area. These changes in the concentration of oxygen and blood flow lead to localized blood oxygenation level-dependent (BOLD) changes in the magnetic resonance signal. Such fluctuations are detected using statistical image process-

ing techniques to produce maps of task-dependent brain function (see Figure C). Because fMRI uses signals intrinsic to the brain without any radioactivity, repeated observations can be made on the same individual—a major advantage over imaging methods such as PET. The spatial resolution (2–3 mm) and temporal resolution (a few seconds) of fMRI are also superior to other functional imaging techniques. MRI has thus emerged as the technology of choice for probing both the structure and function of the living human brain.

References

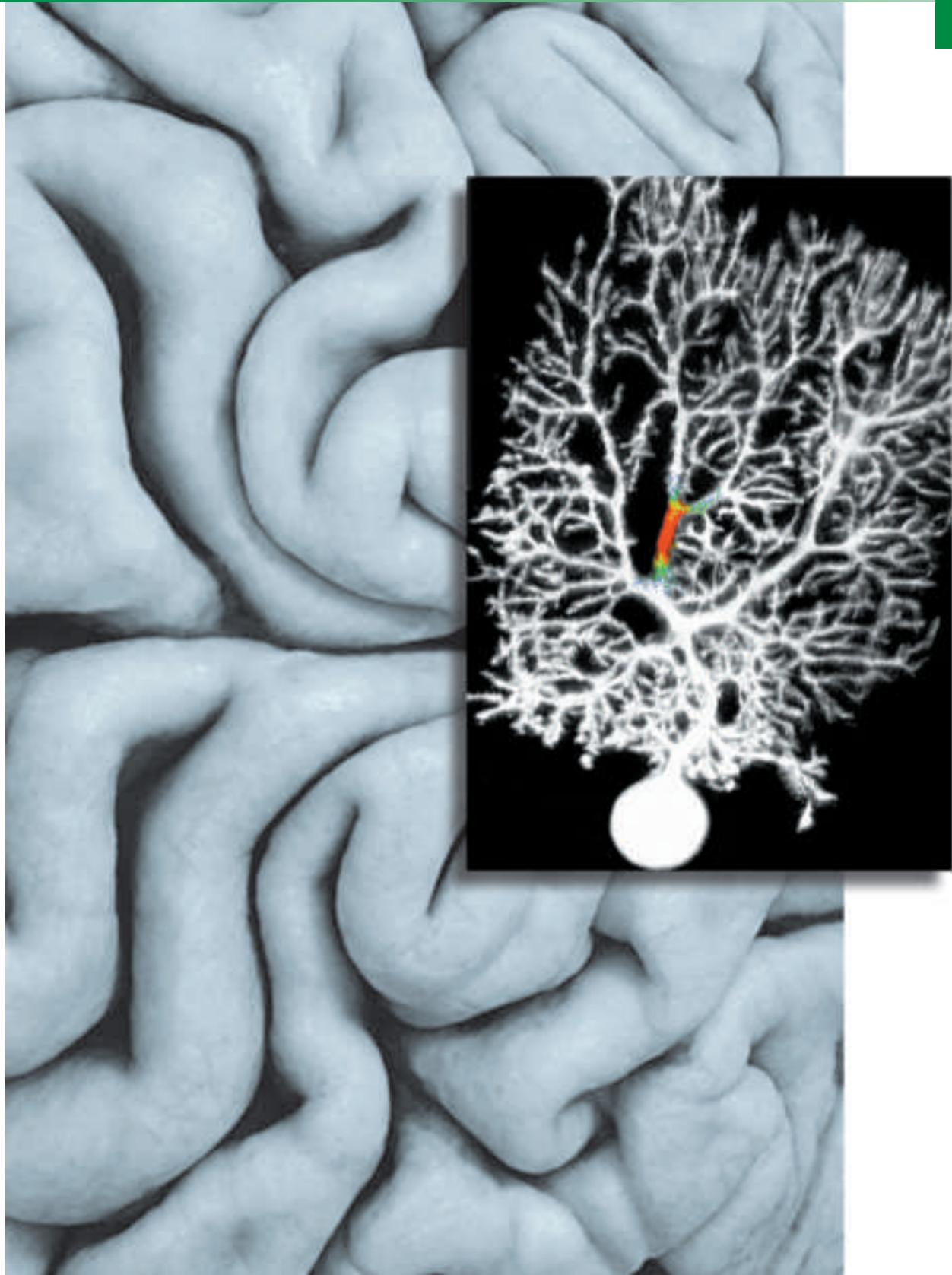
- HUETTEL, S. A., A. W. SONG AND G. McCARTHY (2004) *Functional Magnetic Resonance Imaging*. Sunderland, MA: Sinauer Associates.
- OLDENDORF, W. AND W. OLDENDORF JR. (1988) *Basics of Magnetic Resonance Imaging*. Boston: Kluwer Academic Publishers.
- RAICHLÉ, M. E. (1994) Images of the mind: Studies with modern imaging techniques. *Ann. Rev. Psychol.* 45: 333–356.
- SCHILD, H. (1990) *MRI Made Easy (...Well, Almost)*. Berlin: H. Heineman.

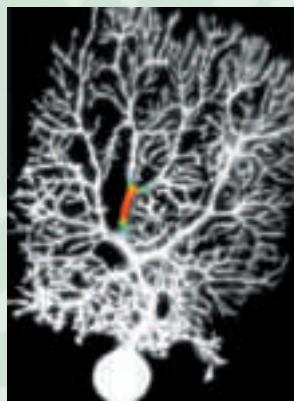
their individual functions, as well as providing a basis for understanding how nerve cells are organized into circuits, and circuits into systems that process specific types of information pertinent to perception and action. Goals that remain include understanding how basic molecular genetic phenomena are linked to cellular, circuit, and system functions; understanding how these processes go awry in neurological and psychiatric diseases; and beginning to understand the especially complex functions of the brain that make us human.

Additional Reading

- BRODAL, P. (1992) *The Central Nervous System: Structure and Function*. New York: Oxford University Press.
- CARPENTER, M. B. AND J. SUTIN (1983) *Human Neuroanatomy*, 8th Ed. Baltimore, MD: Williams and Wilkins.
- ENGLAND, M. A. AND J. WAKELY (1991) *Color Atlas of the Brain and Spinal Cord: An Introduction to Normal Neuroanatomy*. St. Louis: Mosby Yearbook.
- GIBSON, G. AND S. MUSE (2001) *A Primer of Genome Science*. Sunderland, MA: Sinauer Associates.
- HAINES, D. E. (1995) *Neuroanatomy: An Atlas of Structures, Sections, and Systems*, 2nd Ed. Baltimore: Urban and Schwarzenberg.
- MARTIN, J. H. (1996) *Neuroanatomy: Text and Atlas*, 2nd Ed. Stamford, CT: Appleton and Lange.
- NATURE VOL. 409, NO. 6822 (2001) Issue of February 16. Special issue on the human genome.
- NETTER, F. H. (1983) *The CIBA Collection of Medical Illustrations*, Vols. I and II. A. Brass and R. V. Dingle (eds.). Summit, NJ: CIBA Pharmaceutical Co.
- PETERS, A., S. L. PALAY AND H. DE F. WEBSTER (1991) *The Fine Structure of the Nervous System: Neurons and Their Supporting Cells*, 3rd Ed. New York: Oxford University Press.
- POSNER, M. I. AND M. E. RAICHLÉ (1997) *Images of Mind*, 2nd Ed. New York: W. H. Freeman & Co.
- RAMÓN Y CAJAL, S. (1984) *The Neuron and the Glial Cell*. (Transl. by J. de la Torre and W. C. Gibson.) Springfield, IL: Charles C. Thomas.
- RAMÓN Y CAJAL, S. (1990) *New Ideas on the Structure of the Nervous System in Man and Vertebrates*. (Transl. by N. Swanson and L. W. Swanson.) Cambridge, MA: MIT Press.
- SCIENCE VOL. 291, NO. 5507 (2001) Issue of February 16. Special issue on the human genome.
- SHEPHERD, G. M. (1991) *Foundations of the Neuron Doctrine*. History of Neuroscience Series, No. 6. Oxford: Oxford University Press.
- WAXMAN, S. G. AND J. DEGROOT (1995) *Correlative Neuroanatomy*, 22nd Ed. Norwalk, CT: Appleton and Lange.

Neural Signaling





Calcium signaling in a cerebellar Purkinje neuron. An electrode was used to fill the neuron with a fluorescent calcium indicator dye. This dye revealed the release of intracellular calcium ions (color) produced by the actions of the second messenger IP₃. (Courtesy of Elizabeth A. Finch and George J. Augustine.)

UNIT I

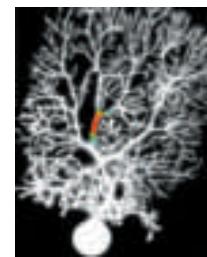
NEURAL SIGNALING

- 2 *Electrical Signals of Nerve Cells*
- 3 *Voltage-Dependent Membrane Permeability*
- 4 *Channels and Transporters*
- 5 *Synaptic Transmission*
- 6 *Neurotransmitters, Receptors, and Their Effects*
- 7 *Molecular Signaling within Neurons*

The brain is remarkably adept at acquiring, coordinating, and disseminating information about the body and its environment. Such information must be processed within milliseconds, yet it also can be stored away as memories that endure for years. Neurons within the central and peripheral nervous systems perform these functions by generating sophisticated electrical and chemical signals. This unit describes these signals and how they are produced. It explains how one type of electrical signal, the action potential, allows information to travel along the length of a nerve cell. It also explains how other types of signals—both electrical and chemical—are generated at synaptic connections between nerve cells. Synapses permit information transfer by interconnecting neurons to form the circuitry on which neural processing depends. Finally, it describes the intricate biochemical signaling events that take place within neurons. Appreciating these fundamental forms of neuronal signaling provides a foundation for appreciating the higher-level functions considered in the rest of the book.

The cellular and molecular mechanisms that give neurons their unique signaling abilities are also targets for disease processes that compromise the function of the nervous system. A working knowledge of the cellular and molecular biology of neurons is therefore fundamental to understanding a variety of brain pathologies, and for developing novel approaches to diagnosing and treating these all too prevalent problems.

Chapter 2



Electrical Signals of Nerve Cells

Overview

Nerve cells generate electrical signals that transmit information. Although neurons are not intrinsically good conductors of electricity, they have evolved elaborate mechanisms for generating these signals based on the flow of ions across their plasma membranes. Ordinarily, neurons generate a negative potential, called the resting membrane potential, that can be measured by recording the voltage between the inside and outside of nerve cells. The action potential transiently abolishes the negative resting potential and makes the transmembrane potential positive. Action potentials are propagated along the length of axons and are the fundamental signal that carries information from one place to another in the nervous system. Still other types of electrical signals are produced by the activation of synaptic contacts between neurons or by the actions of external forms of energy on sensory neurons. All of these electrical signals arise from ion fluxes brought about by nerve cell membranes being selectively permeable to different ions, and from the non-uniform distribution of these ions across the membrane.

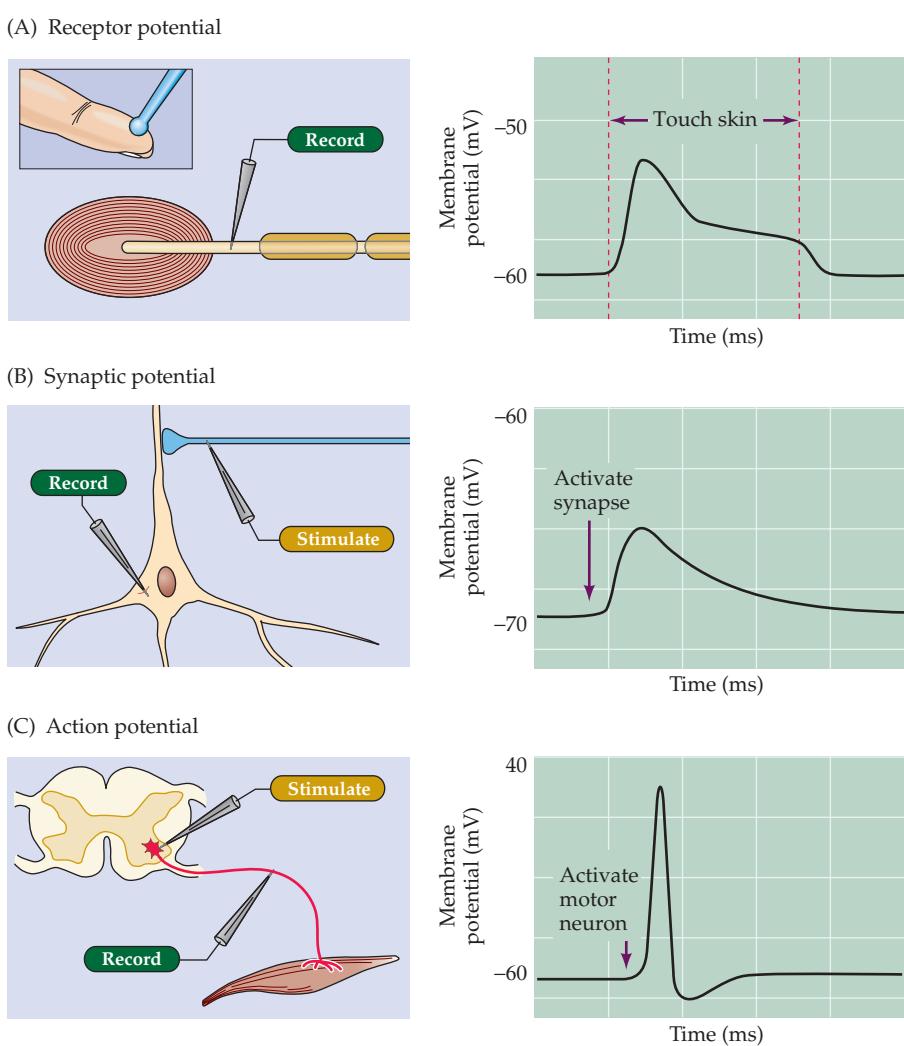
Electrical Potentials across Nerve Cell Membranes

Neurons employ several different types of electrical signal to encode and transfer information. The best way to observe these signals is to use an intracellular microelectrode to measure the electrical potential across the neuronal plasma membrane. A typical microelectrode is a piece of glass tubing pulled to a very fine point (with an opening of less than 1 μm diameter) and filled with a good electrical conductor, such as a concentrated salt solution. This conductive core can then be connected to a voltmeter, such as an oscilloscope, to record the transmembrane voltage of the nerve cell.

The first type of electrical phenomenon can be observed as soon as a microelectrode is inserted through the membrane of the neuron. Upon entering the cell, the microelectrode reports a negative potential, indicating that neurons have a means of generating a constant voltage across their membranes when at rest. This voltage, called the **resting membrane potential**, depends on the type of neuron being examined, but it is always a fraction of a volt (typically -40 to -90 mV).

The electrical signals produced by neurons are caused by responses to stimuli, which then change the resting membrane potential. **Receptor potentials** are due to the activation of sensory neurons by external stimuli, such as light, sound, or heat. For example, touching the skin activates Pacinian corpuscles, receptor neurons that sense mechanical disturbances of the skin. These neurons respond to touch with a receptor potential that changes the resting potential for a fraction of a second (Figure 2.1A). These transient

Figure 2.1 Types of neuronal electrical signals. In all cases, microelectrodes are used to measure changes in the resting membrane potential during the indicated signals. (A) A brief touch causes a receptor potential in a Pacinian corpuscle in the skin. (B) Activation of a synaptic contact onto a hippocampal pyramidal neuron elicits a synaptic potential. (C) Stimulation of a spinal reflex produces an action potential in a spinal motor neuron.



changes in potential are the first step in generating the sensation of vibrations (or “tickles”) of the skin in the somatic sensory system (Chapter 8). Similar sorts of receptor potentials are observed in all other sensory neurons during transduction of sensory signals (Unit II).

Another type of electrical signal is associated with communication between neurons at synaptic contacts. Activation of these synapses generates **synaptic potentials**, which allow transmission of information from one neuron to another. An example of such a signal is shown in Figure 2.1B. In this case, activation of a synaptic terminal innervating a hippocampal pyramidal neuron causes a very brief change in the resting membrane potential in the pyramidal neuron. Synaptic potentials serve as the means of exchanging information in complex neural circuits in both the central and peripheral nervous systems (Chapter 5).

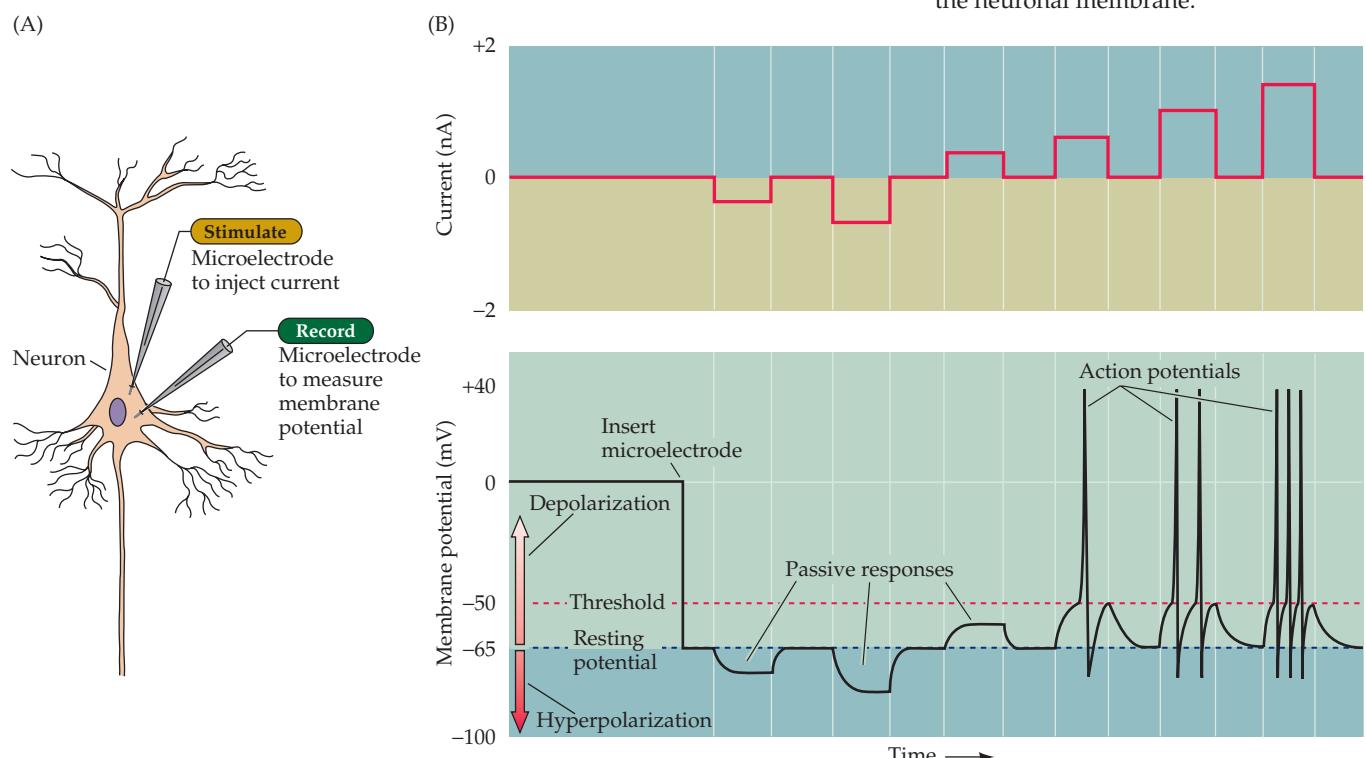
The use of electrical signals—as in sending electricity over wires to provide power or information—presents a series of problems in electrical engineering. A fundamental problem for neurons is that their axons, which can be quite long (remember that a spinal motor neuron can extend for a meter or more), are not good electrical conductors. Although neurons and wires

are both capable of passively conducting electricity, the electrical properties of neurons compare poorly to an ordinary wire. To compensate for this deficiency, neurons have evolved a “booster system” that allows them to conduct electrical signals over great distances despite their intrinsically poor electrical characteristics. The electrical signals produced by this booster system are called **action potentials** (which are also referred to as “spikes” or “impulses”). An example of an action potential recorded from the axon of a spinal motor neuron is shown in Figure 2.1C.

One way to elicit an action potential is to pass electrical current across the membrane of the neuron. In normal circumstances, this current would be generated by receptor potentials or by synaptic potentials. In the laboratory, however, electrical current suitable for initiating an action potential can be readily produced by inserting a second microelectrode into the same neuron and then connecting the electrode to a battery (Figure 2.2A). If the current delivered in this way makes the membrane potential more negative (**hyperpolarization**), nothing very dramatic happens. The membrane potential simply changes in proportion to the magnitude of the injected current (central part of Figure 2.2B). Such hyperpolarizing responses do not require any unique property of neurons and are therefore called passive electrical responses. A much more interesting phenomenon is seen if current of the opposite polarity is delivered, so that the membrane potential of the nerve cell becomes more positive than the resting potential (**depolarization**). In this case, at a certain level of membrane potential, called the **threshold potential**, an action potential occurs (see right side of Figure 2.2B).

The action potential, which is an active response generated by the neuron, is a brief (about 1 ms) change from negative to positive in the transmem-

Figure 2.2 Recording passive and active electrical signals in a nerve cell. (A) Two microelectrodes are inserted into a neuron; one of these measures membrane potential while the other injects current into the neuron. (B) Inserting the voltage-measuring microelectrode into the neuron reveals a negative potential, the resting membrane potential. Injecting current through the current-passing microelectrode alters the neuronal membrane potential. Hyperpolarizing current pulses produce only passive changes in the membrane potential. While small depolarizing currents also elicit only passive responses, depolarizations that cause the membrane potential to meet or exceed threshold additionally evoke action potentials. Action potentials are active responses in the sense that they are generated by changes in the permeability of the neuronal membrane.



brane potential. Importantly, the amplitude of the action potential is independent of the magnitude of the current used to evoke it; that is, larger currents do not elicit larger action potentials. The action potentials of a given neuron are therefore said to be *all-or-none*, because they occur fully or not at all. If the amplitude or duration of the stimulus current is increased sufficiently, multiple action potentials occur, as can be seen in the responses to the three different current intensities shown in Figure 2.2B (right side). It follows, therefore, that the intensity of a stimulus is encoded in the frequency of action potentials rather than in their amplitude. This arrangement differs dramatically from receptor potentials, whose amplitudes are graded in proportion to the magnitude of the sensory stimulus, or synaptic potentials, whose amplitude varies according to the number of synapses activated and the previous amount of synaptic activity.

Because electrical signals are the basis of information transfer in the nervous system, it is essential to understand how these signals arise. Remarkably, all of the neuronal electrical signals described above are produced by similar mechanisms that rely upon the movement of ions across the neuronal membrane. The remainder of this chapter addresses the question of how nerve cells use ions to generate electrical potentials. Chapter 3 explores more specifically the means by which action potentials are produced and how these signals solve the problem of long-distance electrical conduction within nerve cells. Chapter 4 examines the properties of membrane molecules responsible for electrical signaling. Finally, Chapters 5–7 consider how electrical signals are transmitted from one nerve cell to another at synaptic contacts.

How Ionic Movements Produce Electrical Signals

Electrical potentials are generated across the membranes of neurons—and, indeed, all cells—because (1) there are *differences in the concentrations* of specific ions across nerve cell membranes, and (2) the membranes are *selectively permeable* to some of these ions. These two facts depend in turn on two different kinds of proteins in the cell membrane (Figure 2.3). The ion concentration gradients are established by proteins known as **active transporters**, which, as their name suggests, actively move ions into or out of cells against their concentration gradients. The selective permeability of membranes is

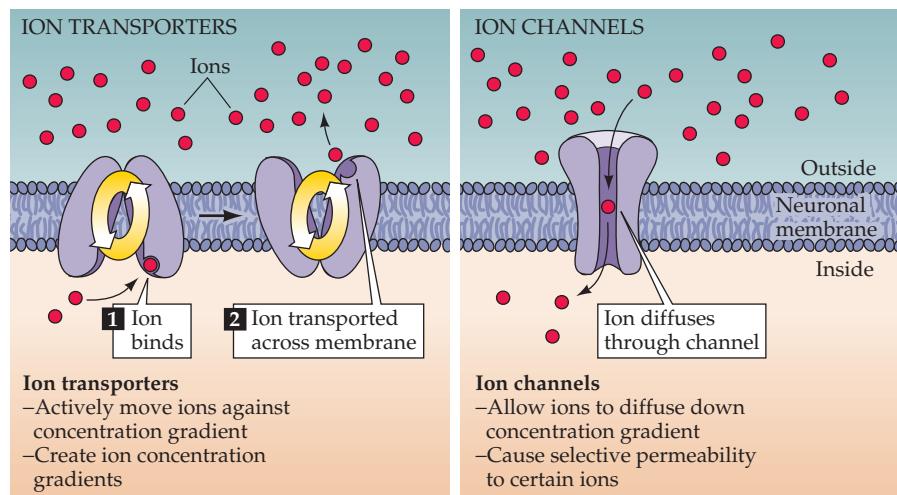


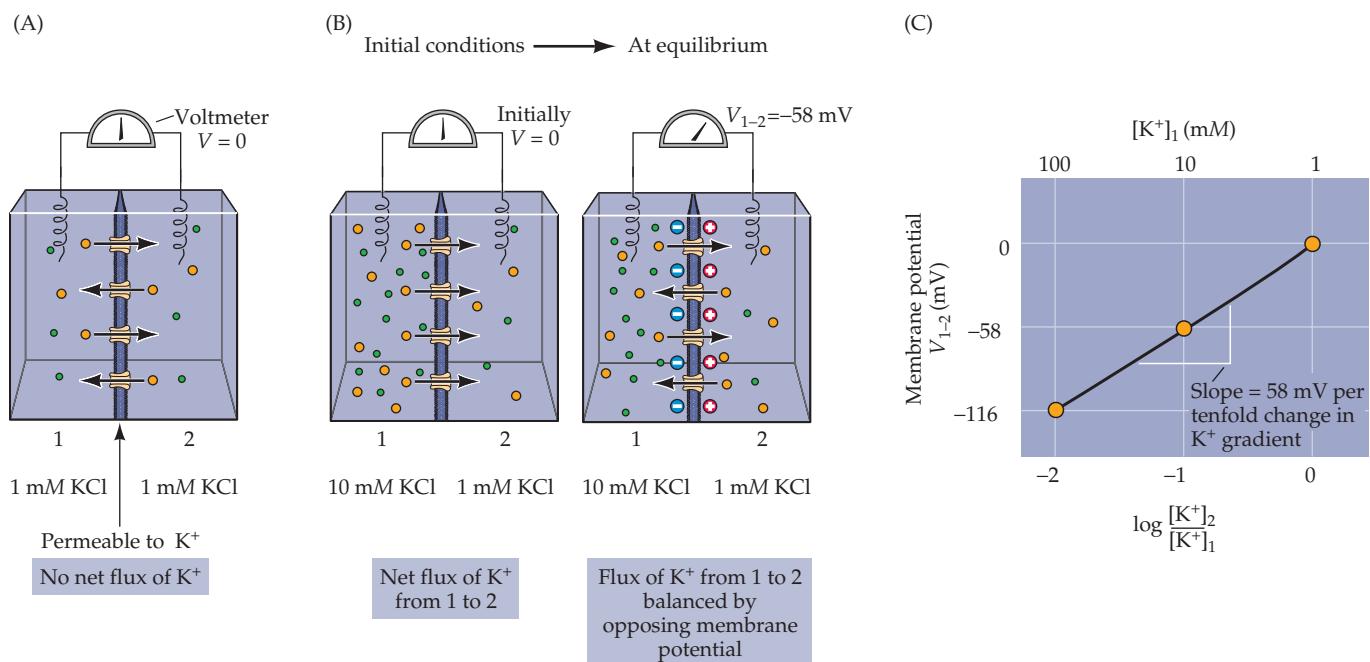
Figure 2.3 Ion transporters and ion channels are responsible for ionic movements across neuronal membranes. Transporters create ion concentration differences by actively transporting ions against their chemical gradients. Channels take advantage of these concentration gradients, allowing selected ions to move, via diffusion, down their chemical gradients.

due largely to **ion channels**, proteins that allow only certain kinds of ions to cross the membrane in the direction of their concentration gradients. Thus, channels and transporters basically work against each other, and in so doing they generate the resting membrane potential, action potentials, and the synaptic potentials and receptor potentials that trigger action potentials. The structure and function of these channels and transporters are described in Chapter 4.

To appreciate the role of ion gradients and selective permeability in generating a membrane potential, consider a simple system in which an artificial membrane separates two compartments containing solutions of ions. In such a system, it is possible to determine the composition of the two solutions and, thereby, control the ion gradients across the membrane. For example, take the case of a membrane that is permeable only to potassium ions (K^+). If the concentration of K^+ on each side of this membrane is equal, then no electrical potential will be measured across it (Figure 2.4A). However, if the concentration of K^+ is not the same on the two sides, then an electrical potential will be generated. For instance, if the concentration of K^+ on one side of the membrane (compartment 1) is 10 times higher than the K^+ concentration on the other side (compartment 2), then the electrical potential of compartment 1 will be negative relative to compartment 2 (Figure 2.4B). This difference in electrical potential is generated because the potassium ions flow down their concentration gradient and take their electrical charge (one positive charge per ion) with them as they go. Because neuronal membranes contain pumps that accumulate K^+ in the cell cytoplasm, and because potassium-permeable channels in the plasma membrane allow a transmembrane flow of K^+ , an analogous situation exists in living nerve cells. A continual resting efflux of K^+ is therefore responsible for the resting membrane potential.

In the hypothetical case just described, an equilibrium will quickly be reached. As K^+ moves from compartment 1 to compartment 2 (the initial conditions on the left of Figure 2.4B), a potential is generated that tends to impede further flow of K^+ . This impediment results from the fact that the

Figure 2.4 Electrochemical equilibrium. (A) A membrane permeable only to K^+ (yellow spheres) separates compartments 1 and 2, which contain the indicated concentrations of KCl. (B) Increasing the KCl concentration in compartment 1 to 10 mM initially causes a small movement of K^+ into compartment 2 (initial conditions) until the electromotive force acting on K^+ balances the concentration gradient, and the net movement of K^+ becomes zero (at equilibrium). (C) The relationship between the transmembrane concentration gradient ($[K^+]_2/[K^+]_1$) and the membrane potential. As predicted by the Nernst equation, this relationship is linear when plotted on semi-logarithmic coordinates, with a slope of 58 mV per tenfold difference in the concentration gradient.



potential gradient across the membrane tends to repel the positive potassium ions that would otherwise move across the membrane. Thus, as compartment 2 becomes positive relative to compartment 1, the increasing positivity makes compartment 2 less attractive to the positively charged K⁺. The net movement (or flux) of K⁺ will stop at the point (at equilibrium on the right of Figure 2.4B) where the potential change across the membrane (the relative positivity of compartment 2) exactly offsets the concentration gradient (the tenfold excess of K⁺ in compartment 1). At this **electrochemical equilibrium**, there is an exact balance between two opposing forces: (1) the concentration gradient that causes K⁺ to move from compartment 1 to compartment 2, taking along positive charge, and (2) an opposing electrical gradient that increasingly tends to stop K⁺ from moving across the membrane (Figure 2.4B). The number of ions that needs to flow to generate this electrical potential is very small (approximately 10⁻¹² moles of K⁺ per cm² of membrane, or 10¹² K⁺ ions). This last fact is significant in two ways. First, it means that the concentrations of permeant ions on each side of the membrane remain essentially constant, even after the flow of ions has generated the potential. Second, the tiny fluxes of ions required to establish the membrane potential do not disrupt chemical electroneutrality because each ion has an oppositely charged counter-ion (chloride ions in the example shown in Figure 2.4) to maintain the neutrality of the solutions on each side of the membrane. The concentration of K⁺ remains equal to the concentration of Cl⁻ in the solutions in compartments 1 and 2, meaning that the separation of charge that creates the potential difference is restricted to the immediate vicinity of the membrane.

The Forces That Create Membrane Potentials

The electrical potential generated across the membrane at electrochemical equilibrium, the **equilibrium potential**, can be predicted by a simple formula called the **Nernst equation**. This relationship is generally expressed as

$$E_X = \frac{RT}{zF} \ln \frac{[X]_2}{[X]_1}$$

where E_X is the equilibrium potential for any ion X, R is the gas constant, T is the absolute temperature (in degrees on the Kelvin scale), z is the valence (electrical charge) of the permeant ion, and F is the Faraday constant (the amount of electrical charge contained in one mole of a univalent ion). The brackets indicate the concentrations of ion X on each side of the membrane and the symbol ln indicates the natural logarithm of the concentration gradient. Because it is easier to perform calculations using base 10 logarithms and to perform experiments at room temperature, this relationship is usually simplified to

$$E_X = \frac{58}{z} \log \frac{[X]_2}{[X]_1}$$

where log indicates the base 10 logarithm of the concentration ratio. Thus, for the example in Figure 2.4B, the potential across the membrane at electrochemical equilibrium is

$$E_K = \frac{58}{z} \log \frac{[K]_2}{[K]_1} = 58 \log \frac{1}{10} = -58 \text{ mV}$$

The equilibrium potential is conventionally defined in terms of the potential difference between the reference compartment, side 2 in Figure 2.4, and the other side. This approach is also applied to biological systems. In this case,

the outside of the cell is the conventional reference point (defined as zero potential). Thus, when the concentration of K^+ is higher inside than out, an inside-negative potential is measured across the K^+ -permeable neuronal membrane.

For a simple hypothetical system with only one permeant ion species, the Nernst equation allows the electrical potential across the membrane at equilibrium to be predicted exactly. For example, if the concentration of K^+ on side 1 is increased to 100 mM, the membrane potential will be -116 mV. More generally, if the membrane potential is plotted against the logarithm of the K^+ concentration gradient ($[K]_2/[K]_1$), the Nernst equation predicts a linear relationship with a slope of 58 mV (actually $58/z$) per tenfold change in the K^+ gradient (Figure 2.4C).

To reinforce and extend the concept of electrochemical equilibrium, consider some additional experiments on the influence of ionic species and ionic permeability that could be performed on the simple model system in Figure 2.4. What would happen to the electrical potential across the membrane (the potential of side 1 relative to side 2) if the potassium on side 2 were replaced with 10 mM sodium (Na^+) and the K^+ in compartment 1 were replaced by 1 mM Na^+ ? No potential would be generated, because no Na^+ could flow across the membrane (which was defined as being permeable only to K^+). However, if under these ionic conditions (10 times more Na^+ in compartment 2) the K^+ -permeable membrane were to be magically replaced by a membrane permeable only to Na^+ , a potential of +58 mV would be measured at equilibrium. If 10 mM calcium (Ca^{2+}) were present in compartment 2 and 1 mM Ca^{2+} in compartment 1, and a Ca^{2+} -selective membrane separated the two sides, what would happen to the membrane potential? A potential of +29 mV would develop, because the valence of calcium is +2. Finally, what would happen to the membrane potential if 10 mM Cl^- were present in compartment 1 and 1 mM Cl^- were present in compartment 2, with the two sides separated by a Cl^- -permeable membrane? Because the valence of this anion is -1, the potential would again be +58 mV.

The balance of chemical and electrical forces at equilibrium means that the electrical potential can determine ionic fluxes across the membrane, just as the ionic gradient can determine the membrane potential. To examine the influence of membrane potential on ionic flux, imagine connecting a battery across the two sides of the membrane to control the electrical potential across the membrane without changing the distribution of ions on the two sides (Figure 2.5). As long as the battery is off, things will be just as in Figure 2.4, with the flow of K^+ from compartment 1 to compartment 2 causing a negative membrane potential (Figure 2.5A, left). However, if the battery is used to make compartment 1 initially more negative relative to compartment 2, there will be less K^+ flux, because the negative potential will tend to keep K^+ in compartment 1. How negative will side 1 need to be before there is no net flux of K^+ ? The answer is -58 mV, the voltage needed to counter the tenfold difference in K^+ concentrations on the two sides of the membrane (Figure 2.5A, center). If compartment 1 is initially made more negative than -58 mV, then K^+ will actually flow from compartment 2 into compartment 1, because the positive ions will be attracted to the more negative potential of compartment 1 (Figure 2.5A, right). This example demonstrates that both the direction and magnitude of ion flux depend on the membrane potential. Thus, in some circumstances the electrical potential can overcome an ionic concentration gradient.

The ability to alter ion flux experimentally by changing either the potential imposed on the membrane (Figure 2.5B) or the transmembrane concen-

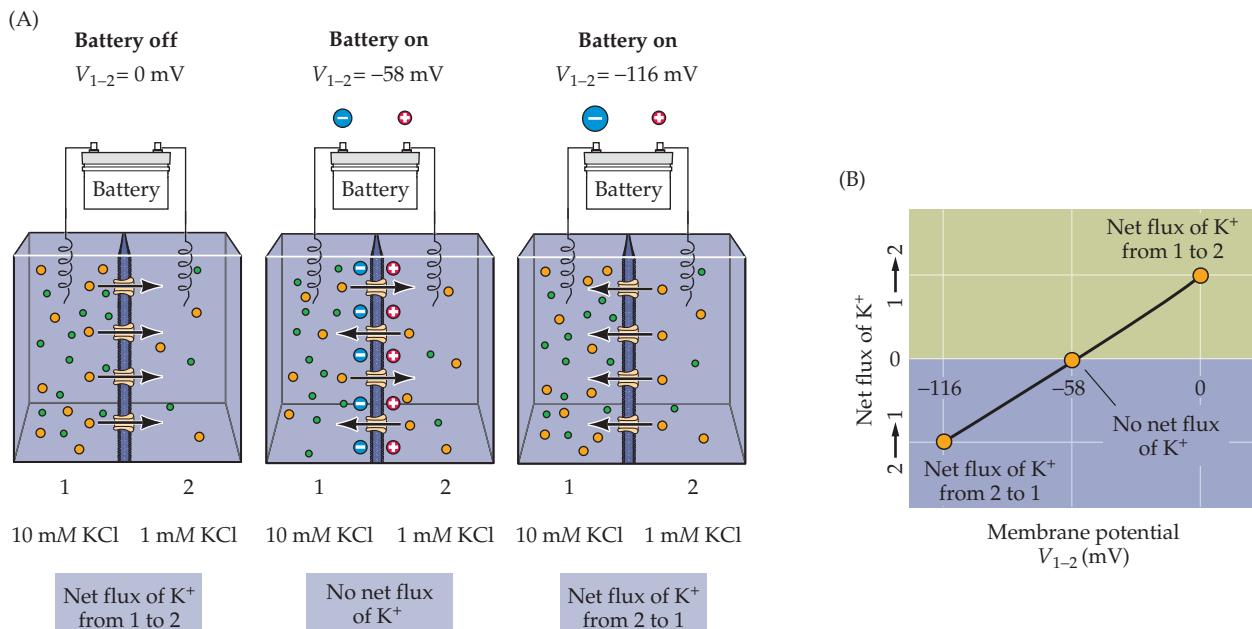


Figure 2.5 Membrane potential influences ion fluxes. (A) Connecting a battery across the K⁺-permeable membrane allows direct control of membrane potential. When the battery is turned off (left), K⁺ ions (yellow) flow simply according to their concentration gradient. Setting the initial membrane potential (V_{1-2}) at the equilibrium potential for K⁺ (center) yields no net flux of K⁺, while making the membrane potential more negative than the K⁺ equilibrium potential (right) causes K⁺ to flow against its concentration gradient. (B) Relationship between membrane potential and direction of K⁺ flux.

concentration gradient for an ion (see Figure 2.4C) provides convenient tools for studying ion fluxes across the plasma membranes of neurons, as will be evident in many of the experiments described in the following chapters.

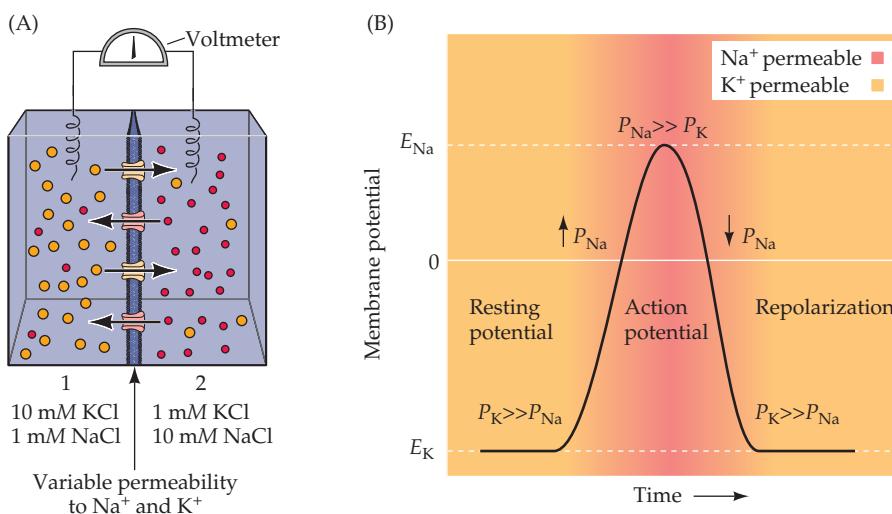
Electrochemical Equilibrium in an Environment with More Than One Permeant Ion

Now consider a somewhat more complex situation in which Na⁺ and K⁺ are unequally distributed across the membrane, as in Figure 2.6A. What would happen if 10 mM K⁺ and 1 mM Na⁺ were present in compartment 1, and 1 mM K⁺ and 10 mM Na⁺ in compartment 2? If the membrane were permeable only to K⁺, the membrane potential would be -58 mV; if the membrane were permeable only to Na⁺, the potential would be +58 mV. But what would the potential be if the membrane were permeable to both K⁺ and Na⁺? In this case, the potential would depend on the relative permeability of the membrane to K⁺ and Na⁺. If it were more permeable to K⁺, the potential would approach -58 mV, and if it were more permeable to Na⁺, the potential would be closer to +58 mV. Because there is no permeability term in the Nernst equation, which only considers the simple case of a single permeant ion species, a more elaborate equation is needed that takes into account both the concentration gradients of the permeant ions and the relative permeability of the membrane to each permeant species.

Such an equation was developed by David Goldman in 1943. For the case most relevant to neurons, in which K⁺, Na⁺, and Cl⁻ are the primary permeant ions, the **Goldman equation** is written

$$V = 58 \log \frac{P_K[K]_2 + P_{Na}[Na]_2 + P_{Cl}[Cl]_1}{P_K[K]_1 + P_{Na}[Na]_1 + P_{Cl}[Cl]_2}$$

where V is the voltage across the membrane (again, compartment 1 relative to the reference compartment 2) and P indicates the permeability of the



membrane to each ion of interest. The Goldman equation is thus an extended version of the Nernst equation that takes into account the relative permeabilities of each of the ions involved. The relationship between the two equations becomes obvious in the situation where the membrane is permeable only to one ion, say, K^+ ; in this case, the Goldman expression collapses back to the simpler Nernst equation. In this context, it is important to note that the valence factor (z) in the Nernst equation has been eliminated; this is why the concentrations of negatively charged chloride ions, Cl^- , have been inverted relative to the concentrations of the positively charged ions [remember that $-\log(A/B) = \log(B/A)$].

If the membrane in Figure 2.6A is permeable to K^+ and Na^+ only, the terms involving Cl^- drop out because $P_{\text{Cl}} = 0$. In this case, solution of the Goldman equation yields a potential of -58 mV when only K^+ is permeant, $+58 \text{ mV}$ when only Na^+ is permeant, and some intermediate value if both ions are permeant. For example, if K^+ and Na^+ were equally permeant, then the potential would be 0 mV .

With respect to neural signaling, it is particularly pertinent to ask what would happen if the membrane started out being permeable to K^+ , and then temporarily switched to become most permeable to Na^+ . In this circumstance, the membrane potential would start out at a negative level, become positive while the Na^+ permeability remained high, and then fall back to a negative level as the Na^+ permeability decreased again. As it turns out, this last case essentially describes what goes on in a neuron during the generation of an action potential. In the resting state, P_{K} of the neuronal plasma membrane is much higher than P_{Na} ; since, as a result of the action of ion transporters, there is always more K^+ inside the cell than outside (Table 2.1), the resting potential is negative (Figure 2.6B). As the membrane potential is depolarized (by synaptic action, for example), P_{Na} increases. The transient increase in Na^+ permeability causes the membrane potential to become even more positive (red region in Figure 2.6B), because Na^+ rushes in (there is much more Na^+ outside a neuron than inside, again as a result of ion pumps). Because of this positive feedback loop, an action potential occurs. The rise in Na^+ permeability during the action potential is transient, however; as the membrane permeability to K^+ is restored, the membrane potential quickly returns to its resting level.

Figure 2.6 Resting and action potentials entail permeabilities to different ions. (A) Hypothetical situation in which a membrane variably permeable to Na^+ (red) and K^+ (yellow) separates two compartments that contain both ions. For simplicity, Cl^- ions are not shown in the diagram. (B) Schematic representation of the membrane ionic permeabilities associated with resting and action potentials. At rest, neuronal membranes are more permeable to K^+ (yellow) than to Na^+ (red); accordingly, the resting membrane potential is negative and approaches the equilibrium potential for K^+ , E_{K} . During an action potential, the membrane becomes very permeable to Na^+ (red); thus the membrane potential becomes positive and approaches the equilibrium potential for Na^+ , E_{Na} . The rise in Na^+ permeability is transient, however, so that the membrane again becomes primarily permeable to K^+ (yellow), causing the potential to return to its negative resting value. Notice that at the equilibrium potential for a given ion, there is no net flux of that ion across the membrane.

TABLE 2.1
Extracellular and Intracellular Ion Concentrations

<i>Ion</i>	<i>Concentration (mM)</i>	
	<i>Intracellular</i>	<i>Extracellular</i>
Squid neuron		
Potassium (K^+)	400	20
Sodium (Na^+)	50	440
Chloride (Cl^-)	40–150	560
Calcium (Ca^{2+})	0.0001	10
Mammalian neuron		
Potassium (K^+)	140	5
Sodium (Na^+)	5–15	145
Chloride (Cl^-)	4–30	110
Calcium (Ca^{2+})	0.0001	1–2

Armed with an appreciation of these simple electrochemical principles, it will be much easier to understand the following, more detailed account of how neurons generate resting and action potentials.

The Ionic Basis of the Resting Membrane Potential

The action of ion transporters creates substantial transmembrane gradients for most ions. Table 2.1 summarizes the ion concentrations measured directly in an exceptionally large nerve cell found in the nervous system of the squid (Box A). Such measurements are the basis for stating that there is much more K^+ inside the neuron than out, and much more Na^+ outside than in. Similar concentration gradients occur in the neurons of most animals, including humans. However, because the ionic strength of mammalian blood is lower than that of sea-dwelling animals such as squid, in mammals the concentrations of each ion are several times lower. These transporter-dependent concentration gradients are, indirectly, the source of the resting neuronal membrane potential and the action potential.

Once the ion concentration gradients across various neuronal membranes are known, the Nernst equation can be used to calculate the equilibrium potential for K^+ and other major ions. Since the resting membrane potential of the squid neuron is approximately -65 mV , K^+ is the ion that is closest to being in electrochemical equilibrium when the cell is at rest. This fact implies that the resting membrane is more permeable to K^+ than to the other ions listed in Table 2.1, and that this permeability is the source of resting potentials.

It is possible to test this guess, as Alan Hodgkin and Bernard Katz did in 1949, by asking what happens to the resting membrane potential if the concentration of K^+ outside the neuron is altered. If the resting membrane were permeable only to K^+ , then the Goldman equation (or even the simpler Nernst equation) predicts that the membrane potential will vary in proportion to the logarithm of the K^+ concentration gradient across the membrane. Assuming that the internal K^+ concentration is unchanged during the experiment, a plot of membrane potential against the logarithm of the external K^+ concentration should yield a straight line with a slope of 58 mV per tenfold change in external K^+ concentration at room temperature (see Figure 2.4C). (The slope becomes about 61 mV at mammalian body temperatures.)

Box A

The Remarkable Giant Nerve Cells of Squid

Many of the initial insights into how ion concentration gradients and changes in membrane permeability produce electrical signals came from experiments performed on the extraordinarily large nerve cells of the squid. The axons of these nerve cells can be up to 1 mm in diameter—100 to 1000 times larger than mammalian axons. Thus, squid axons are large enough to allow experiments that would be impossible on most other nerve cells. For example, it is not difficult to insert simple wire electrodes inside these giant axons and make reliable electrical measurements. The relative ease of this approach yielded the first intracellular recordings of action potentials from nerve cells and, as discussed in the next chapter, the first experimental measure-

ments of the ion currents that produce action potentials. It also is practical to extrude the cytoplasm from giant axons and measure its ionic composition (see Table 2.1). In addition, some giant nerve cells form synaptic contacts with other giant nerve cells, producing very large synapses that have been extraordinarily valuable in understanding the fundamental mechanisms of synaptic transmission (see Chapter 5).

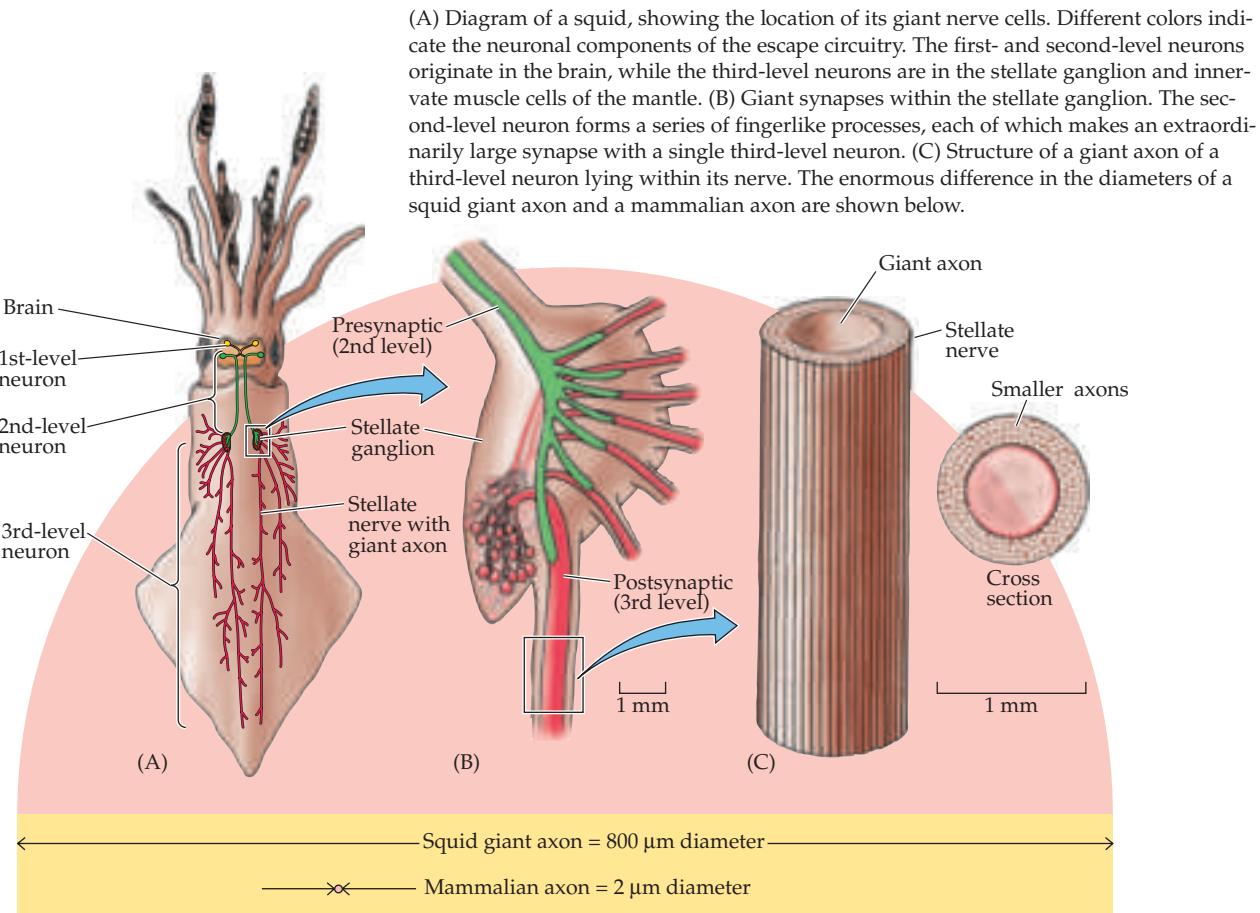
Giant neurons evidently evolved in squid because they enhanced survival. These neurons participate in a simple neural circuit that activates the contraction of the mantle muscle, producing a jet propulsion effect that allows the squid to move away from predators at a remarkably fast speed. As discussed in

Chapter 3, larger axonal diameter allows faster conduction of action potentials. Thus, presumably these huge nerve cells help squid escape more successfully from their numerous enemies.

Today—nearly 70 years after their discovery by John Z. Young at University College London—the giant nerve cells of squid remain useful experimental systems for probing basic neuronal functions.

References

- LLINÁS, R. (1999) *The Squid Synapse: A Model for Chemical Transmission*. Oxford: Oxford University Press.
YOUNG, J. Z. (1939) Fused neurons and synaptic contacts in the giant nerve fibres of cephalopods. Phil. Trans. R. Soc. Lond. 229(B): 465–503.



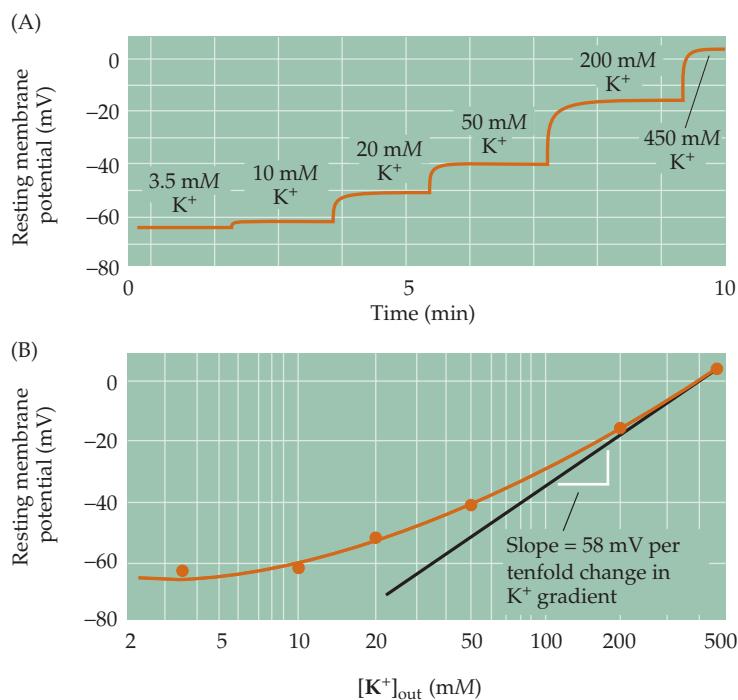


Figure 2.7 Experimental evidence that the resting membrane potential of a squid giant axon is determined by the K^+ concentration gradient across the membrane. (A) Increasing the external K^+ concentration makes the resting membrane potential more positive. (B) Relationship between resting membrane potential and external K^+ concentration, plotted on a semi-logarithmic scale. The straight line represents a slope of 58 mV per tenfold change in concentration, as given by the Nernst equation. (After Hodgkin and Katz, 1949.)

When Hodgkin and Katz carried out this experiment on a living squid neuron, they found that the resting membrane potential did indeed change when the external K^+ concentration was modified, becoming less negative as external K^+ concentration was raised (Figure 2.7A). When the external K^+ concentration was raised high enough to equal the concentration of K^+ inside the neuron, thus making the K^+ equilibrium potential 0 mV, the resting membrane potential was also approximately 0 mV. In short, the resting membrane potential varied as predicted with the logarithm of the K^+ concentration, with a slope that approached 58 mV per tenfold change in K^+ concentration (Figure 2.7B). The value obtained was not exactly 58 mV because other ions, such as Cl^- and Na^+ , are also slightly permeable, and thus influence the resting potential to a small degree. The contribution of these other ions is particularly evident at low external K^+ levels, again as predicted by the Goldman equation. In general, however, manipulation of the external concentrations of these other ions has only a small effect, emphasizing that K^+ permeability is indeed the primary source of the resting membrane potential.

In summary, Hodgkin and Katz showed that the inside-negative resting potential arises because (1) the membrane of the resting neuron is more permeable to K^+ than to any of the other ions present, and (2) there is more K^+ inside the neuron than outside. The selective permeability to K^+ is caused by K^+ -permeable membrane channels that are open in resting neurons, and the

large K⁺ concentration gradient is, as noted, produced by membrane transporters that selectively accumulate K⁺ within neurons. Many subsequent studies have confirmed the general validity of these principles.

The Ionic Basis of Action Potentials

What causes the membrane potential of a neuron to depolarize during an action potential? Although a general answer to this question has been given (increased permeability to Na⁺), it is well worth examining some of the experimental support for this concept. Given the data presented in Table 2.1, one can use the Nernst equation to calculate that the equilibrium potential for Na⁺ (E_{Na}) in neurons, and indeed in most cells, is positive. Thus, if the membrane were to become highly permeable to Na⁺, the membrane potential would approach E_{Na} . Based on these considerations, Hodgkin and Katz hypothesized that the action potential arises because the neuronal membrane becomes temporarily permeable to Na⁺.

Taking advantage of the same style of ion substitution experiment they used to assess the resting potential, Hodgkin and Katz tested the role of Na⁺ in generating the action potential by asking what happens to the action potential when Na⁺ is removed from the external medium. They found that lowering the external Na⁺ concentration reduces both the rate of rise of the action potential and its peak amplitude (Figure 2.8A–C). Indeed, when they examined this Na⁺ dependence quantitatively, they found a more-or-less linear relationship between the amplitude of the action potential and the logarithm of the external Na⁺ concentration (Figure 2.8D). The slope of this rela-

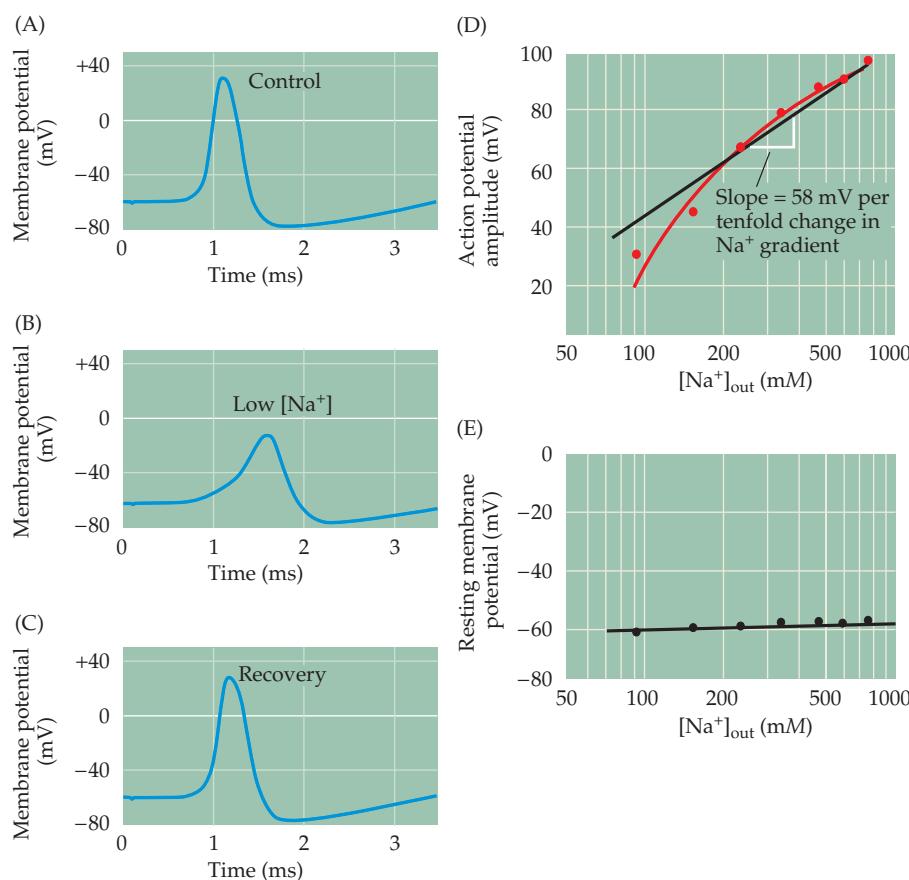


Figure 2.8 The role of sodium in the generation of an action potential in a squid giant axon. (A) An action potential evoked with the normal ion concentrations inside and outside the cell. (B) The amplitude and rate of rise of the action potential diminish when external sodium concentration is reduced to one-third of normal, but (C) recover when the Na⁺ is replaced. (D) While the amplitude of the action potential is quite sensitive to the external concentration of Na⁺, the resting membrane potential (E) is little affected by changing the concentration of this ion. (After Hodgkin and Katz, 1949.)

Box B

Action Potential Form and Nomenclature

The action potential of the squid giant axon has a characteristic shape, or waveform, with a number of different phases (Figure A). During the rising phase, the membrane potential rapidly depolarizes. In fact, action potentials cause the membrane potential to depolarize so much that the membrane potential transiently becomes positive with respect to the external medium, producing an overshoot. The overshoot of the action potential gives way to a falling phase in which the membrane potential rapidly repolarizes. Repolarization takes the membrane potential to levels even more negative than the resting membrane potential for a short time; this brief period of hyperpolarization is called the undershoot.

Although the waveform of the squid action potential is typical, the details of the action potential form vary widely from neuron to neuron in different animals. In myelinated axons of vertebrate motor neurons (Figure B), the action potential is virtually indistinguishable from that of the squid axon. However, the action potential recorded in the cell body of this same motor neuron (Figure

C) looks rather different. Thus, the action potential waveform can vary even within the same neuron. More complex action potentials are seen in other central neurons. For example, action potentials recorded from the cell bodies of neurons in the mammalian inferior olive (a region of the brainstem involved in motor control) last tens of milliseconds (Figure D). These action potentials exhibit a pronounced plateau during their falling phase, and their undershoot lasts even longer than that of the motor neuron. One of the most dramatic types of action potentials occurs in the cell bodies of cerebellar Purkinje neurons (Figure E). These potentials have several complex phases that result from the summation of multiple, discrete action potentials.

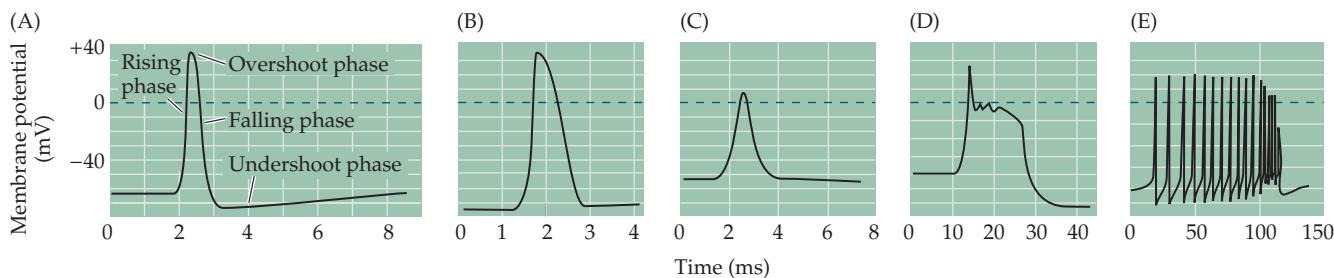
The variety of action potential waveforms could mean that each type of neuron has a different mechanism of action potential production. Fortunately, however, these diverse waveforms all result from relatively minor variations in the scheme used by the squid giant axon. For example, plateaus in the repolarization phase result from the presence of

ion channels that are permeable to Ca^{2+} , and long-lasting undershoots result from the presence of additional types of membrane K^+ channels. The complex action potential of the Purkinje cell results from these extra features plus the fact that different types of action potentials are generated in various parts of the Purkinje neuron—cell body, dendrites, and axons—and are summed together in recordings from the cell body. Thus, the lessons learned from the squid axon are applicable to, and indeed essential for, understanding action potential generation in all neurons.

References

- BARRETT, E. F. AND J. N. BARRETT (1976) Separation of two voltage-sensitive potassium currents, and demonstration of a tetrodotoxin-resistant calcium current in frog motoneurones. *J. Physiol. (Lond.)* 255: 737–774.
- DODGE, F. A. AND B. FRANKENHAEUSER (1958) Membrane currents in isolated frog nerve fibre under voltage clamp conditions. *J. Physiol. (Lond.)* 143: 76–90.
- HODGKIN, A. L. AND A. F. HUXLEY (1939) Action potentials recorded from inside a nerve fibre. *Nature* 144: 710–711.
- LLINÁS, R. AND M. SUGIMORI (1980) Electrophysiological properties of *in vitro* Purkinje cell dendrites in mammalian cerebellar slices. *J. Physiol. (Lond.)* 305: 197–213.
- LLINÁS, R. AND Y. YAROM (1981) Electrophysiology of mammalian inferior olivary neurones *in vitro*. Different types of voltage-dependent ionic conductances. *J. Physiol. (Lond.)* 315: 549–567.

(A) The phases of an action potential of the squid giant axon. (B) Action potential recorded from a myelinated axon of a frog motor neuron. (C) Action potential recorded from the cell body of a frog motor neuron. The action potential is smaller and the undershoot prolonged in comparison to the action potential recorded from the axon of this same neuron (B). (D) Action potential recorded from the cell body of a neuron from the inferior olive of a guinea pig. This action potential has a pronounced plateau during its falling phase. (E) Action potential recorded from the cell body of a Purkinje neuron in the cerebellum of a guinea pig. (A after Hodgkin and Huxley, 1939; B after Dodge and Frankenhaeuser, 1958; C after Barrett and Barrett, 1976; D after Llinás and Yarom, 1981; E after Llinás and Sugimori, 1980.)



tionship approached a value of 58 mV per tenfold change in Na^+ concentration, as expected for a membrane selectively permeable to Na^+ . In contrast, lowering Na^+ concentration had very little effect on the resting membrane potential (Figure 2.8E). Thus, while the resting neuronal membrane is only slightly permeable to Na^+ , the membrane becomes extraordinarily permeable to Na^+ during the **rising phase** and **overshoot phase** of the action potential (see Box B for an explanation of action potential nomenclature). This temporary increase in Na^+ permeability results from the opening of Na^+ -selective channels that are essentially closed in the resting state. Membrane pumps maintain a large electrochemical gradient for Na^+ , which is in much higher concentration outside the neuron than inside. When the Na^+ channels open, Na^+ flows into the neuron, causing the membrane potential to depolarize and approach E_{Na} .

The time that the membrane potential lingers near E_{Na} (about +58 mV) during the overshoot phase of an action potential is brief because the increased membrane permeability to Na^+ itself is short-lived. The membrane potential rapidly repolarizes to resting levels and is actually followed by a transient **undershoot**. As will be described in Chapter 3, these latter events in the action potential are due to an inactivation of the Na^+ permeability and an increase in the K^+ permeability of the membrane. During the undershoot, the membrane potential is transiently hyperpolarized because K^+ permeability becomes even greater than it is at rest. The action potential ends when this phase of enhanced K^+ permeability subsides, and the membrane potential thus returns to its normal resting level.

The ion substitution experiments carried out by Hodgkin and Katz provided convincing evidence that the resting membrane potential results from a high resting membrane permeability to K^+ , and that depolarization during an action potential results from a transient rise in membrane Na^+ permeability. Although these experiments identified the ions that flow during an action potential, they did not establish *how* the neuronal membrane is able to change its ionic permeability to generate the action potential, or what mechanisms trigger this critical change. The next chapter addresses these issues, documenting the surprising conclusion that the neuronal membrane potential itself affects membrane permeability.

Summary

Nerve cells generate electrical signals to convey information over substantial distances and to transmit it to other cells by means of synaptic connections. These signals ultimately depend on changes in the resting electrical potential across the neuronal membrane. A resting potential occurs because nerve cell membranes are permeable to one or more ion species subject to an electrochemical gradient. More specifically, a negative membrane potential at rest results from a net efflux of K^+ across neuronal membranes that are predominantly permeable to K^+ . In contrast, an action potential occurs when a transient rise in Na^+ permeability allows a net flow of Na^+ in the opposite direction across the membrane that is now predominantly permeable to Na^+ . The brief rise in membrane Na^+ permeability is followed by a secondary, transient rise in membrane K^+ permeability that repolarizes the neuronal membrane and produces a brief undershoot of the action potential. As a result of these processes, the membrane is depolarized in an all-or-none fashion during an action potential. When these active permeability changes subside, the membrane potential returns to its resting level because of the high resting membrane permeability to K^+ .

Additional Reading

Reviews

HODGKIN, A. L. (1951) The ionic basis of electrical activity in nerve and muscle. *Biol. Rev.* 26: 339–409.

HODGKIN, A. L. (1958) The Croonian Lecture: Ionic movements and electrical activity in giant nerve fibres. *Proc. R. Soc. Lond. (B)* 148: 1–37.

Important Original Papers

BAKER, P. F., A. L. HODGKIN AND T. I. SHAW (1962) Replacement of the axoplasm of giant nerve fibres with artificial solutions. *J. Physiol. (London)* 164: 330–354.

COLE, K. S. AND H. J. CURTIS (1939) Electric impedance of the squid giant axon during activity. *J. Gen. Physiol.* 22: 649–670.

GOLDMAN, D. E. (1943) Potential, impedance, and rectification in membranes. *J. Gen. Physiol.* 27: 37–60.

HODGKIN, A. L. AND P. HOROWICZ (1959) The influence of potassium and chloride ions on the membrane potential of single muscle fibres. *J. Physiol. (London)* 148: 127–160.

HODGKIN, A. L. AND B. KATZ (1949) The effect of sodium ions on the electrical activity of the giant axon of the squid. *J. Physiol. (London)* 108: 37–77.

HODGKIN, A. L. AND R. D. KEYNES (1953) The mobility and diffusion coefficient of potassium in giant axons from *Sepia*. *J. Physiol. (London)* 119: 513–528.

KEYNES, R. D. (1951) The ionic movements during nervous activity. *J. Physiol. (London)* 114: 119–150.

Books

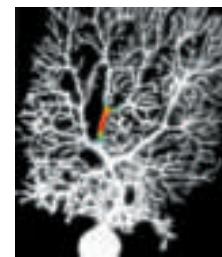
HODGKIN, A. L. (1967) *The Conduction of the Nervous Impulse*. Springfield, IL: Charles C. Thomas.

HODGKIN, A. L. (1992) *Chance and Design*. Cambridge: Cambridge University Press.

JUNGE, D. (1992) *Nerve and Muscle Excitation*, 3rd Ed. Sunderland, MA: Sinauer Associates.

KATZ, B. (1966) *Nerve, Muscle, and Synapse*. New York: McGraw-Hill.

Chapter 3



Voltage-Dependent Membrane Permeability

Overview

The action potential, the primary electrical signal generated by nerve cells, reflects changes in membrane permeability to specific ions. Present understanding of these changes in ionic permeability is based on evidence obtained by the voltage clamp technique, which permits detailed characterization of permeability changes as a function of membrane potential and time. For most types of axons, these changes consist of a rapid and transient rise in sodium (Na^+) permeability, followed by a slower but more prolonged rise in potassium (K^+) permeability. Both permeabilities are voltage-dependent, increasing as the membrane potential depolarizes. The kinetics and voltage dependence of Na^+ and K^+ permeabilities provide a complete explanation of action potential generation. Depolarizing the membrane potential to the threshold level causes a rapid, self-sustaining increase in Na^+ permeability that produces the rising phase of the action potential; however, the Na^+ permeability increase is short-lived and is followed by a slower increase in K^+ permeability that restores the membrane potential to its usual negative resting level. A mathematical model that describes the behavior of these ionic permeabilities predicts virtually all of the observed properties of action potentials. Importantly, this same ionic mechanism permits action potentials to be propagated along the length of neuronal axons, explaining how electrical signals are conveyed throughout the nervous system.

Ionic Currents Across Nerve Cell Membranes

The previous chapter introduced the idea that nerve cells generate electrical signals by virtue of a membrane that is differentially permeable to various ion species. In particular, a transient increase in the permeability of the neuronal membrane to Na^+ initiates the action potential. This chapter considers exactly how this increase in Na^+ permeability occurs. A key to understanding this phenomenon is the observation that action potentials are initiated *only* when the neuronal membrane potential becomes more positive than a threshold level. This observation suggests that the mechanism responsible for the increase in Na^+ permeability is sensitive to the membrane potential. Therefore, if one could understand how a change in membrane potential activates Na^+ permeability, it should be possible to explain how action potentials are generated.

The fact that the Na^+ permeability that generates the membrane potential change is itself sensitive to the membrane potential presents both conceptual and practical obstacles to studying the mechanism of the action potential. A practical problem is the difficulty of systematically varying the membrane

Box A

The Voltage Clamp Method

Breakthroughs in scientific research often rely on the development of new technologies. In the case of the action potential, detailed understanding came only after the invention of the voltage clamp technique by Kenneth Cole in the 1940s. This device is called a voltage clamp because it controls, or clamps, membrane potential (or voltage) at any level desired by the experimenter. The method measures the membrane potential with a microelectrode (or other type of electrode) placed inside the cell (1), and electronically compares this voltage to the voltage to be maintained (called the *command voltage*) (2). The clamp circuitry then passes a current back into the cell through another intracellular elec-

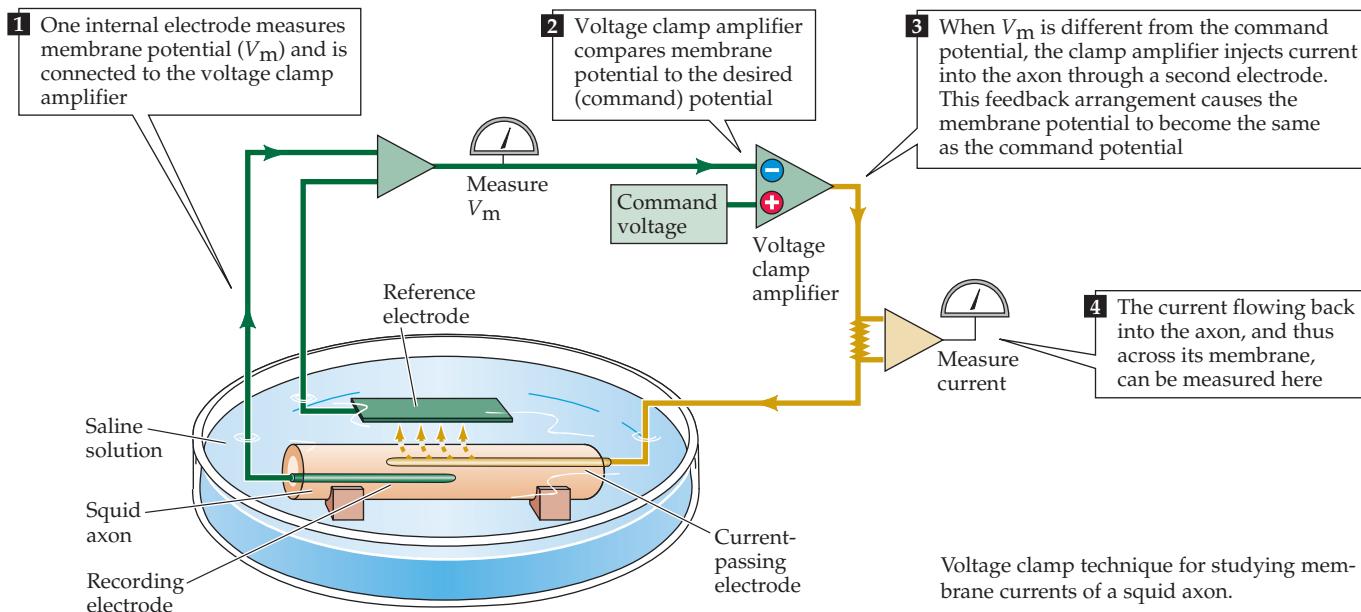
trode (3). This electronic feedback circuit holds the membrane potential at the desired level, even in the face of permeability changes that would normally alter the membrane potential (such as those generated during the action potential). Most importantly, the device permits the simultaneous measurement of the current needed to keep the cell at a given voltage (4). This current is exactly equal to the amount of current flowing across the neuronal membrane, allowing direct measurement of these membrane currents. Therefore, the voltage clamp technique can indicate how membrane potential influences ionic current flow across the membrane. This information gave Hodgkin and Huxley the key

insights that led to their model for action potential generation.

Today, the voltage clamp method remains widely used to study ionic currents in neurons and other cells. The most popular contemporary version of this approach is the patch clamp technique, a method that can be applied to virtually any cell and has a resolution high enough to measure the minute electrical currents flowing through single ion channels (see Box A in Chapter 4).

References

COLE, K. S. (1968) *Membranes, Ions and Impulses: A Chapter of Classical Biophysics*. Berkeley, CA: University of California Press.



Voltage clamp technique for studying membrane currents of a squid axon.

potential to study the permeability change, because such changes in membrane potential will produce an action potential, which causes further, uncontrolled changes in the membrane potential. Historically, then, it was not really possible to understand action potentials until a technique was developed that allowed experimenters to control membrane potential *and* simultaneously measure the underlying permeability changes. This tech-

nique, the **voltage clamp method** (Box A), provides the information needed to define the ionic permeability of the membrane at any level of membrane potential.

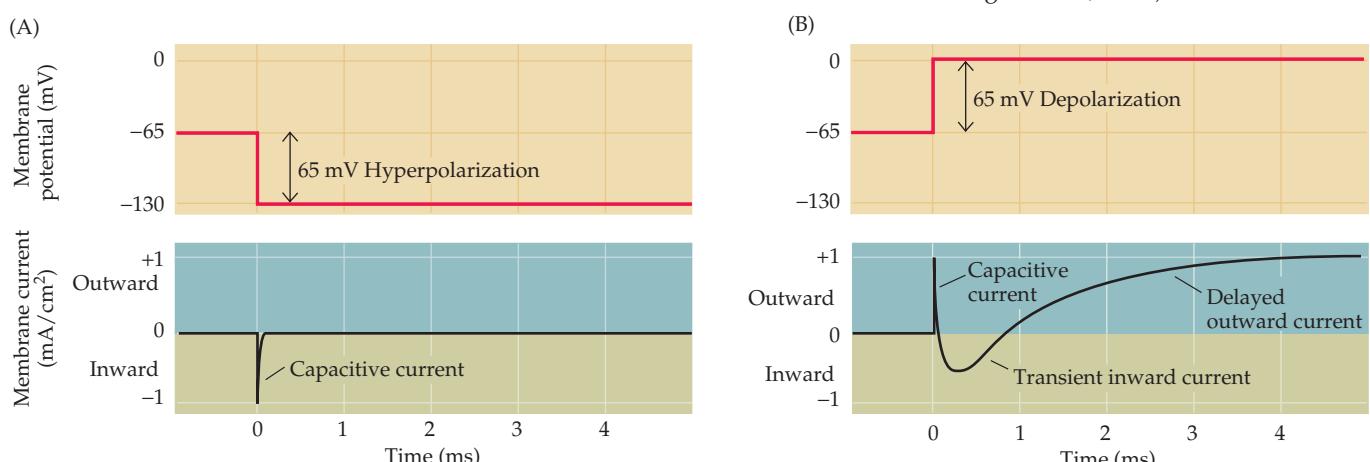
In the late 1940s, Alan Hodgkin and Andrew Huxley working at the University of Cambridge used the voltage clamp technique to work out the permeability changes underlying the action potential. They again chose to use the giant neuron of the squid because its large size (up to 1 mm in diameter; see Box A in Chapter 2) allowed insertion of the electrodes necessary for voltage clamping. They were the first investigators to test directly the hypothesis that potential-sensitive Na^+ and K^+ permeability changes are both necessary and sufficient for the production of action potentials.

Hodgkin and Huxley's first goal was to determine whether neuronal membranes do, in fact, have voltage-dependent permeabilities. To address this issue, they asked whether ionic currents flow across the membrane when its potential is changed. The result of one such experiment is shown in Figure 3.1. Figure 3.1A illustrates the currents produced by a squid axon when its membrane potential, V_m , is hyperpolarized from the resting level of -65 mV to -130 mV . The initial response of the axon results from the redistribution of charge across the axonal membrane. This capacitive current is nearly instantaneous, ending within a fraction of a millisecond. Aside from this brief event, very little current flows when the membrane is hyperpolarized. However, when the membrane potential is depolarized from -65 mV to 0 mV , the response is quite different (Figure 3.1B). Following the capacitive current, the axon produces a rapidly rising inward ionic current (inward refers to a positive charge entering the cell—that is, cations in or anions out), which gives way to a more slowly rising, delayed outward current. The fact that membrane depolarization elicits these ionic currents establishes that the membrane permeability of axons is indeed voltage-dependent.

Two Types of Voltage-Dependent Ionic Current

The results shown in Figure 3.1 demonstrate that the ionic permeability of neuronal membranes is voltage-sensitive, but the experiments do not identify how many types of permeability exist, or which ions are involved. As discussed in Chapter 2 (see Figure 2.5), varying the potential across a membrane makes it possible to deduce the equilibrium potential for the ionic fluxes through the membrane, and thus to identify the ions that are flowing.

Figure 3.1 Current flow across a squid axon membrane during a voltage clamp experiment. (A) A 65 mV hyperpolarization of the membrane potential produces only a very brief capacitive current. (B) A 65 mV depolarization of the membrane potential also produces a brief capacitive current, which is followed by a longer lasting but transient phase of inward current and a delayed but sustained outward current. (After Hodgkin et al., 1952.)



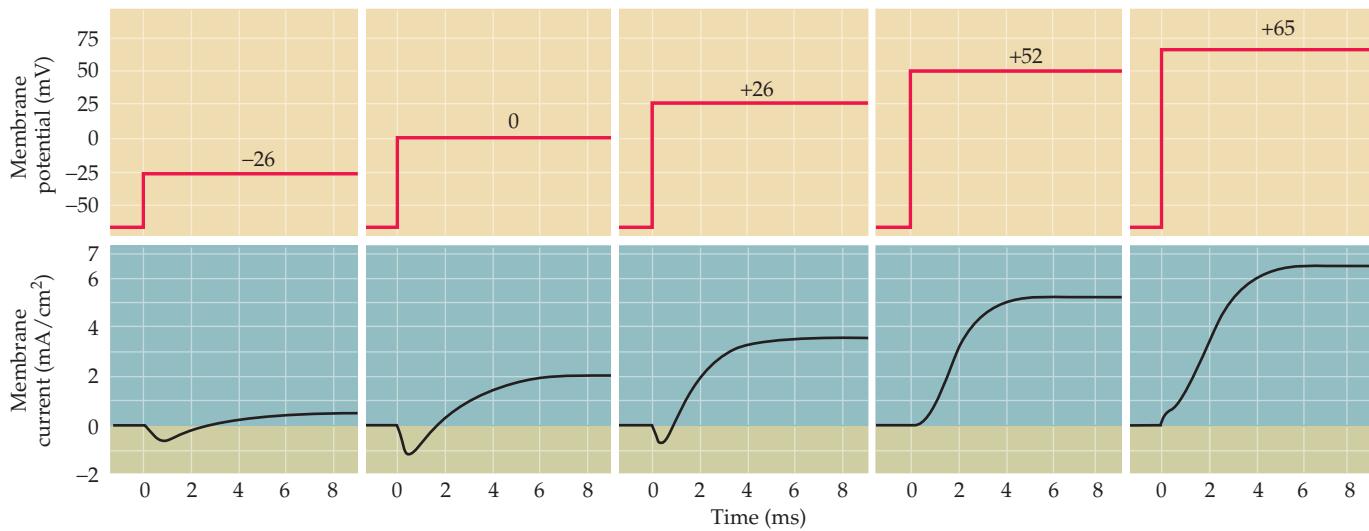


Figure 3.2 Current produced by membrane depolarizations to several different potentials. The early current first increases, then decreases in magnitude as the depolarization increases; note that this current is actually reversed in polarity at potentials more positive than about +55 mV. In contrast, the late current increases monotonically with increasing depolarization. (After Hodgkin et al., 1952.)

Because the voltage clamp method allows the membrane potential to be changed while ionic currents are being measured, it was a straightforward matter for Hodgkin and Huxley to determine ionic permeability by examining how the properties of the early inward and late outward currents changed as the membrane potential was varied (Figure 3.2). As already noted, no appreciable ionic currents flow at membrane potentials more negative than the resting potential. At more positive potentials, however, the currents not only flow but change in magnitude. The early current has a U-shaped dependence on membrane potential, increasing over a range of depolarizations up to approximately 0 mV but decreasing as the potential is depolarized further. In contrast, the late current increases monotonically with increasingly positive membrane potentials. These different responses to membrane potential can be seen more clearly when the magnitudes of the two current components are plotted as a function of membrane potential, as in Figure 3.3.

The voltage sensitivity of the early inward current gives an important clue about the nature of the ions carrying the current, namely, that no current flows when the membrane potential is clamped at +52 mV. For the squid neurons studied by Hodgkin and Huxley, the external Na^+ concentration is 440 mM, and the internal Na^+ concentration is 50 mM. For this concentration gradient, the Nernst equation predicts that the equilibrium poten-

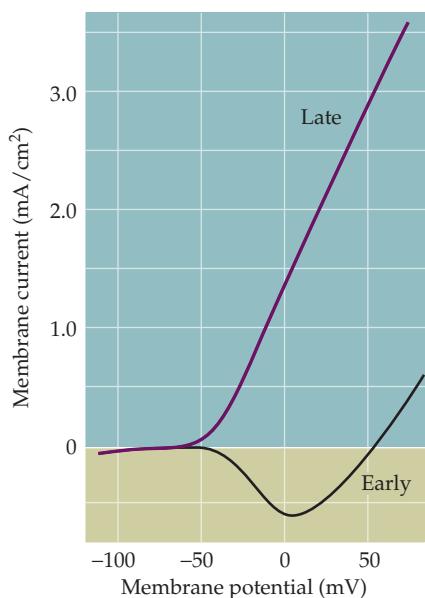


Figure 3.3 Relationship between current amplitude and membrane potential, taken from experiments such as the one shown in Figure 3.2. Whereas the late outward current increases steeply with increasing depolarization, the early inward current first increases in magnitude, but then decreases and reverses to outward current at about +55 mV (the sodium equilibrium potential). (After Hodgkin et al., 1952.)

Figure 3.4 Dependence of the early inward current on sodium. In the presence of normal external concentrations of Na^+ , depolarization of a squid axon to 0 mV produces an inward initial current. However, removal of external Na^+ causes the initial inward current to become outward, an effect that is reversed by restoration of external Na^+ . (After Hodgkin and Huxley, 1952a.)

tial for Na^+ should be +55 mV. Recall further from Chapter 2 that at the Na^+ equilibrium potential there is no net flux of Na^+ across the membrane, even if the membrane is highly permeable to Na^+ . Thus, the experimental observation that no current flows at the membrane potential where Na^+ cannot flow is a strong indication that the early inward current is carried by entry of Na^+ into the axon.

An even more demanding way to test whether Na^+ carries the early inward current is to examine the behavior of this current after removing external Na^+ . Removing the Na^+ outside the axon makes E_{Na} negative; if the permeability to Na^+ is increased under these conditions, current should flow outward as Na^+ leaves the neuron, due to the reversed electrochemical gradient. When Hodgkin and Huxley performed this experiment, they obtained the result shown in Figure 3.4. Removing external Na^+ caused the early inward current to reverse its polarity and become an outward current at a membrane potential that gave rise to an inward current when external Na^+ was present. This result demonstrates convincingly that the early inward current measured when Na^+ is present in the external medium must be due to Na^+ entering the neuron.

Notice that removal of external Na^+ in the experiment shown in Figure 3.4 has little effect on the outward current that flows after the neuron has been kept at a depolarized membrane voltage for several milliseconds. This further result shows that the late outward current must be due to the flow of an ion other than Na^+ . Several lines of evidence presented by Hodgkin, Huxley, and others showed that this late outward current is caused by K^+ exiting the neuron. Perhaps the most compelling demonstration of K^+ involvement is that the amount of K^+ efflux from the neuron, measured by loading the neuron with radioactive K^+ , is closely correlated with the magnitude of the late outward current.

Taken together, these experiments using the voltage clamp show that changing the membrane potential to a level more positive than the resting potential produces two effects: an early influx of Na^+ into the neuron, followed by a delayed efflux of K^+ . The early influx of Na^+ produces a transient inward current, whereas the delayed efflux of K^+ produces a sustained outward current. The differences in the time course and ionic selectivity of the two fluxes suggest that two different ionic permeability mechanisms are activated by changes in membrane potential. Confirmation that there are indeed two distinct mechanisms has come from pharmacological studies of drugs that specifically affect these two currents (Figure 3.5). **Tetrodotoxin**, an alkaloid neurotoxin found in certain puffer fish, tropical frogs, and salamanders, blocks the Na^+ current without affecting the K^+ current. Conversely, **tetraethylammonium ions** block K^+ currents without affecting Na^+ currents. The differential sensitivity of Na^+ and K^+ currents to these drugs provides strong additional evidence that Na^+ and K^+ flow through independent permeability pathways. As discussed in Chapter 4, it is now known that these pathways are ion channels that are selectively permeable to either Na^+ or K^+ . In fact, tetrodotoxin, tetraethylammonium, and other drugs that interact with spe-

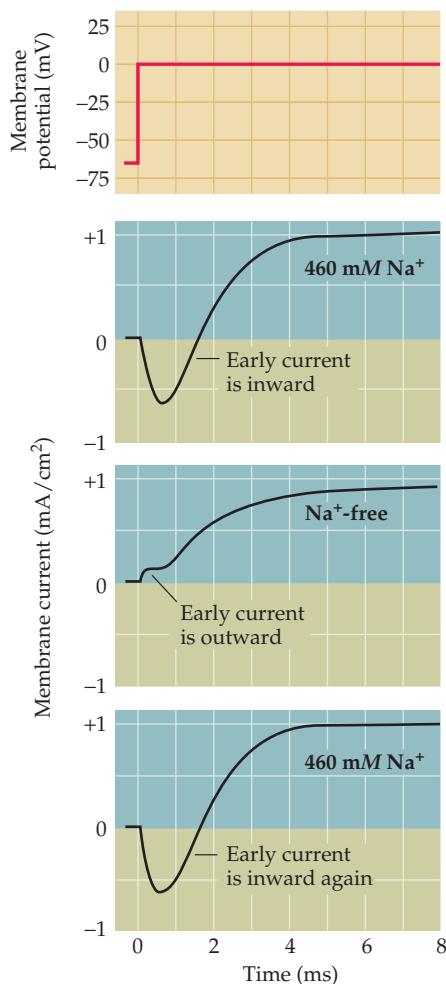
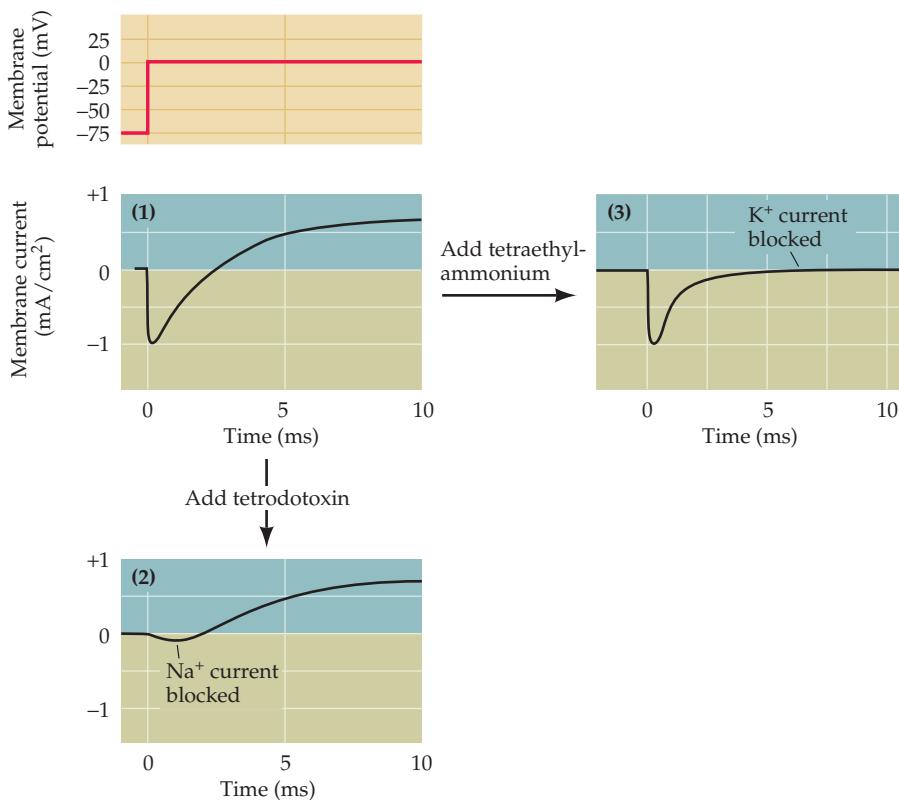


Figure 3.5 Pharmacological separation of Na^+ and K^+ currents into sodium and potassium components. Panel (1) shows the current that flows when the membrane potential of a squid axon is depolarized to 0 mV in control conditions. (2) Treatment with tetrodotoxin causes the early Na^+ currents to disappear but spares the late K^+ currents. (3) Addition of tetraethylammonium blocks the K^+ currents without affecting the Na^+ currents. (After Moore et al., 1967 and Armstrong and Binstock, 1965.)



cific types of ion channels have been extraordinarily useful tools in characterizing these channel molecules (see Chapter 4).

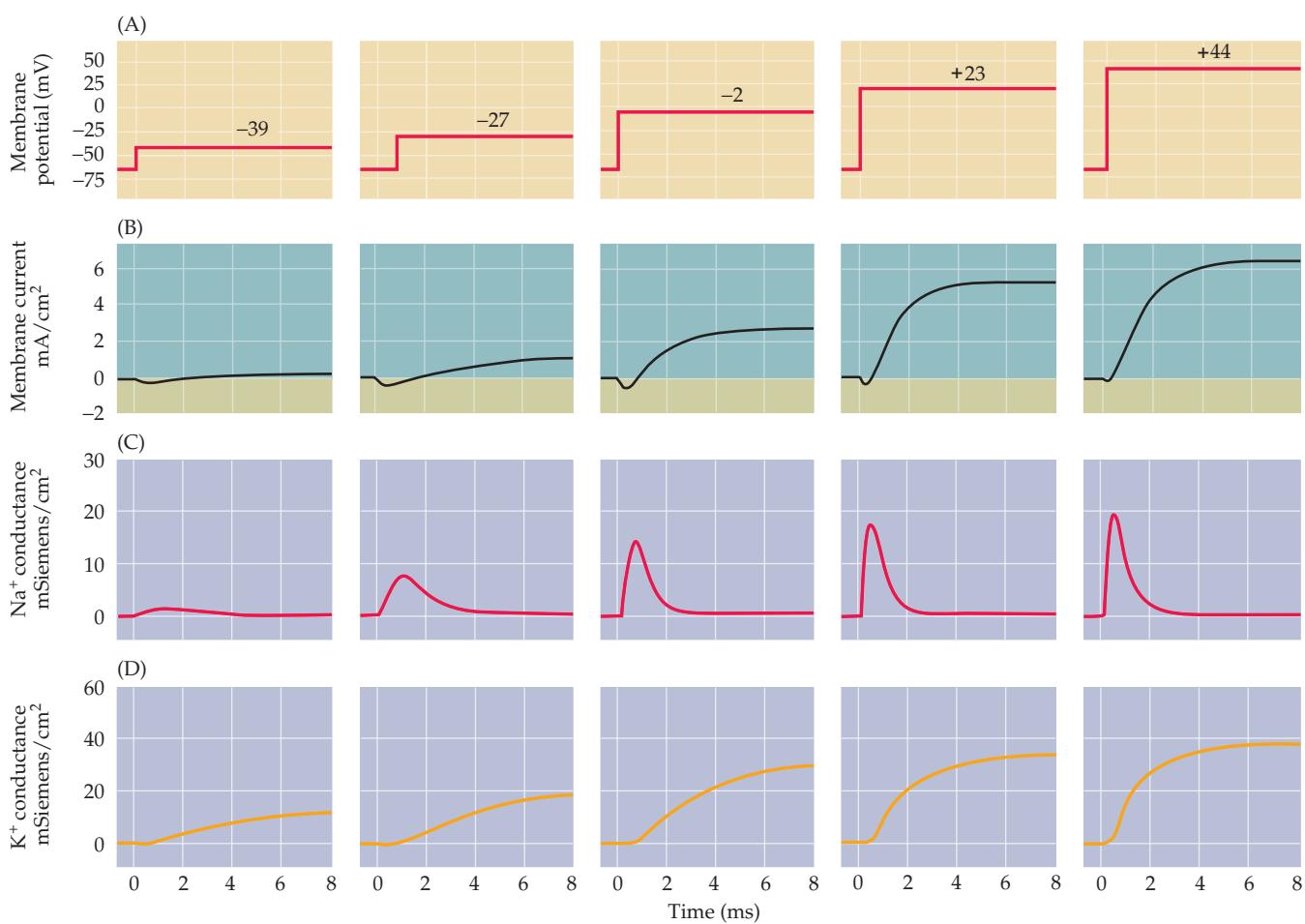
Two Voltage-Dependent Membrane Conductances

The next goal Hodgkin and Huxley set for themselves was to describe Na^+ and K^+ permeability changes mathematically. To do this, they assumed that the ionic currents are due to a change in **membrane conductance**, defined as the reciprocal of the membrane resistance. Membrane conductance is thus closely related, although not identical, to membrane permeability. When evaluating ionic movements from an electrical standpoint, it is convenient to describe them in terms of ionic conductances rather than ionic permeabilities. For present purposes, permeability and conductance can be considered synonymous. If membrane conductance (g) obeys Ohm's Law (which states that voltage is equal to the product of current and resistance), then the ionic current that flows during an increase in membrane conductance is given by

$$I_{\text{ion}} = g_{\text{ion}} (V_m - E_{\text{ion}})$$

where I_{ion} is the ionic current, V_m is the membrane potential, and E_{ion} is the equilibrium potential for the ion flowing through the conductance, g_{ion} . The difference between V_m and E_{ion} is the electrochemical driving force acting on the ion.

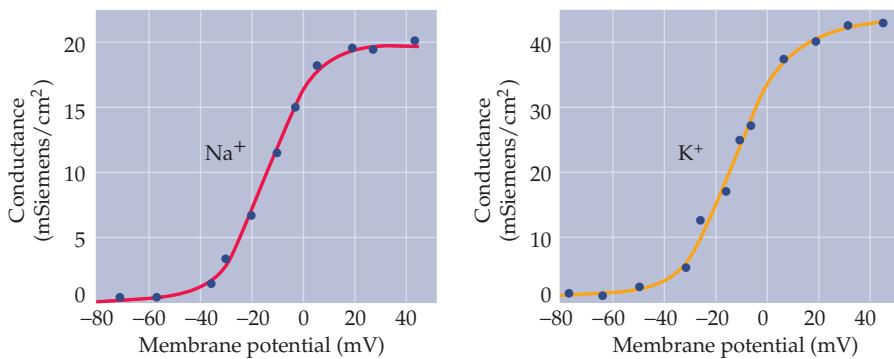
Hodgkin and Huxley used this simple relationship to calculate the dependence of Na^+ and K^+ conductances on time and membrane potential. They knew V_m , which was set by their voltage clamp device (Figure 3.6A), and could determine E_{Na} and E_{K} from the ionic concentrations on the two sides



of the axonal membrane (see Table 2.1). The currents carried by Na^+ and K^+ — I_{Na} and I_{K} —could be determined separately from recordings of the membrane currents resulting from depolarization (Figure 3.6B) by measuring the difference between currents recorded in the presence and absence of external Na^+ (as shown in Figure 3.4). From these measurements, Hodgkin and Huxley were able to calculate g_{Na} and g_{K} (Figure 3.6C,D), from which they drew two fundamental conclusions. The first conclusion is that the Na^+ and K^+ conductances change over time. For example, both Na^+ and K^+ conductances require some time to **activate**, or turn on. In particular, the K^+ conductance has a pronounced delay, requiring several milliseconds to reach its maximum (Figure 3.6D), whereas the Na^+ conductance reaches its maximum more rapidly (Figure 3.6C). The more rapid activation of the Na^+ conductance allows the resulting inward Na^+ current to precede the delayed outward K^+ current (see Figure 3.6B). Although the Na^+ conductance rises rapidly, it quickly declines, even though the membrane potential is kept at a depolarized level. This fact shows that depolarization not only causes the Na^+ conductance to activate, but also causes it to decrease over time, or **inactivate**. The K^+ conductance of the squid axon does not inactivate in this way; thus, while the Na^+ and K^+ conductances share the property of time-dependent activation, only the Na^+ conductance inactivates. (Inactivating K^+ conductances have since been discovered in other types of nerve cells; see Chapter 4.) The time courses of the Na^+ and K^+ conductances are voltage-

Figure 3.6 Membrane conductance changes underlying the action potential are time- and voltage-dependent. Depolarizations to various membrane potentials (A) elicit different membrane currents (B). Below are shown the Na^+ (C) and K^+ (D) conductances calculated from these currents. Both peak Na^+ conductance and steady-state K^+ conductance increase as the membrane potential becomes more positive. In addition, the activation of both conductances, as well as the rate of inactivation of the Na^+ conductance, occur more rapidly with larger depolarizations. (After Hodgkin and Huxley, 1952b.)

Figure 3.7 Depolarization increases Na^+ and K^+ conductances of the squid giant axon. The peak magnitude of Na^+ conductance and steady-state value of K^+ conductance both increase steeply as the membrane potential is depolarized. (After Hodgkin and Huxley, 1952b.)



dependent, with the speed of both activation and inactivation increasing at more depolarized potentials. This finding accounts for more rapid time courses of membrane currents measured at more depolarized potentials.

The second conclusion derived from Hodgkin and Huxley's calculations is that both the Na^+ and K^+ conductances are voltage-dependent—that is, both conductances increase progressively as the neuron is depolarized. Figure 3.7 illustrates this by plotting the relationship between peak value of the conductances (from Figure 3.6C,D) against the membrane potential. Note the similar voltage dependence for each conductance; both conductances are quite small at negative potentials, maximal at very positive potentials, and exquisitely dependent on membrane voltage at intermediate potentials. The observation that these conductances are sensitive to changes in membrane potential shows that the mechanism underlying the conductances somehow “senses” the voltage across the membrane.

All told, the voltage clamp experiments carried out by Hodgkin and Huxley showed that the ionic currents that flow when the neuronal membrane is depolarized are due to three different voltage-sensitive processes: (1) activation of Na^+ conductance, (2) activation of K^+ conductance, and (3) inactivation of Na^+ conductance.

Reconstruction of the Action Potential

From their experimental measurements, Hodgkin and Huxley were able to construct a detailed mathematical model of the Na^+ and K^+ conductance changes. The goal of these modeling efforts was to determine whether the Na^+ and K^+ conductances alone are sufficient to produce an action potential. Using this information, they could in fact generate the form and time course of the action potential with remarkable accuracy (Figure 3.8A). Further, the Hodgkin-Huxley model predicted other features of action potential behavior in the squid axon, such as how the delay before action potential generation changes in response to stimulating currents of different intensities (Figure 3.8B,C). The model also predicted that the axon membrane would become refractory to further excitation for a brief period following an action potential, as was experimentally observed.

The Hodgkin-Huxley model also provided many insights into how action potentials are generated. Figure 3.8A shows a reconstructed action potential, together with the time courses of the underlying Na^+ and K^+ conductances. The coincidence of the initial increase in Na^+ conductance with the rapid rising phase of the action potential demonstrates that a selective increase in

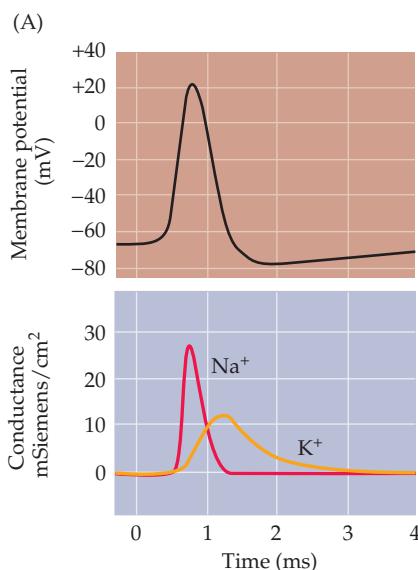
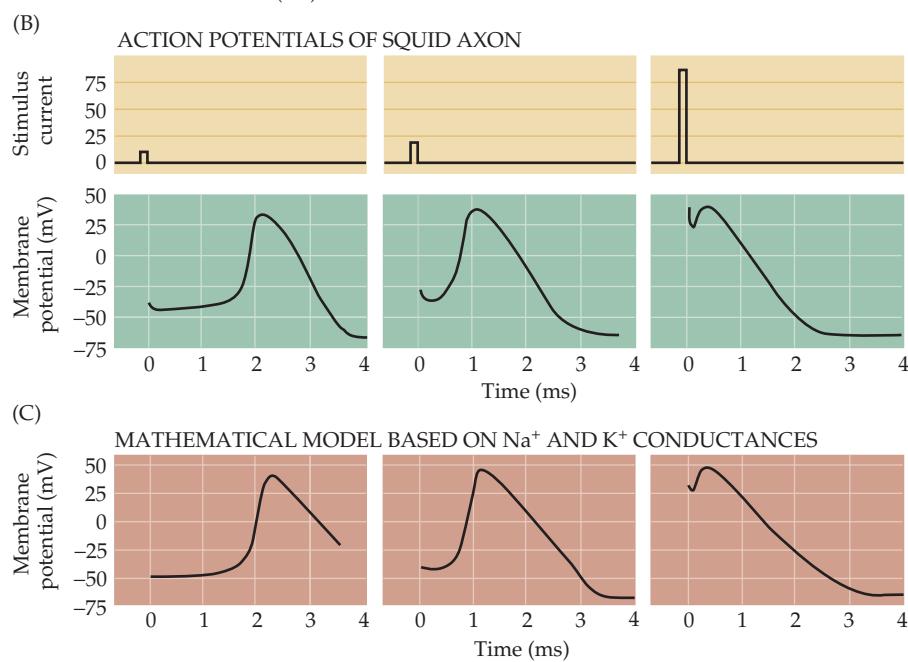


Figure 3.8 Mathematical reconstruction of the action potential. (A) Reconstruction of an action potential (black curve) together with the underlying changes in Na^+ (red curve) and K^+ (yellow curve) conductance. The size and time course of the action potential were calculated using only the properties of g_{Na} and g_{K} measured in voltage clamp experiments. Real action potentials evoked by brief current pulses of different intensities (B) are remarkably similar to those generated by the mathematical model (C). The reconstructed action potentials shown in (A) and (C) differ in duration because (A) simulates an action potential at 19°C, whereas (C) simulates an action potential at 6°C. (After Hodgkin and Huxley, 1952d.)



Na^+ conductance is responsible for action potential initiation. The increase in Na^+ conductance causes Na^+ to enter the neuron, thus depolarizing the membrane potential, which approaches E_{Na} . The rate of depolarization subsequently falls both because the electrochemical driving force on Na^+ decreases and because the Na^+ conductance inactivates. At the same time, depolarization slowly activates the voltage-dependent K^+ conductance, causing K^+ to leave the cell and repolarizing the membrane potential toward E_{K} . Because the K^+ conductance becomes temporarily higher than it is in the resting condition, the membrane potential actually becomes briefly more negative than the normal resting potential (the **undershoot**). The hyperpolarization of the membrane potential causes the voltage-dependent K^+ conductance (and any Na^+ conductance not inactivated) to turn off, allowing the membrane potential to return to its resting level.

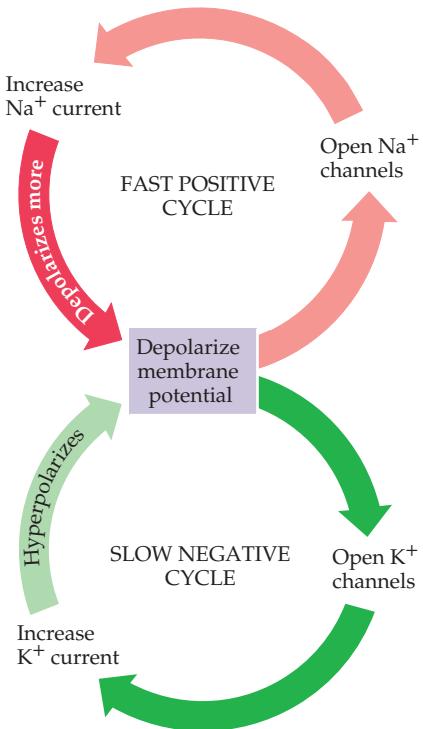


Figure 3.9 Feedback cycles responsible for membrane potential changes during an action potential. Membrane depolarization rapidly activates a positive feedback cycle fueled by the voltage-dependent activation of Na⁺ conductance. This phenomenon is followed by the slower activation of a negative feedback loop as depolarization activates a K⁺ conductance, which helps to repolarize the membrane potential and terminate the action potential.

This mechanism of action potential generation represents a positive feedback loop: Activating the voltage-dependent Na⁺ conductance increases Na⁺ entry into the neuron, which makes the membrane potential depolarize, which leads to the activation of still more Na⁺ conductance, more Na⁺ entry, and still further depolarization (Figure 3.9). Positive feedback continues unabated until Na⁺ conductance inactivation and K⁺ conductance activation restore the membrane potential to the resting level. Because this positive feedback loop, once initiated, is sustained by the intrinsic properties of the neuron—namely, the voltage dependence of the ionic conductances—the action potential is self-supporting, or **regenerative**. This regenerative quality explains why action potentials exhibit all-or-none behavior (see Figure 2.1), and why they have a threshold (Box B). The delayed activation of the K⁺ conductance represents a negative feedback loop that eventually restores the membrane to its resting state.

Hodgkin and Huxley's reconstruction of the action potential and all its features shows that the properties of the voltage-sensitive Na⁺ and K⁺ conductances, together with the electrochemical driving forces created by ion transporters, are sufficient to explain action potentials. Their use of both empirical and theoretical methods brought an unprecedented level of rigor to a long-standing problem, setting a standard of proof that is achieved only rarely in biological research.

Long-Distance Signaling by Means of Action Potentials

The voltage-dependent mechanisms of action potential generation also explain the long-distance transmission of these electrical signals. Recall from Chapter 2 that neurons are relatively poor conductors of electricity, at least compared to a wire. Current conduction by wires, and by neurons in the absence of action potentials, is called **passive current flow** (Box C). The passive electrical properties of a nerve cell axon can be determined by measuring the voltage change resulting from a current pulse passed across the axonal membrane (Figure 3.10A). If this current pulse is not large enough to generate action potentials, the magnitude of the resulting potential change decays exponentially with increasing distance from the site of current injection (Figure 3.10B). Typically, the potential falls to a small fraction of its initial value at a distance of no more than a couple of millimeters away from the site of injection (Figure 3.10C). The progressive decrease in the amplitude of the induced potential change occurs because the injected current leaks out across the axonal membrane; accordingly, less current is available to change the membrane potential farther along the axon. Thus, the leakiness of the axonal membrane prevents effective passive transmission of electrical signals in all but the shortest axons (those 1 mm or less in length). Likewise, the leakiness of the membrane slows the time course of the responses measured at increasing distances from the site where current was injected (Figure 3.10D).

Box B

Threshold

An important—and potentially puzzling—property of the action potential is its initiation at a particular membrane potential, called threshold. Indeed, action potentials never occur without a depolarizing stimulus that brings the membrane to this level. The depolarizing “trigger” can be one of several events: a synaptic input, a receptor potential generated by specialized receptor organs, the endogenous pacemaker activity of cells that generate action potentials spontaneously, or the local current that mediates the spread of the action potential down the axon.

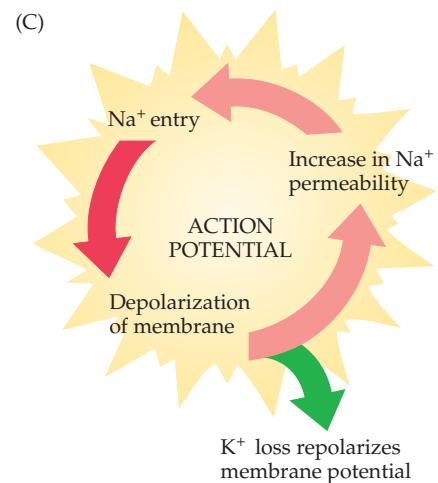
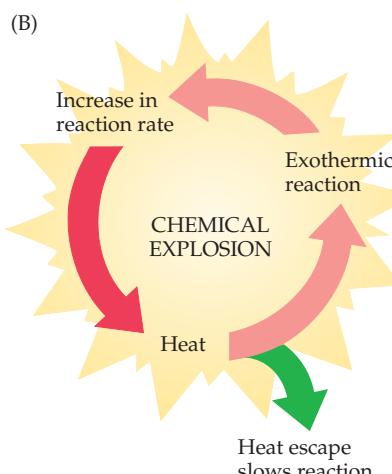
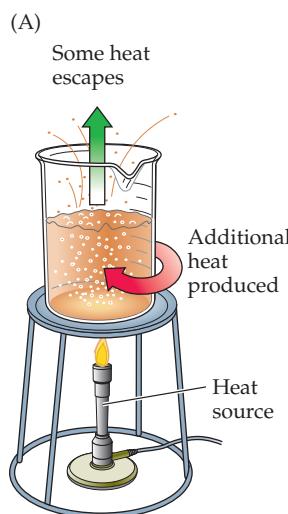
Why the action potential “takes off” at a particular level of depolarization can be understood by comparing the underlying events to a chemical explosion (Figure A). Exogenous heat (analogous to the initial depolarization of the membrane potential) stimulates an exothermic chemical reaction, which produces more heat, which further enhances the reaction (Figure B). As a result of this positive feedback loop, the rate of the reaction builds up exponentially—the definition of an explosion. In any such

process, however, there is a threshold, that is, a point up to which heat can be supplied without resulting in an explosion. The threshold for the chemical explosion diagrammed here is the point at which the amount of heat supplied exogenously is just equal to the amount of heat that can be dissipated by the circumstances of the reaction (such as escape of heat from the beaker).

The threshold of action potential initiation is, in principle, similar (Figure C). There is a range of “subthreshold” depolarization, within which the rate of increased sodium entry is less than the rate of potassium exit (remember that the membrane at rest is highly permeable to K^+ , which therefore flows out as the membrane is depolarized). The point at which Na^+ inflow just equals K^+ outflow represents an unstable equilibrium analogous to the ignition point of an explosive mixture. The behavior of the membrane at threshold reflects this instability: The membrane potential may linger at the threshold level for a variable period before either returning to the resting level or flaring up into a full-blown

action potential. In theory at least, if there is a net internal gain of a single Na^+ ion, an action potential occurs; conversely, the net loss of a single K^+ ion leads to repolarization. A more precise definition of threshold, therefore, is that value of membrane potential, in depolarizing from the resting potential, at which the current carried by Na^+ entering the neuron is exactly equal to the K^+ current that is flowing out. Once the triggering event depolarizes the membrane beyond this point, the positive feedback loop of Na^+ entry on membrane potential closes and the action potential “fires.”

Because the Na^+ and K^+ conductances change dynamically over time, the threshold potential for producing an action potential also varies as a consequence of the previous activity of the neuron. For example, following an action potential, the membrane becomes temporarily refractory to further excitation because the threshold for firing an action potential transiently rises. There is, therefore, no specific value of membrane potential that defines the threshold for a given nerve cell in all circumstances.



A positive feedback loop underlying the action potential explains the phenomenon of threshold.

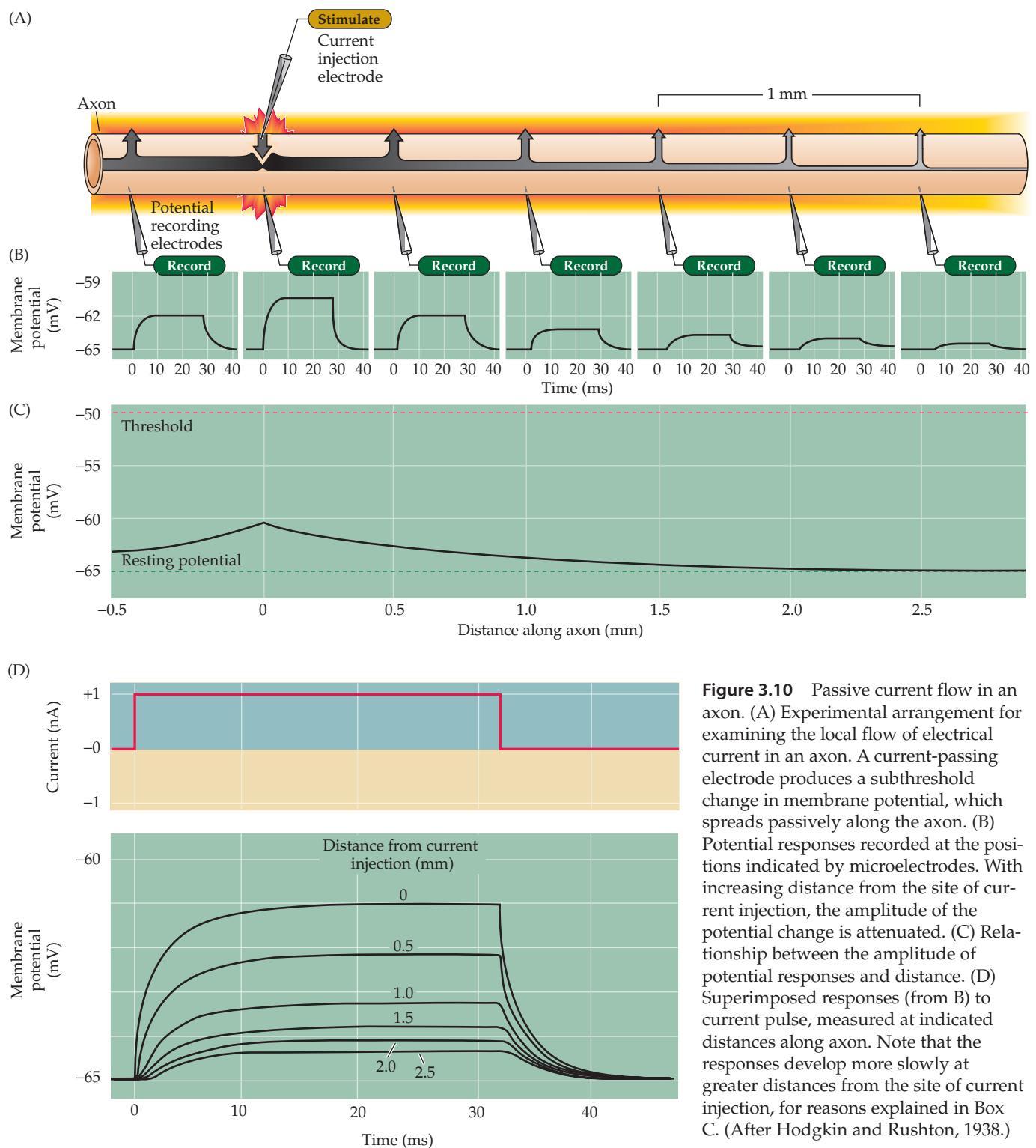
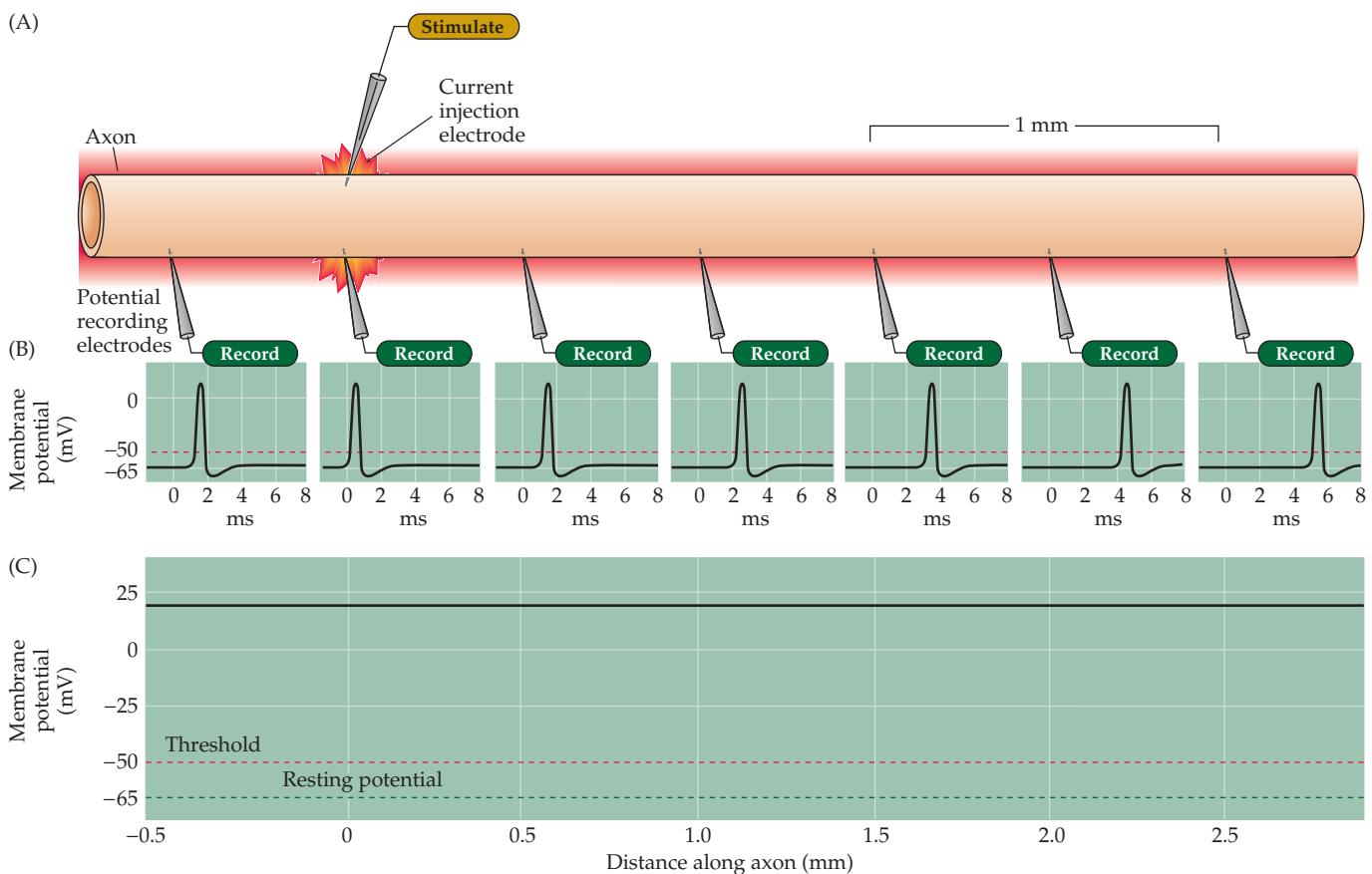


Figure 3.10 Passive current flow in an axon. (A) Experimental arrangement for examining the local flow of electrical current in an axon. A current-passing electrode produces a subthreshold change in membrane potential, which spreads passively along the axon. (B) Potential responses recorded at the positions indicated by microelectrodes. With increasing distance from the site of current injection, the amplitude of the potential change is attenuated. (C) Relationship between the amplitude of potential responses and distance. (D) Superimposed responses (from B) to current pulse, measured at indicated distances along axon. Note that the responses develop more slowly at greater distances from the site of current injection, for reasons explained in Box C. (After Hodgkin and Rushton, 1938.)

If the experiment shown in Figure 3.10 is repeated with a depolarizing current pulse large enough to produce an action potential, the result is dramatically different (Figure 3.11A). In this case, an action potential occurs without decrement along the entire length of the axon, which in humans



may be a distance of a meter or more (Figure 3.11B). Thus, action potentials somehow circumvent the inherent leakiness of neurons.

How, then, do action potentials traverse great distances along such a poor passive conductor? The answer is in part provided by the observation that the amplitude of the action potentials recorded at different distances is constant. This all-or-none behavior indicates that more than simple passive flow of current must be involved in action potential propagation. A second clue comes from examination of the time of occurrence of the action potentials recorded at different distances from the site of stimulation: Action potentials occur later and later at greater distances along the axon (Figure 3.11B). Thus, the action potential has a measurable rate of transmission, called the **conduction velocity**. The delay in the arrival of the action potential at successively more distant points along the axon differs from the case shown in Figure 3.10, in which the electrical changes produced by passive current flow occur at more or less the same time at successive points.

The mechanism of action potential propagation is easy to grasp once one understands how action potentials are generated and how current passively flows along an axon (Figure 3.12). A depolarizing stimulus—a synaptic potential or a receptor potential in an intact neuron, or an injected current pulse in an experiment—locally depolarizes the axon, thus opening the voltage-sensitive Na^+ channels in that region. The opening of Na^+ channels causes inward movement of Na^+ , and the resultant depolarization of the membrane potential generates an action potential at that site. Some of the local current generated by the action potential will then flow passively down

Figure 3.11 Propagation of an action potential. (A) In this experimental arrangement, an electrode evokes an action potential by injecting a suprathreshold current. (B) Potential responses recorded at the positions indicated by microelectrodes. The amplitude of the action potential is constant along the length of the axon, although the time of appearance of the action potential is delayed with increasing distance. (C) The constant amplitude of an action potential (solid black line) measured at different distances.

Box C

Passive Membrane Properties

The passive flow of electrical current plays a central role in action potential propagation, synaptic transmission, and all other forms of electrical signaling in nerve cells. Therefore, it is worthwhile understanding in quantitative terms how passive current flow varies with distance along a neuron. For the case of a cylindrical axon, such as the one depicted in Figure 3.10, subthreshold current injected into one part of the axon spreads passively along the axon until the current is dissipated by leakage out across the axon membrane. The decrement in the current flow with distance (Figure A) is described by a simple exponential function:

$$V_x = V_0 e^{-x/\lambda}$$

where V_x is the voltage response at any distance x along the axon, V_0 is the voltage change at the point where current is injected into the axon, e is the base of natural logarithms (approximately 2.7), and λ is the length constant of the axon. As evident in this relationship, the length constant is the distance where the initial voltage response (V_0) decays to $1/e$ (or 37%) of its value. The length constant is thus a way to characterize how far passive current flow spreads before it leaks out of the axon, with leakier axons having shorter length constants.

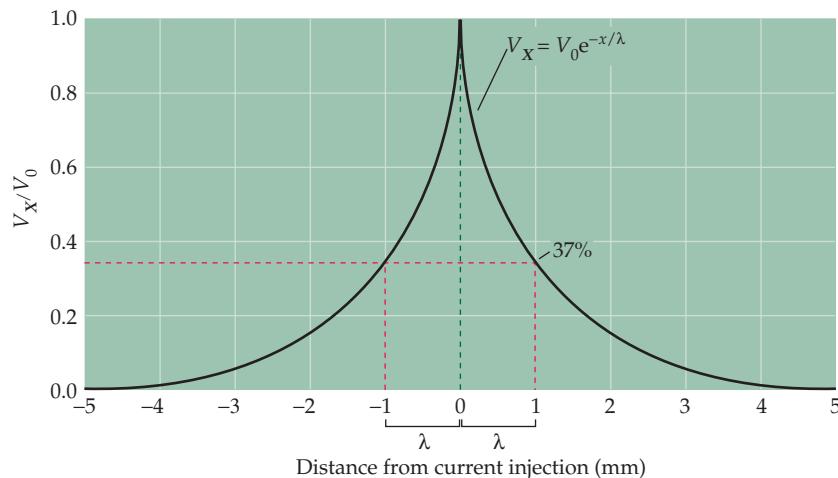
The length constant depends upon the physical properties of the axon, in particular the relative resistances of the

plasma membrane (r_m), the intracellular axoplasm (r_i), and the extracellular medium (r_0). The relationship between these parameters is:

$$\lambda = \sqrt{\frac{r_m}{r_0 + r_i}}$$

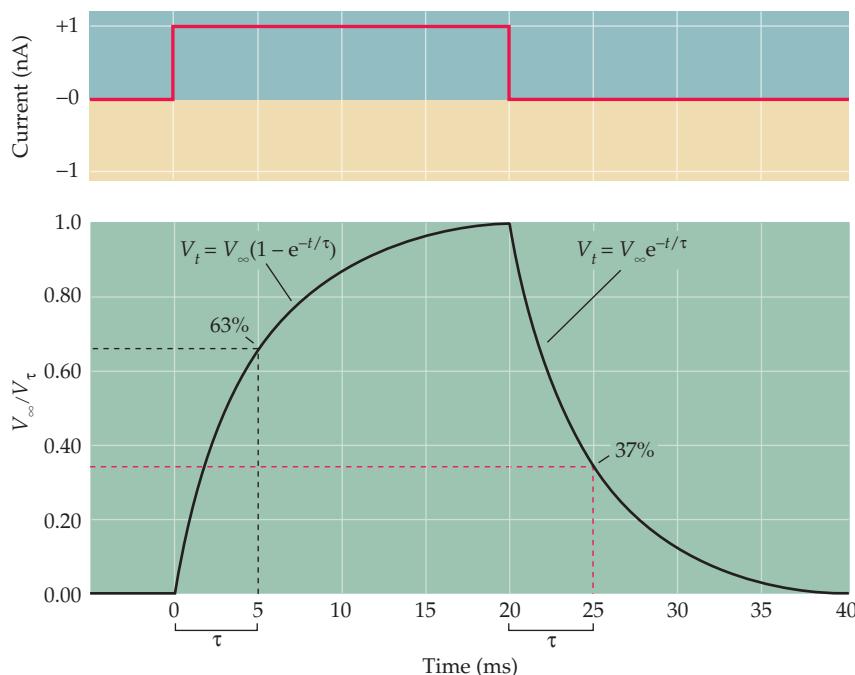
Hence, to improve the passive flow of current along an axon, the resistance of the plasma membrane should be as high as possible and the resistances of the axoplasm and extracellular medium should be low.

Another important consequence of the passive properties of neurons is that currents flowing across a membrane do not immediately change the membrane potential. For example, when a rectangular current pulse is injected into the axon shown in the experiment illustrated in Figure 3.10A, the membrane potential depolarizes slowly over a few milliseconds and then repolarizes over a similar time course when the current pulse ends (see Figure 3.10D). These delays in changing the membrane potential are due to the fact that the plasma mem-



(A) Spatial decay of membrane potential along a cylindrical axon. A current pulse injected at one point in the axon (0 mm) produces voltage responses (V_x) that decay exponentially with distance. The distance where the voltage response is $1/e$ of its initial value (V_0) is the length constant, λ .

the axon, in the same way that subthreshold currents spread along the axon (see Figure 3.10). Note that this passive current flow does not require the movement of Na^+ along the axon but, instead, occurs by a shuttling of charge, somewhat similar to what happens when wires passively conduct electricity by transmission of electron charge. This passive current flow depolarizes the membrane potential in the adjacent region of the axon, thus opening the Na^+ channels in the neighboring membrane. The local depolarization triggers an action potential in this region, which then spreads again in a continuing cycle until the end of the axon is reached. Thus, action potential propagation requires the coordinated action of two forms of current



(B) Time course of potential changes produced in a spatially uniform cell by a current pulse. The rise and fall of the membrane potential (V_t) can be described as exponential functions, with the time constant τ defining the time required for the response to rise to $1 - (1/e)$ of the steady-state value (V_∞), or to decline to $1/e$ of V_∞ .

brane behaves as a capacitor, storing the initial charge that flows at the beginning and end of the current pulse. For the case of a cell whose membrane potential is spatially uniform, the change in the membrane potential at any time, V_t , after beginning the current pulse (Figure B) can also be described by an exponential relationship:

$$V_t = V_\infty(1 - e^{-t/\tau})$$

where V_∞ is the steady-state value of the

membrane potential change, t is the time after the current pulse begins, and τ is the membrane time constant. The time constant is thus defined as the time when the voltage response (V_t) rises to $1 - (1/e)$ (or 63%) of V_∞ . After the current pulse ends, the membrane potential change also declines exponentially according to the relationship

$$V_t = V_\infty e^{-t/\tau}$$

During this decay, the membrane poten-

tial returns to $1/e$ of V_∞ at a time equal to t . For cells with more complex geometries than the axon in Figure 3.10, the time courses of the changes in membrane potential are not simple exponentials, but nonetheless depend on the membrane time constant. Thus, the time constant characterizes how rapidly current flow changes the membrane potential. The membrane time constant also depends on the physical properties of the nerve cell, specifically on the resistance (r_m) and capacitance (c_m) of the plasma membrane such that:

$$\tau = r_m c_m$$

The values of r_m and c_m depend, in part, on the size of the neuron, with larger cells having lower resistances and larger capacitances. In general, small nerve cells tend to have long time constants and large cells brief time constants.

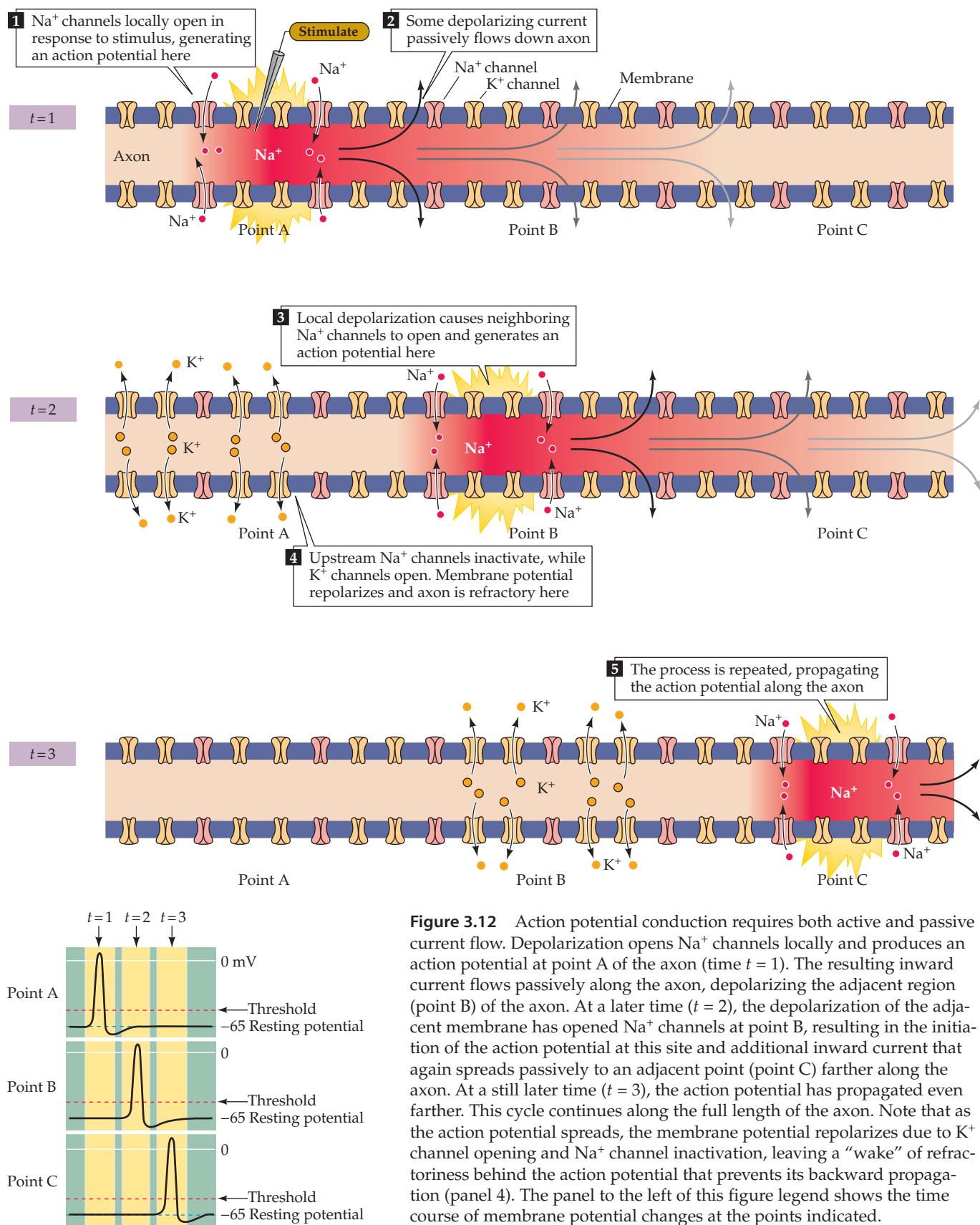
References

- HODGKIN, A. L. AND W. A. H. RUSHTON (1938) The electrical constants of a crustacean nerve fibre. Proc. R. Soc. Lond. 133: 444–479.
- JOHNSTON, D. AND S. M.-S. WU (1995) *Foundations of Cellular Neurophysiology*. Cambridge, MA: MIT Press.
- RAILL, W. (1977) Core conductor theory and cable properties of neurons. In *Handbook of Physiology*, Section 1: *The Nervous System*, Vol. 1: *Cellular Biology of Neurons*. E. R. Kandel (ed.). Bethesda, MD: American Physiological Society, pp. 39–98.

flow—the passive flow of current as well as active currents flowing through voltage-dependent ion channels. The regenerative properties of Na^+ channel opening allow action potentials to propagate in an all-or-none fashion by acting as a booster at each point along the axon, thus ensuring the long-distance transmission of electrical signals.

The Refractory Period

Recall that the depolarization that produces Na^+ channel opening also causes delayed activation of K^+ channels and Na^+ channel inactivation, lead-

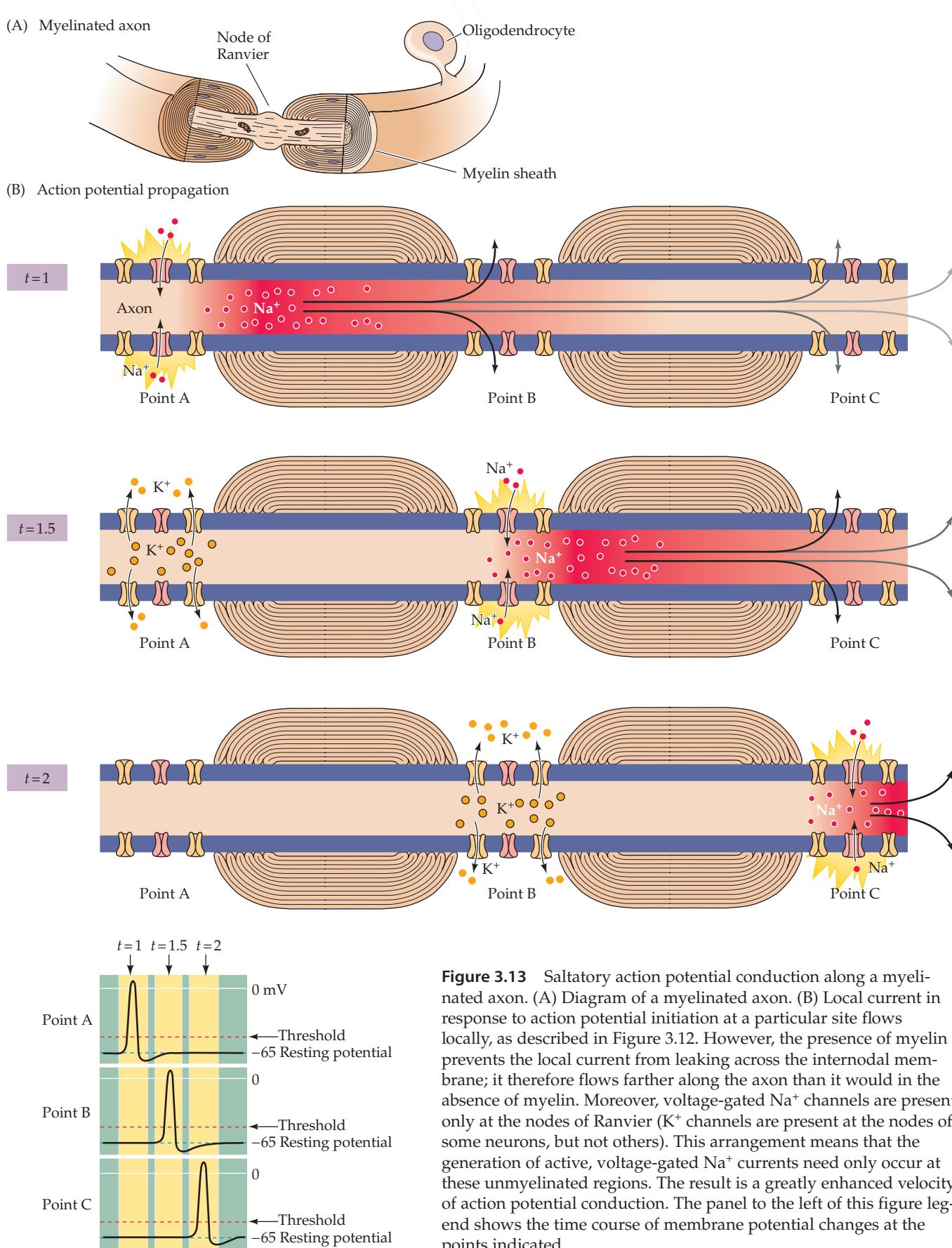


ing to repolarization of the membrane potential as the action potential sweeps along the length of an axon (see Figure 3.12). In its wake, the action potential leaves the Na^+ channels inactivated and K^+ channels activated for a brief time. These transitory changes make it harder for the axon to produce subsequent action potentials during this interval, which is called the **refractory period**. Thus, the refractory period limits the number of action potentials that a given nerve cell can produce per unit time. As might be expected, different types of neurons have different maximum rates of action potential firing due to different types and densities of ion channels. The refractoriness of the membrane in the wake of the action potential also explains why action potentials do not propagate back toward the point of their initiation as they travel along an axon.

Increased Conduction Velocity as a Result of Myelination

The rate of action potential conduction limits the flow of information within the nervous system. It is not surprising, then, that various mechanisms have evolved to optimize the propagation of action potentials along axons. Because action potential conduction requires passive and active flow of current (see Figure 3.12), the rate of action potential propagation is determined by both of these phenomena. One way of improving passive current flow is to increase the diameter of an axon, which effectively decreases the internal resistance to passive current flow (see Box C). The consequent increase in action potential conduction velocity presumably explains why giant axons evolved in invertebrates such as squid, and why rapidly conducting axons in all animals tend to be larger than slowly conducting ones.

Another strategy to improve the passive flow of electrical current is to insulate the axonal membrane, reducing the ability of current to leak out of the axon and thus increasing the distance along the axon that a given local current can flow passively (see Box C). This strategy is evident in the **myelination** of axons, a process by which oligodendrocytes in the central nervous system (and Schwann cells in the peripheral nervous system) wrap the axon in **myelin**, which consists of multiple layers of closely opposed glial membranes (Figure 3.13; see also Chapter 1). By acting as an electrical insulator, myelin greatly speeds up action potential conduction (Figure 3.14). For example, whereas unmyelinated axon conduction velocities range from about 0.5 to 10 m/s, myelinated axons can conduct at velocities of up to 150 m/s. The major reason underlying this marked increase in speed is that the time-consuming process of action potential generation occurs only at specific points along the axon, called **nodes of Ranvier**, where there is a gap in the myelin wrapping (see Figure 1.4F). If the entire surface of an axon were insulated, there would be no place for current to flow out of the axon and action potentials could not be generated. As it happens, an action potential generated at one node of Ranvier elicits current that flows passively within the myelinated segment until the next node is reached. This local current flow then generates an action potential in the neighboring segment, and the cycle is repeated along the length of the axon. Because current flows across the neuronal membrane only at the nodes (see Figure 3.13), this type of propagation is called **saltatory**, meaning that the action potential jumps from node to node. Not surprisingly, loss of myelin, as occurs in diseases such as multiple sclerosis, causes a variety of serious neurological problems (Box D).



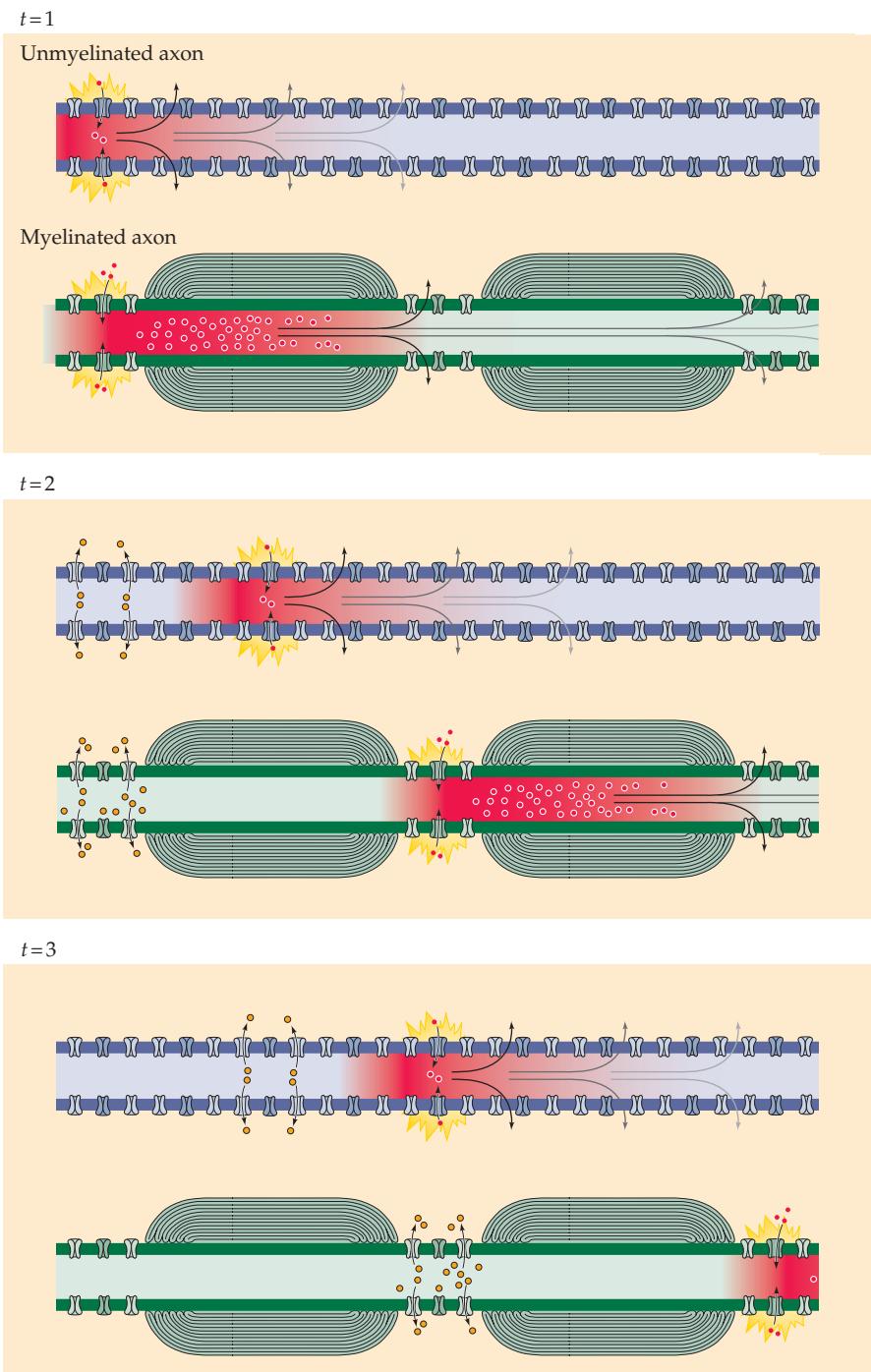


Figure 3.14 Comparison of speed of action potential conduction in unmyelinated (upper) and myelinated (lower) axons.

Summary

The action potential and all its complex properties can be explained by time- and voltage-dependent changes in the Na^+ and K^+ permeabilities of neuronal membranes. This conclusion derives primarily from evidence obtained by a device called the voltage clamp. The voltage clamp technique is an electronic feedback method that allows control of neuronal membrane potential

Box D

Multiple Sclerosis

Multiple sclerosis (MS) is a disease of the central nervous system characterized by a variety of clinical problems arising from multiple regions of demyelination and inflammation along axonal pathways. The disorder commonly begins between ages 20 and 40, characterized by the abrupt onset of neurological deficits that typically persist for days or weeks and then remit. The clinical course ranges from patients with no persistent neurological loss, some of whom experience only occasional later exacerbations, to others who progressively deteriorate as a result of extensive and relentless central nervous system involvement.

The signs and symptoms of MS are determined by the location of the affected regions. Particularly common are monocular blindness (due to lesions of the optic nerve), motor weakness or paralysis (due to lesions of the corticospinal tracts), abnormal somatic sensations (due to lesions of somatic sensory pathways, often in the posterior columns), double vision (due to lesions of medial longitudinal fasciculus), and dizziness (due to lesions of vestibular pathways). Abnormalities are often apparent in the cerebrospinal fluid, which usually contains an abnormal number of cells associated with inflammation and an increased content of antibodies (a sign of an altered immune response). The diagnosis of MS generally relies on the presence of a neurological problem that remits and then returns at an unrelated site. Confirmation can sometimes be obtained from magnetic resonance imaging (MRI), or functional evidence of lesions in a particular pathway by abnormal evoked potentials. The histological hallmark of MS at postmortem exam is multiple lesions at different sites showing loss of myelin associated with infiltration of inflammatory

cells and, in some instances, loss of axons themselves.

The concept of MS as a demyelinating disease is deeply embedded in the clinical literature, although precisely how the demyelination translates into functional deficits is poorly understood. The loss of the myelin sheath surrounding many axons clearly compromises action potential conduction, and the abnormal patterns of nerve conduction that result presumably produce most of the clinical deficits in the disease. However, MS may have effects that extend beyond loss of the myelin sheath. It is clear that some axons are actually destroyed, probably as a result of inflammatory processes in the overlying myelin and/or loss of trophic support of the axon by oligodendrocytes. Thus, axon loss also contributes to the functional deficits in MS, especially in the chronic, progressive forms of the disease.

The ultimate cause of MS remains unclear. The immune system undoubtedly contributes to the damage and new immunoregulatory therapies provide substantial benefits to many patients. Precisely how the immune system is activated to cause the injury is not known. The most popular hypothesis is that MS is an autoimmune disease (i.e., a disease in which the immune system attacks the body's proper constituents). The fact that immunization of experimental animals with any one of several molecular constituents of the myelin sheath can induce a demyelinating disease (called experimental allergic encephalomyelitis) shows that an autoimmune attack on the myelin membrane is sufficient to produce a picture similar to MS. A possible explanation of the human disease is that a genetically susceptible individual becomes transiently infected (by a minor viral illness, for example) with a microorganism that expresses a molecule struc-

turally similar to a component of myelin. An immune response to this antigen is mounted to attack the invader, but the failure of the immune system to discriminate between the foreign protein and self results in destruction of otherwise normal myelin, a scenario occurring in mice infected with Theiler's virus.

An alternative hypothesis is that MS is caused by a persistent infection by a virus or other microorganism. In this interpretation, the immune system's ongoing efforts to get rid of the pathogen cause the damage to myelin. Tropical spastic paraparesis (TSP) provides a precedent for this idea. TSP is a disease characterized by the gradual progression of weakness of the legs and impaired control of bladder function associated with increased deep tendon reflexes and a positive Babinski sign (see Chapter 16). This clinical picture is similar to that of rapidly advancing MS. TSP is known to be caused by persistent infection with a retrovirus (human T lymphotropic virus-1). This precedent notwithstanding, proving the persistent viral infection hypothesis for MS requires unambiguous demonstration of the presence of a virus. Despite periodic reports of a virus associated with MS, convincing evidence has not been forthcoming. In sum, MS remains a daunting clinical challenge.

References

- ADAMS, R. D. AND M. VICTOR (2001) *Principles of Neurology*, 7th Ed. New York: McGraw-Hill, pp. 954–982.
- MILLER, D. H. AND 9 OTHERS. (2003) A controlled trial of natalizumab for relapsing multiple sclerosis. *N. Engl. J. Med.* 348: 15–23.
- ZANVIL, S. S. AND L. STEINMAN (2003) Diverse targets for intervention during inflammatory and neurodegenerative phases of multiple sclerosis. *Neuron* 38: 685–688.

and, simultaneously, direct measurement of the voltage-dependent fluxes of Na^+ and K^+ that produce the action potential. Voltage clamp experiments show that a transient rise in Na^+ conductance activates rapidly and then inactivates during a sustained depolarization of the membrane potential. Such experiments also demonstrate a rise in K^+ conductance that activates in a delayed fashion and, in contrast to the Na^+ conductance, does not inactivate. Mathematical modeling of the properties of these conductances indicates that they, and they alone, are responsible for the production of all-or-none action potentials in the squid axon. Action potentials propagate along the nerve cell axons initiated by the voltage gradient between the active and inactive regions of the axon by virtue of the local current flow. In this way, action potentials compensate for the relatively poor passive electrical properties of nerve cells and enable neural signaling over long distances. These classical electrophysiological findings provide a solid basis for considering the functional and ultimately molecular variations on neural signaling taken up in the next chapter.

Additional Reading

Reviews

ARMSTRONG, C. M. AND B. HILLE (1998) Voltage-gated ion channels and electrical excitability. *Neuron* 20: 371–80.

NEHER, E. (1992) Ion channels for communication between and within cells. *Science* 256: 498–502.

Important Original Papers

ARMSTRONG, C. M. AND L. BINSTOCK (1965) Anomalous rectification in the squid giant axon injected with tetraethylammonium chloride. *J. Gen. Physiol.* 48: 859–872.

HODGKIN, A. L. AND A. F. HUXLEY (1952a) Currents carried by sodium and potassium ions through the membrane of the giant axon of *Loligo*. *J. Physiol.* 116: 449–472.

HODGKIN, A. L. AND A. F. HUXLEY (1952b) The components of membrane conductance in the giant axon of *Loligo*. *J. Physiol.* 116: 473–496.

HODGKIN, A. L. AND A. F. HUXLEY (1952c) The dual effect of membrane potential on sodium conductance in the giant axon of *Loligo*. *J. Physiol.* 116: 497–506.

HODGKIN, A. L. AND A. F. HUXLEY (1952d) A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 116: 507–544.

HODGKIN, A. L. AND W. A. H. RUSHTON (1938) The electrical constants of a crustacean nerve fibre. *Proc. R. Soc. Lond.* 133: 444–479.

HODGKIN, A. L., A. F. HUXLEY AND B. KATZ (1952) Measurements of current–voltage relations in the membrane of the giant axon of *Loligo*. *J. Physiol.* 116: 424–448.

MOORE, J. W., M. P. BLAUSTEIN, N. C. ANDERSON AND T. NARAHASHI (1967) Basis of tetrodotoxin's selectivity in blockage of squid axons. *J. Gen. Physiol.* 50: 1401–1411.

Books

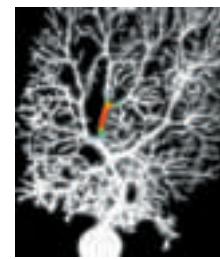
AIDLEY, D. J. AND P. R. STANFIELD (1996) *Ion Channels: Molecules in Action*. Cambridge: Cambridge University Press.

HILLE, B. (2001) *Ion Channels of Excitable Membranes*, 3rd Ed. Sunderland, MA: Sinauer Associates.

JOHNSTON, D. AND S. M.-S. WU (1995) *Foundations of Cellular Neurophysiology*. Cambridge, MA: MIT Press.

JUNGE, D. (1992) *Nerve and Muscle Excitation*, 3rd Ed. Sunderland, MA: Sinauer Associates.

Chapter 4



Channels and Transporters

Overview

The generation of electrical signals in neurons requires that plasma membranes establish concentration gradients for specific ions and that these membranes undergo rapid and selective changes in the membrane permeability to these ions. The membrane proteins that create and maintain ion gradients are called active transporters, whereas other proteins called ion channels give rise to selective ion permeability changes. As their name implies, ion channels are transmembrane proteins that contain a specialized structure, called a pore, that permits particular ions to cross the neuronal membrane. Some of these channels also contain other structures that are able to sense the electrical potential across the membrane. Such voltage-gated channels open or close in response to the magnitude of the membrane potential, allowing the membrane permeability to be regulated by changes in this potential. Other types of ion channels are gated by extracellular chemical signals such as neurotransmitters, and some by intracellular signals such as second messengers. Still others respond to mechanical stimuli, temperature changes, or a combination of such effects. Many types of ion channels have now been characterized at both the gene and protein level, resulting in the identification of a large number of ion channel subtypes that are expressed differentially in neuronal and non-neuronal cells. The specific expression pattern of ion channels in each cell type can generate a wide spectrum of electrical characteristics. In contrast to ion channels, active transporters are membrane proteins that produce and maintain ion concentration gradients. The most important of these is the Na^+ pump, which hydrolyzes ATP to regulate the intracellular concentrations of both Na^+ and K^+ . Other active transporters produce concentration gradients for the full range of physiologically important ions, including Cl^- , Ca^{2+} , and H^+ . From the perspective of electrical signaling, active transporters and ion channels are complementary: Transporters create the concentration gradients that help drive ion fluxes through open ion channels, thus generating electrical signals.

Ion Channels Underlying Action Potentials

Although Hodgkin and Huxley had no knowledge of the physical nature of the conductance mechanisms underlying action potentials, they nonetheless proposed that nerve cell membranes have channels that allow ions to pass selectively from one side of the membrane to the other (see Chapter 3). Based on the ionic conductances and currents measured in voltage clamp experiments, the postulated channels had to have several properties. First, because the ionic currents are quite large, the channels had to be capable of allowing ions to move across the membrane at high rates. Second, because

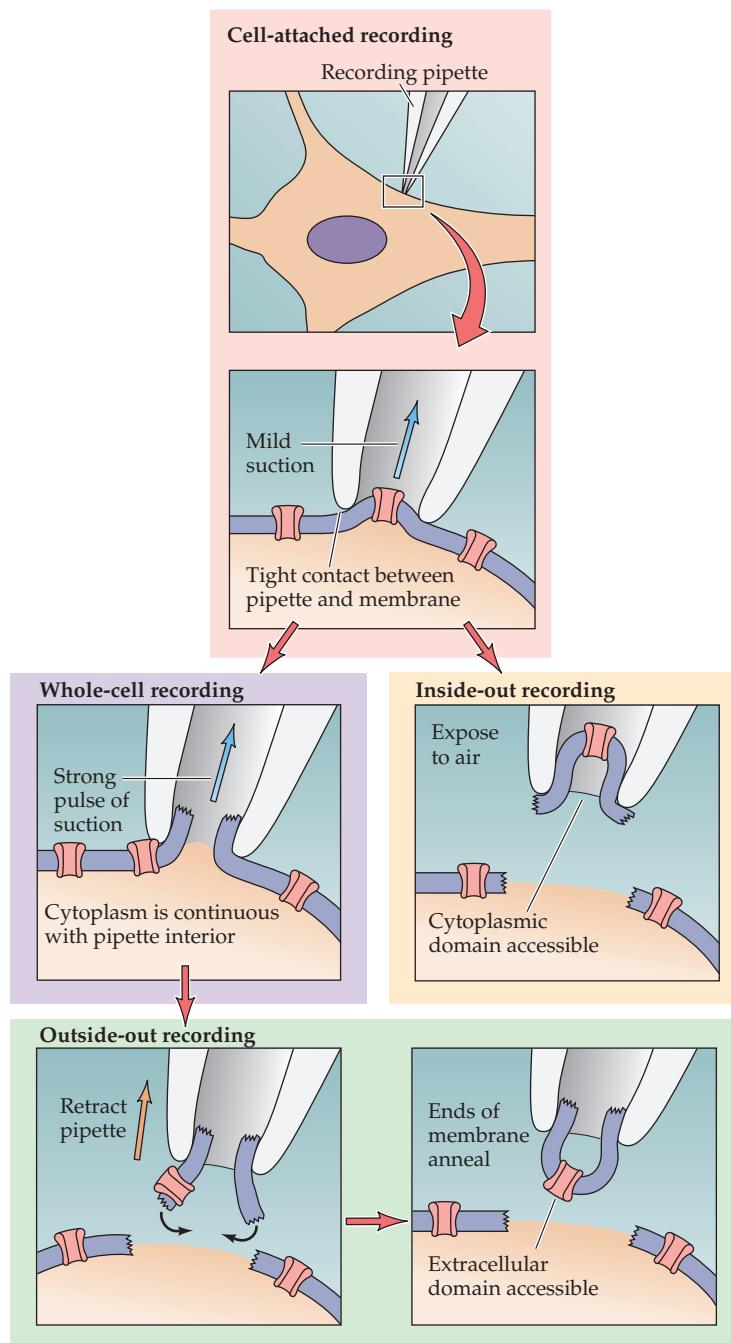
Box A

The Patch Clamp Method

A wealth of new information about ion channels resulted from the invention of the patch clamp method in the 1970s. This technique is based on a very simple idea. A glass pipette with a very small opening is used to make tight contact with a tiny area, or patch, of neuronal membrane. After the application of a small amount of suction to the back of the pipette, the seal between pipette and membrane becomes so tight that no ions can flow between the pipette and the membrane. Thus, all the ions that flow when a single ion channel opens must flow into the pipette. The resulting electrical current, though small, can be measured with an ultrasensitive electronic amplifier connected to the pipette. Based on the geometry involved, this arrangement usually is called the *cell-attached patch clamp recording method*. As with the conventional voltage clamp method, the patch clamp method allows experimental control of the membrane potential to characterize the voltage dependence of membrane currents.

Although the ability to record currents flowing through single ion channels is an important advantage of the cell-attached patch clamp method, minor technical modifications yield still other advantages. For example, if the membrane patch within the pipette is disrupted by briefly applying strong suction, the interior of the pipette becomes continuous with the cytoplasm of the cell. This arrangement allows measurements of electrical potentials and currents from the entire cell and is therefore called the *whole-cell recording method*. The whole-cell configuration also allows diffusional exchange between the pipette and the cytoplasm, producing a convenient way to inject substances into the interior of a “patched” cell.

Two other variants of the patch clamp method originate from the finding that once a tight seal has formed between the



Four configurations in patch clamp measurements of ionic currents.

membrane and the glass pipette, small pieces of membrane can be pulled away from the cell without disrupting the seal; this yields a preparation that is free of the complications imposed by the rest of the cell. Simply retracting a pipette that

is in the cell-attached configuration causes a small vesicle of membrane to remain attached to the pipette. By exposing the tip of the pipette to air, the vesicle opens to yield a small patch of membrane with its (former) intracellular sur-

face exposed. This arrangement, called the inside-out patch recording configuration, allows the measurement of single-channel currents with the added benefit of making it possible to change the medium to which the intracellular surface of the membrane is exposed. Thus, the inside-out configuration is particularly valuable when studying the influence of intracellular molecules on ion channel function. Alternatively, if the pipette is retracted while it is in the

whole-cell configuration, a membrane patch is produced that has its extracellular surface exposed. This arrangement, called the outside-out recording configuration, is optimal for studying how channel activity is influenced by extracellular chemical signals, such as neurotransmitters (see Chapter 5). This range of possible configurations makes the patch clamp method an unusually versatile technique for studies of ion channel function.

References

- HAMILL, O. P., A. MARTY, E. NEHER, B. SAKMANN AND F. J. SIGWORTH (1981) Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Arch.* 391: 85–100.
- LEOIS, R. A. AND J. L. RAE (1998) Low-noise patch-clamp techniques. *Meth. Enzym.* 293: 218–266.
- SIGWORTH, F. J. (1986) The patch clamp is more useful than anyone had expected. *Fed. Proc.* 45: 2673–2677.

the ionic currents depend on the electrochemical gradient across the membrane, the channels had to make use of these gradients. Third, because Na^+ and K^+ flow across the membrane independently of each other, different channel types had to be capable of discriminating between Na^+ and K^+ , allowing only one of these ions to flow across the membrane under the relevant conditions. Finally, given that the conductances are voltage-dependent, the channels had to be able to sense the voltage drop across the membrane, opening only when the voltage reached appropriate levels. While this concept of channels was highly speculative in the 1950s, later experimental work established beyond any doubt that transmembrane proteins called voltage-sensitive ion channels indeed exist and are responsible for all of the ionic conductance phenomena described in Chapter 3.

The first direct evidence for the presence of voltage-sensitive, ion-selective channels in nerve cell membranes came from measurements of the ionic currents flowing through individual ion channels. The voltage-clamp apparatus used by Hodgkin and Huxley could only resolve the *aggregate* current resulting from the flow of ions through many thousands of channels. A technique capable of measuring the currents flowing through single channels was devised in 1976 by Erwin Neher and Bert Sakmann at the Max Planck Institute in Goettingen. This remarkable approach, called patch clamping (Box A), revolutionized the study of membrane currents. In particular, the patch clamp method provided the means to test directly Hodgkin and Huxley's proposals about the characteristics of ion channels.

Currents flowing through Na^+ channels are best examined in experimental circumstances that prevent the flow of current through other types of channels that are present in the membrane (e.g., K^+ channels). Under such conditions, depolarizing a patch of membrane from a squid giant axon causes tiny inward currents to flow, but only occasionally (Figure 4.1). The size of these currents is minuscule—approximately 1–2 pA (i.e., 10^{-12} ampere), which is orders of magnitude smaller than the Na^+ currents measured by voltage clamping the entire axon. The currents flowing through single channels are called **microscopic currents** to distinguish them from the **macroscopic currents** flowing through a large number of channels distributed over a much more extensive region of surface membrane. Although microscopic currents are certainly small, a current of 1 pA nonetheless reflects the flow of thousands of ions per millisecond. Thus, as predicted, a single channel can let many ions pass through the membrane in a very short time.

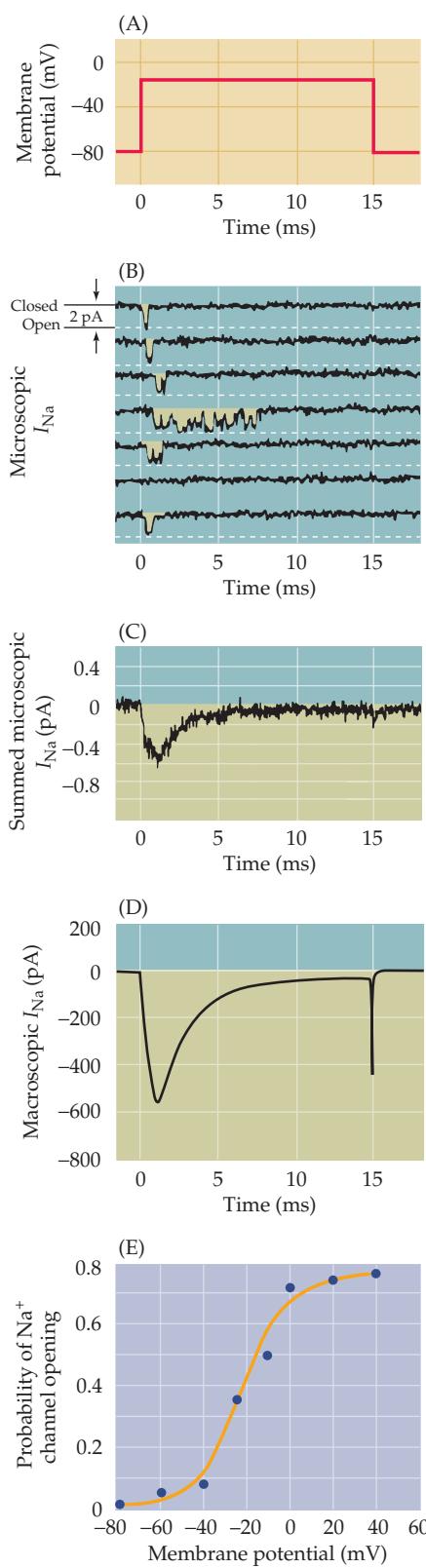


Figure 4.1 Patch clamp measurements of ionic currents flowing through single Na^+ channels in a squid giant axon. In these experiments, Cs^+ was applied to the axon to block voltage-gated K^+ channels. Depolarizing voltage pulses (A) applied to a patch of membrane containing a single Na^+ channel result in brief currents (B, downward deflections) in the seven successive recordings of membrane current (I_{Na}). (C) The sum of many such current records shows that most channels open in the initial 1–2 ms following depolarization of the membrane, after which the probability of channel openings diminishes because of channel inactivation. (D) A macroscopic current measured from another axon shows the close correlation between the time courses of microscopic and macroscopic Na^+ currents. (E) The probability of an Na^+ channel opening depends on the membrane potential, increasing as the membrane is depolarized. (B,C after Bezanilla and Correa, 1995; D after Vandenberg and Bezanilla, 1991; E after Correa and Bezanilla, 1994.)

Several observations further proved that the microscopic currents in Figure 4.1B are due to the opening of single, voltage-activated Na^+ channels. First, the currents are carried by Na^+ ; thus, they are directed inward when the membrane potential is more negative than $E_{Na'}$, reverse their polarity at $E_{Na'}$, are outward at more positive potentials, and are reduced in size when the Na^+ concentration of the external medium is decreased. This behavior exactly parallels that of the macroscopic Na^+ currents described in Chapter 3. Second, the channels have a time course of opening, closing, and inactivating that matches the kinetics of macroscopic Na^+ currents. This correspondence is difficult to appreciate in the measurement of microscopic currents flowing through a single open channel, because individual channels open and close in a stochastic (random) manner, as can be seen by examining the individual traces in Figure 4.1B. However, repeated depolarization of the membrane potential causes each Na^+ channel to open and close many times. When the current responses to a large number of such stimuli are averaged together, the collective response has a time course that looks much like the macroscopic Na^+ current (Figure 4.1C). In particular, the channels open mostly at the beginning of a prolonged depolarization, showing that they subsequently inactivate, as predicted from the macroscopic Na^+ current (compare Figures 4.1C and 4.1D). Third, both the opening and closing of the channels are voltage-dependent; thus, the channels are closed at -80 mV but open when the membrane potential is depolarized. In fact, the probability that any given channel will be open varies with membrane potential (Figure 4.1E), again as predicted from the macroscopic Na^+ conductance (see Figure 3.7). Finally, tetrodotoxin, which blocks the macroscopic Na^+ current (see Box C), also blocks microscopic Na^+ currents. Taken together, these results show that the macroscopic Na^+ current measured by Hodgkin and Huxley does indeed arise from the aggregate effect of many thousands of microscopic Na^+ currents, each representing the opening of a single voltage-sensitive Na^+ channel.

Patch clamp experiments have also revealed the properties of the channels responsible for the macroscopic K^+ currents associated with action potentials. When the membrane potential is depolarized (Figure 4.2A), microscopic outward currents (Figure 4.2B) can be observed under conditions that block Na^+ channels. The microscopic outward currents exhibit all the features expected for currents flowing through action-potential-related K^+ channels. Thus, the microscopic currents (Figure 4.2C), like their macroscopic counterparts (Figure 4.2D), fail to inactivate during brief depolarizations. Moreover, these single-channel currents are sensitive to ionic manipu-

Figure 4.2 Patch clamp measurements of ionic currents flowing through single K⁺ channels in a squid giant axon. In these experiments, tetrodotoxin was applied to the axon to block voltage-gated Na⁺ channels. Depolarizing voltage pulses (A) applied to a patch of membrane containing a single K⁺ channel results in brief currents (B, upward deflections) whenever the channel opens. (C) The sum of such current records shows that most channels open with a delay, but remain open for the duration of the depolarization. (D) A macroscopic current measured from another axon shows the correlation between the time courses of microscopic and macroscopic K⁺ currents. (E) The probability of a K⁺ channel opening depends on the membrane potential, increasing as the membrane is depolarized. (B and C after Augustine and Bezanilla, in Hille 1992; D after Augustine and Bezanilla, 1990; E after Perozo et al., 1991.)

lations and drugs that affect the macroscopic K⁺ currents and, like the macroscopic K⁺ currents, are voltage-dependent (Figure 4.2E). This and other evidence shows that macroscopic K⁺ currents associated with action potentials arise from the opening of many voltage-sensitive K⁺ channels.

In summary, patch clamping has allowed direct observation of microscopic ionic currents flowing through single ion channels, confirming that voltage sensitive Na⁺ and K⁺ channels are responsible for the macroscopic conductances and currents that underlie the action potential. Measurements of the behavior of single ion channels has also provided some insight into the molecular attributes of these channels. For example, single channel studies show that the membrane of the squid axon contains at least two types of channels—one selectively permeable to Na⁺ and a second selectively permeable to K⁺. Both channel types are **voltage-gated**, meaning that their opening is influenced by membrane potential (Figure 4.3). For each channel, depolarization increases the probability of channel opening, whereas hyperpolarization closes them (see Figures 4.1E and 4.2E). Thus, both channel types must have a **voltage sensor** that detects the potential across the membrane (Figure 4.3). However, these channels differ in important respects. In addition to their different ion selectivities, depolarization also inactivates the Na⁺ channel but not the K⁺ channel, causing Na⁺ channels to pass into a nonconducting state. The Na⁺ channel must therefore have an additional molecular mechanism responsible for **inactivation**. And, as expected from the macroscopic behavior of the Na⁺ and K⁺ currents described in Chapter 3, the kinetic properties of the gating of the two channels differs. This information about the physiology of single channels set the stage for subsequent studies of the molecular diversity of ion channels in various cell types, and of their detailed functional characteristics.

The Diversity of Ion Channels

Molecular genetic studies, in conjunction with the patch clamp method and other techniques, have led to many additional advances in understanding ion channels. Genes encoding Na⁺ and K⁺ channels, as well as many other channel types, have now been identified and cloned. A surprising fact that has emerged from these molecular studies is the diversity of genes that code for ion channels. Well over 100 ion channel genes have now been discovered, a number that could not have been anticipated from early studies of ion channel function. To understand the functional significance of this multitude of ion channel genes, the channels can be selectively expressed in well-

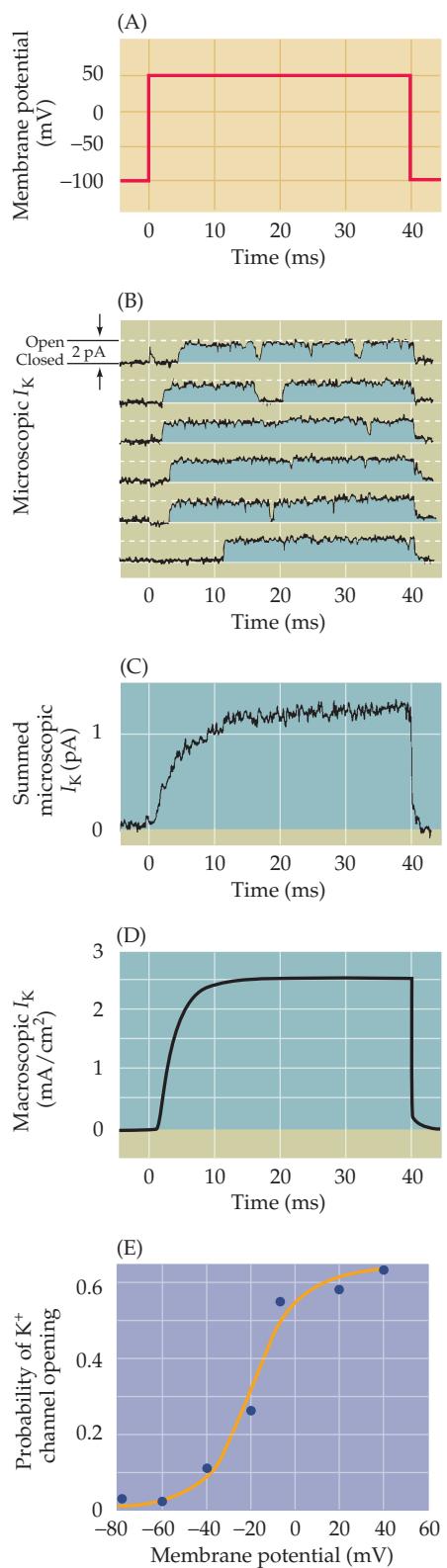
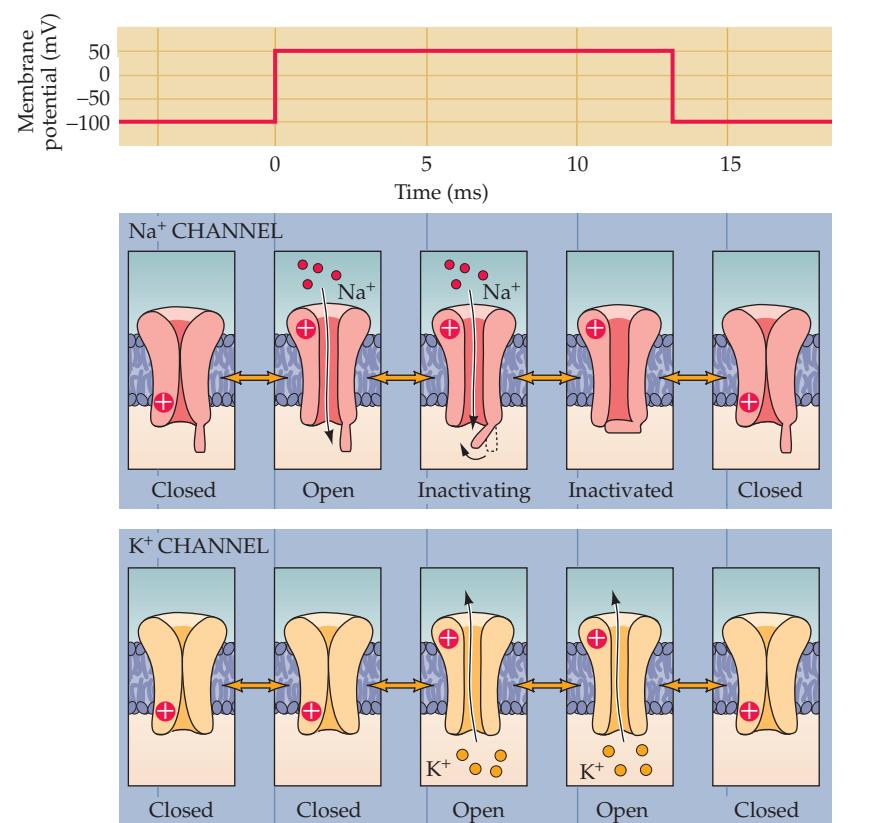


Figure 4.3 Functional states of voltage-gated Na^+ and K^+ channels. The gates of both channels are closed when the membrane potential is hyperpolarized. When the potential is depolarized, voltage sensors (indicated by +) allow the channel gates to open—first the Na^+ channels and then the K^+ channels. Na^+ channels also inactivate during prolonged depolarization, whereas many types of K^+ channels do not.



defined experimental systems, such as in cultured cells or frog oocytes (Box B), and then studied with patch clamping and other physiological techniques. Such studies have found many voltage-gated channels that respond to membrane potential in much the same way as the Na^+ and K^+ channels that underlie the action potential. Other channels, however, are gated by chemical signals that bind to extracellular or intracellular domains on these proteins and are insensitive to membrane voltage. Still others are sensitive to mechanical displacement, or to changes in temperature.

Further magnifying this diversity of ion channels are a number of mechanisms that can produce functionally different types of ion channels from a single gene. Ion channel genes contain a large number of coding regions that can be spliced together in different ways, giving rise to channel proteins that can have dramatically different functional properties. RNAs encoding ion channels also can be edited, modifying their base composition after transcription from the gene. For example, editing the RNA encoding of some receptors for the neurotransmitter glutamate (Chapter 6) changes a single amino acid within the receptor, which in turn gives rise to channels that differ in their selectivity for cations and in their conductance. Channel proteins can also undergo posttranslational modifications, such as phosphorylation by protein kinases (see Chapter 7), which can further change their functional characteristics. Thus, although the basic electrical signals of the nervous system are relatively stereotyped, the proteins responsible for generating these signals are remarkably diverse, conferring specialized signaling properties to many of the neuronal cell types that populate the nervous system. These channels also are involved in a broad range of neurological diseases.

Box B

Expression of Ion Channels in *Xenopus* Oocytes

Bridging the gap between the sequence of an ion channel gene and understanding channel function is a challenge. To meet this challenge, it is essential to have an experimental system in which the gene product can be expressed efficiently, and in which the function of the resulting channel can be studied with methods such as the patch clamp technique. Ideally, the vehicle for expression should be readily available, have few endogenous channels, and be large enough to permit mRNA and DNA to be microinjected with ease. Oocytes (immature eggs) from the clawed African frog, *Xenopus laevis* (Figure A), fulfill all these demands. These huge cells (approximately 1 mm in diameter; Figure B) are easily harvested from the female

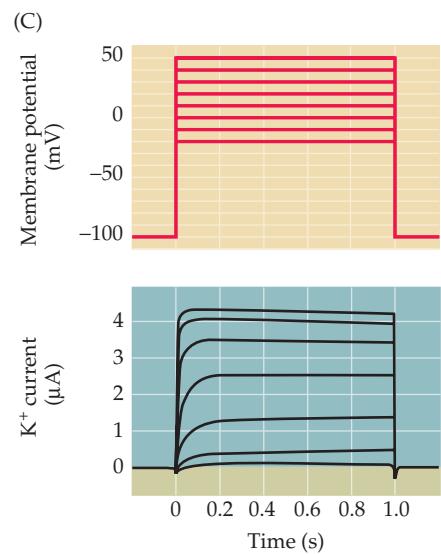
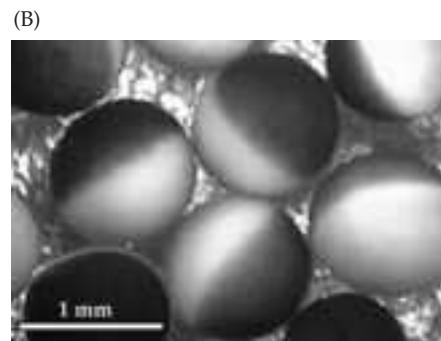
Xenopus. Work performed in the 1970s by John Gurdon, a developmental biologist, showed that injection of exogenous mRNA into frog oocytes causes them to synthesize foreign protein in prodigious quantities. In the early 1980s, Ricardo Miledi, Eric Barnard, and other neurobiologists demonstrated that *Xenopus* oocytes could express exogenous ion channels, and that physiological methods could be used to study the ionic currents generated by the newly-synthesized channels (Figure C).

As a result of these pioneering studies, heterologous expression experiments have now become a standard way of studying ion channels. The approach has been especially valuable in deciphering the relationship between channel structure and function. In such experiments, defined mutations (often affecting a single nucleotide) are made in the part of the channel gene that encodes a structure of interest; the resulting channel proteins are then expressed in oocytes to assess the functional consequences of the mutation.

The ability to combine molecular and physiological methods in a single cell system has made *Xenopus* oocytes a powerful experimental tool. Indeed, this system has been as valuable to contemporary studies of voltage-gated ion channels as the squid axon was to such studies in the 1950s and 1960s.

References

- GUNDERSEN, C. B., R. MILEDI AND I. PARKER (1984) Slowly inactivating potassium channels induced in *Xenopus* oocytes by messenger ribonucleic acid from *Torpedo* brain. *J. Physiol. (Lond.)* 353: 231–248.
 GURDON, J. B., C. D. LANE, H. R. WOODLAND AND G. MARBAIX (1971) Use of frog eggs and oocytes for the study of messenger RNA and its translation in living cells. *Nature* 233: 177–182.
 STÜHMER, W. (1998) Electrophysiological recordings from *Xenopus* oocytes. *Meth. Enzym.* 293: 280–300.
 SUMIKAWA, K., M. HOUGHTON, J. S. EMTAGE, B. M. RICHARDS AND E. A. BARNARD (1981) Active multi-subunit ACh receptor assembled by translation of heterologous mRNA in *Xenopus* oocytes. *Nature* 292: 862–864.



(A) The clawed African frog, *Xenopus laevis*. (B) Several oocytes from *Xenopus* highlighting the dark coloration of the original pole and the lighter coloration of the vegetal pole. (Courtesy of P. Reinhart.) (C) Results of a voltage clamp experiment showing K⁺ currents produced following injection of K⁺-channel mRNA into an oocyte. (After Gundersen et al., 1984.)

Voltage-Gated Ion Channels

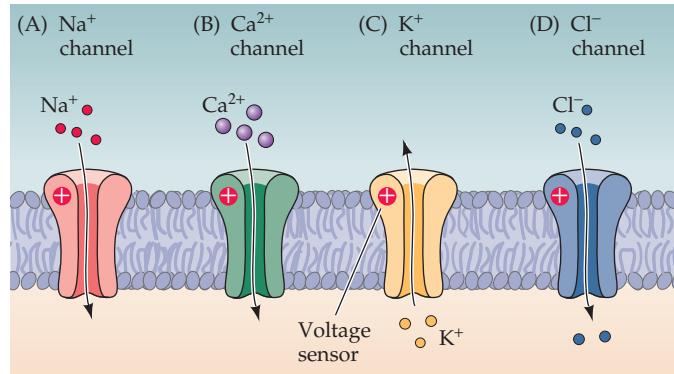
Voltage-gated ion channels that are selectively permeable to each of the major physiological ions— Na^+ , K^+ , Ca^{2+} , and Cl^- —have now been discovered (Figure 4.4 A–D). Indeed, many different genes have been discovered for each type of voltage-gated ion channel. An example is the identification of 10 human Na^+ channel genes. This finding was unexpected because Na^+ channels from many different cell types have similar functional properties, consistent with their origin from a single gene. It is now clear, however, that all of these Na^+ channel genes (called SCN genes) produce proteins that differ in their structure, function, and distribution in specific tissues. For instance, in addition to the rapidly inactivating Na^+ channels discovered by Hodgkin and Huxley in squid axon, a voltage-sensitive Na^+ channel that does *not* inactivate has been identified in mammalian axons. As might be expected, this channel gives rise to action potentials of long duration and is a target of local anesthetics such as benzocaine and lidocaine.

Other electrical responses in neurons entail the activation of voltage-gated Ca^{2+} channels (Figure 4.4B). In some neurons, voltage-gated Ca^{2+} channels give rise to action potentials in much the same way as voltage-sensitive Na^+ channels. In other neurons, Ca^{2+} channels control the shape of action potentials generated primarily by Na^+ conductance changes. More generally, by affecting intracellular Ca^{2+} concentrations, the activity of Ca^{2+} channels regulates an enormous range of biochemical processes within cells (see Chapter 7). Perhaps the most important of the processes regulated by voltage-sensitive Ca^{2+} channels is the release of neurotransmitters at synapses (see Chapter 5). Given these crucial functions, it is perhaps not surprising that 16 different Ca^{2+} channel genes (called CACNA genes) have been identified. Like Na^+ channels, Ca^{2+} channels differ in their activation and inactivation properties, allowing subtle variations in both electrical and chemical signaling processes mediated by Ca^{2+} . As a result, drugs that block voltage-gated Ca^{2+} channels are especially valuable in treating a variety of conditions ranging from heart disease to anxiety disorders.

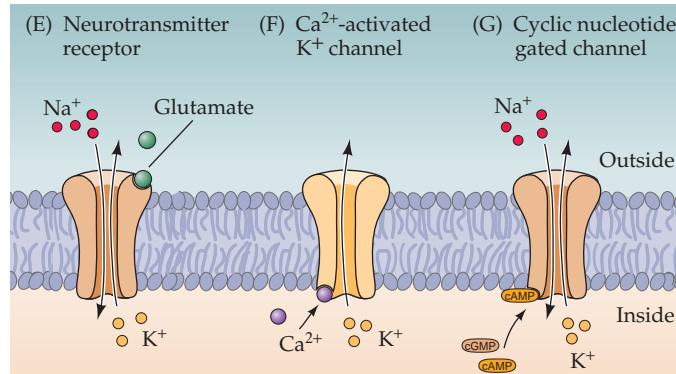
By far the largest and most diverse class of voltage-gated ion channels are the K^+ channels (Figure 4.4C). Nearly 100 K^+ channel genes are now known, and these fall into several distinct groups that differ substantially in their activation, gating, and inactivation properties. Some take minutes to inactivate, as in the case of squid axon K^+ channels studied by Hodgkin and Huxley (Figure 4.5A). Others inactivate within milliseconds, as is typical of most voltage-gated Na^+ channels (Figure 4.5B). These properties influence the

Figure 4.4 Types of voltage-gated ion channels. Examples of voltage-gated channels include those selectively permeable to Na^+ (A), Ca^{2+} (B), K^+ (C), and Cl^- (D). Ligand-gated ion channels include those activated by the extracellular presence of neurotransmitters, such as glutamate (E). Other ligand-gated channels are activated by intracellular second messengers, such as Ca^{2+} (F) or the cyclic nucleotides, cAMP and cGMP (G).

VOLTAGE-GATED CHANNELS



LIGAND-GATED CHANNELS



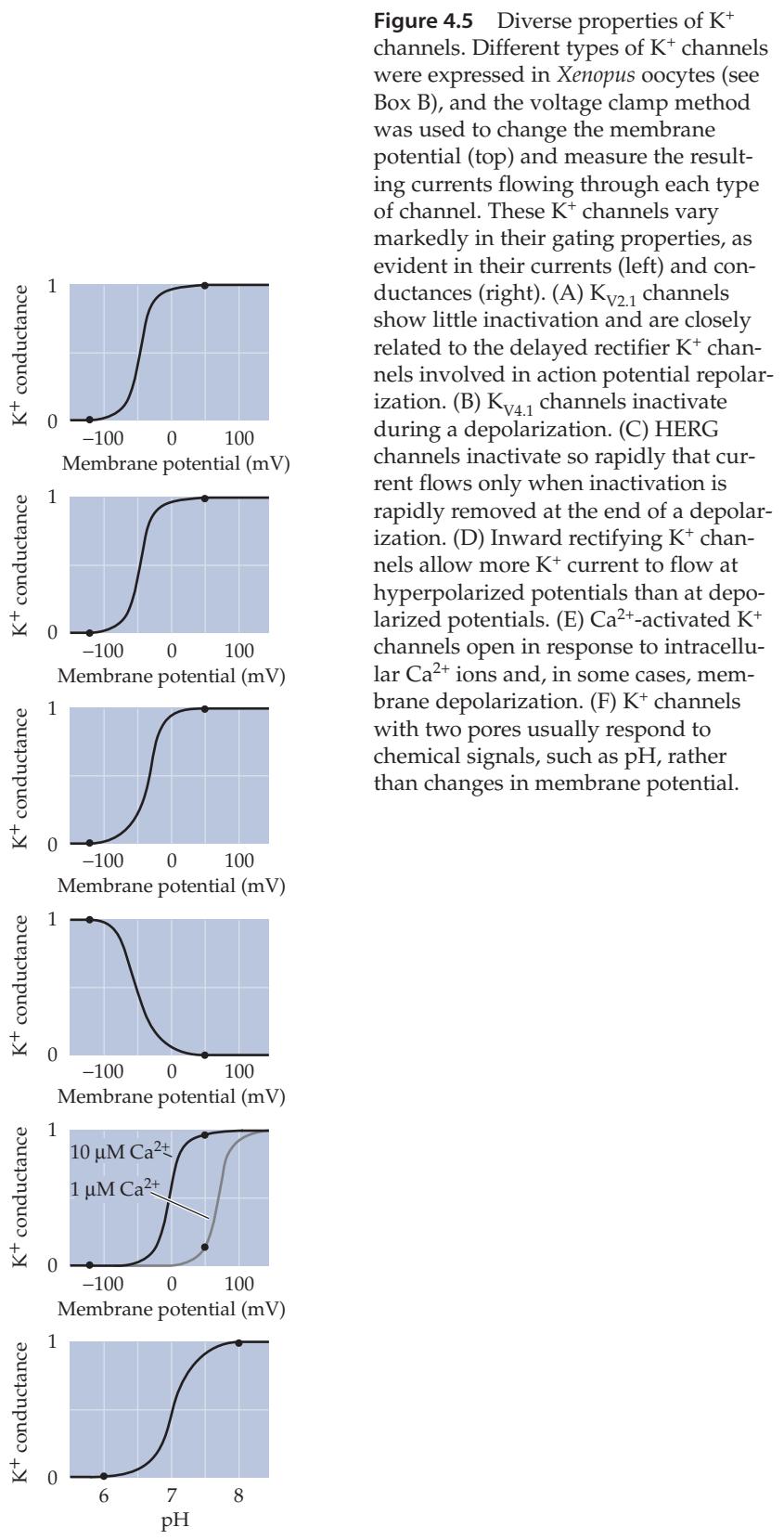
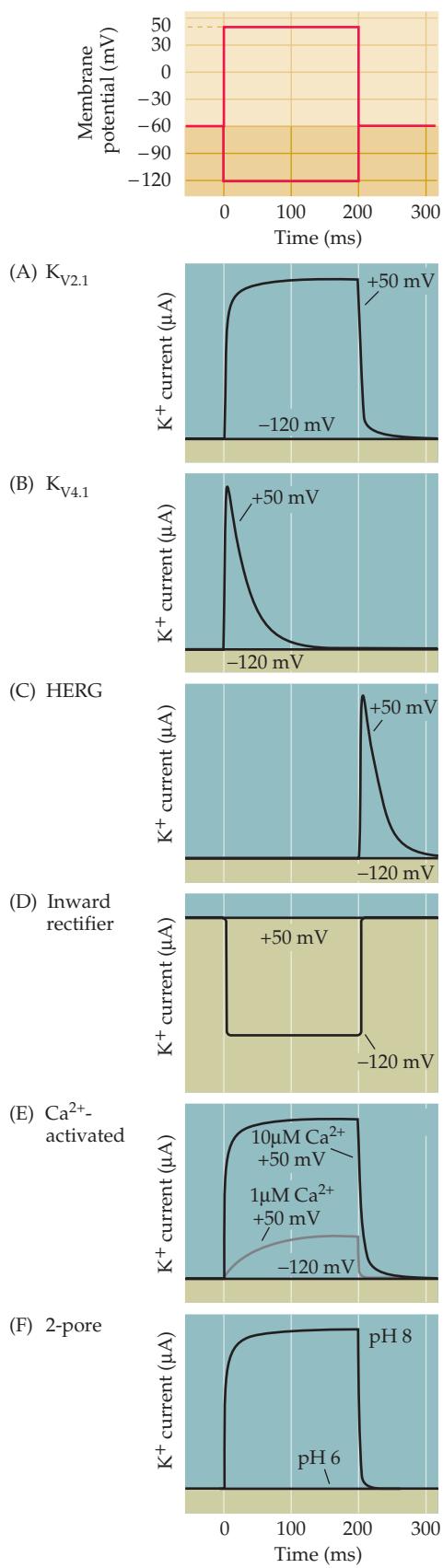


Figure 4.5 Diverse properties of K^+ channels. Different types of K^+ channels were expressed in *Xenopus* oocytes (see Box B), and the voltage clamp method was used to change the membrane potential (top) and measure the resulting currents flowing through each type of channel. These K^+ channels vary markedly in their gating properties, as evident in their currents (left) and conductances (right). (A) $K_{V2.1}$ channels show little inactivation and are closely related to the delayed rectifier K^+ channels involved in action potential repolarization. (B) $K_{V4.1}$ channels inactivate during a depolarization. (C) HERG channels inactivate so rapidly that current flows only when inactivation is rapidly removed at the end of a depolarization. (D) Inward rectifying K^+ channels allow more K^+ current to flow at hyperpolarized potentials than at depolarized potentials. (E) Ca^{2+} -activated K^+ channels open in response to intracellular Ca^{2+} ions and, in some cases, membrane depolarization. (F) K^+ channels with two pores usually respond to chemical signals, such as pH, rather than changes in membrane potential.

duration and rate of action potential firing, with important consequences for axonal conduction and synaptic transmission. Perhaps the most important function of K⁺ channels is the role they play in generating the resting membrane potential (see Chapter 2). At least two families of K⁺ channels that are open at substantially negative membrane voltage levels contribute to setting the resting membrane potential (Figure 4.5D).

Finally, several types of voltage-gated Cl⁻ channel have been identified (see Figure 4.4D). These channels are present in every type of neuron, where they control excitability, contribute to the resting membrane potential, and help regulate cell volume.

Ligand-Gated Ion Channels

Many types of ion channels respond to chemical signals (ligands) rather than to changes in the membrane potential (Figure 4.4E–G). The most important of these **ligand-gated ion channels** in the nervous system is the class activated by binding neurotransmitters (Figure 4.4E). These channels are essential for synaptic transmission and other forms of cell-cell signaling phenomena discussed in Chapters 5–7. Whereas the voltage-gated ion channels underlying the action potential typically allow only one type of ion to permeate, channels activated by extracellular ligands are usually less selective, allowing two or more types of ions to pass through the channel pore.

Other ligand-gated channels are sensitive to chemical signals arising within the cytoplasm of neurons (see Chapter 7), and can be selective for specific ions such as K⁺ or Cl⁻, or permeable to all physiological cations. Such channels are distinguished by ligand-binding domains on their *intracellular* surfaces that interact with second messengers such as Ca²⁺, the cyclic nucleotides cAMP and cGMP, or protons. Examples of channels that respond to intracellular cues include Ca²⁺-activated K⁺ channels (Figure 4.4F), the cyclic nucleotide gated cation channel (Figure 4.4G), or acid-sensing ion channels (ASICs). The main function of these channels is to convert intracellular chemical signals into electrical information. This process is particularly important in sensory transduction, where channels gated by cyclic nucleotides convert odors and light, for example, into electrical signals. Although many of these ligand-gated ion channels are located in the cell surface membrane, others are in membranes of intracellular organelles such as mitochondria or the endoplasmic reticulum. Some of these latter channels are selectively permeable to Ca²⁺ and regulate the release of Ca²⁺ from the lumen of the endoplasmic reticulum into the cytoplasm, where this second messenger can then trigger a spectrum of cellular responses such as described in Chapter 7.

Stretch- and Heat-Activated Channels

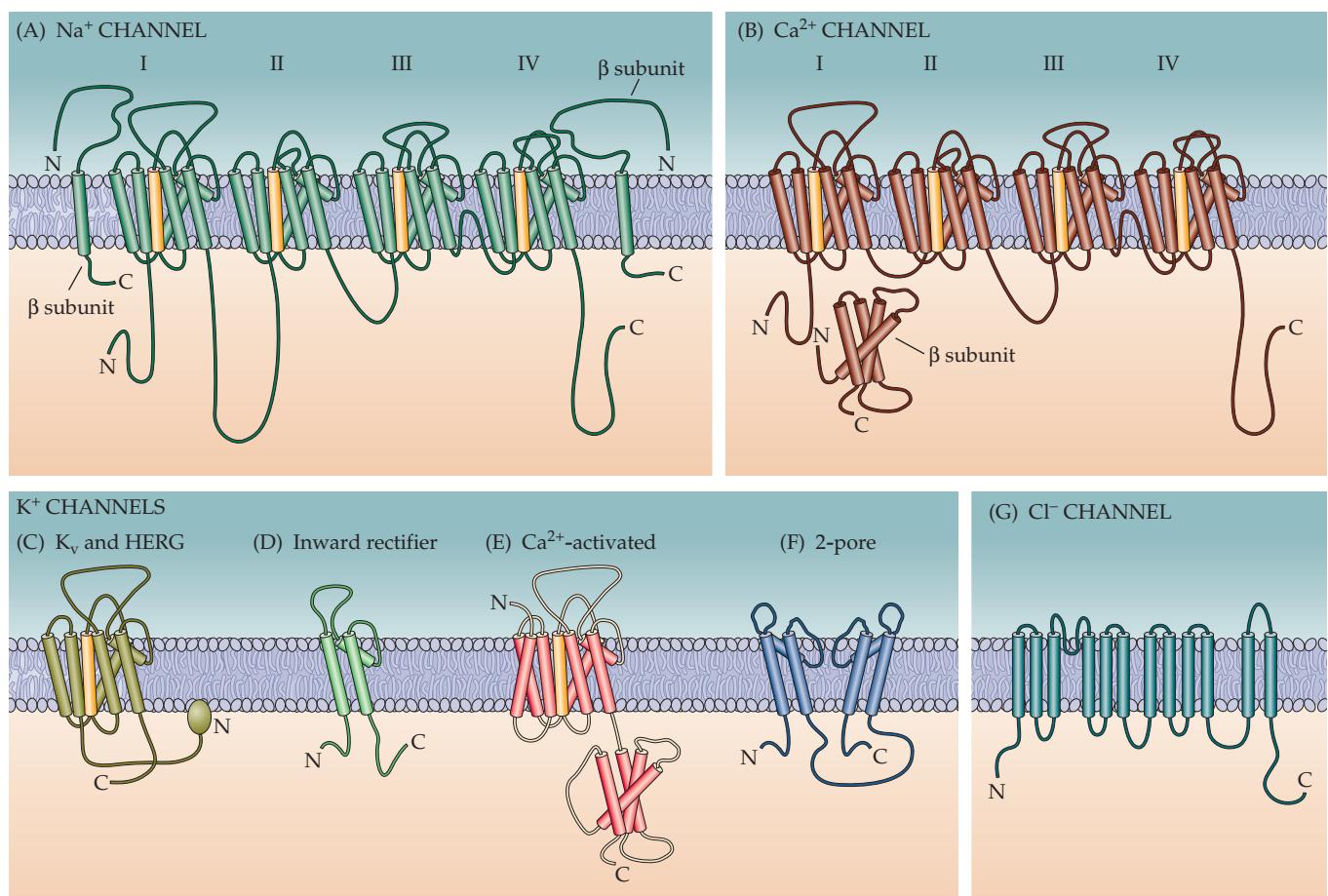
Still other ion channels respond to heat or membrane deformation. Heat-activated ion channels, such as some members of the transient receptor potential (TRP) gene family, contribute to the sensations of pain and temperature and help mediate inflammation (see Chapter 9). These channels are often specialized to detect specific temperature ranges, and some are even activated by cold. Other ion channels respond to mechanical distortion of the plasma membrane and are the basis of stretch receptors and neuromuscular stretch reflexes (see Chapters 8, 15 and 16). A specialized form of these channels enables hearing by allowing auditory hair cells to respond to sound waves (see Chapter 12).

In summary, this tremendous variety of ion channels allows neurons to generate electrical signals in response to changes in membrane potential, synaptic input, intracellular second messengers, light, odors, heat, sound, touch, and many other stimuli.

The Molecular Structure of Ion Channels

Understanding the physical structure of ion channels is obviously the key to sorting out how they actually work. Until recently, most information about channel structure was derived indirectly from studies of the amino acid composition and physiological properties of these proteins. For example, a great deal has been learned by exploring the functions of particular amino acids within the proteins using **mutagenesis** and the expression of such channels in *Xenopus* oocytes (see Box B). Such studies have discovered a general transmembrane architecture common to all the major ion channel families. Thus, these molecules are all integral membrane proteins that span the plasma membrane repeatedly. Na^+ (and Ca^{2+}) channel proteins, consist of repeating motifs of 6 membrane-spanning regions that are repeated 4 times, for a total of 24 transmembrane regions (Figure 4.6A,B). Na^+ (or Ca^{2+}) channels can be produced by just one of these proteins, although other accessory proteins, called β subunits, can regulate the function of these channels. K^+ channel proteins typically span the membrane six times (Figure 4.6C),

Figure 4.6 Topology of the principal subunits of voltage-gated Na^+ , Ca^{2+} , K^+ , and Cl^- channels. Repeating motifs of Na^+ (A) and Ca^{2+} (B) channels are labeled I, II, III, and IV; (C–F) K^+ channels are more diverse. In all cases, four subunits combine to form a functional channel. (G) Chloride channels are structurally distinct from all other voltage-gated channels.



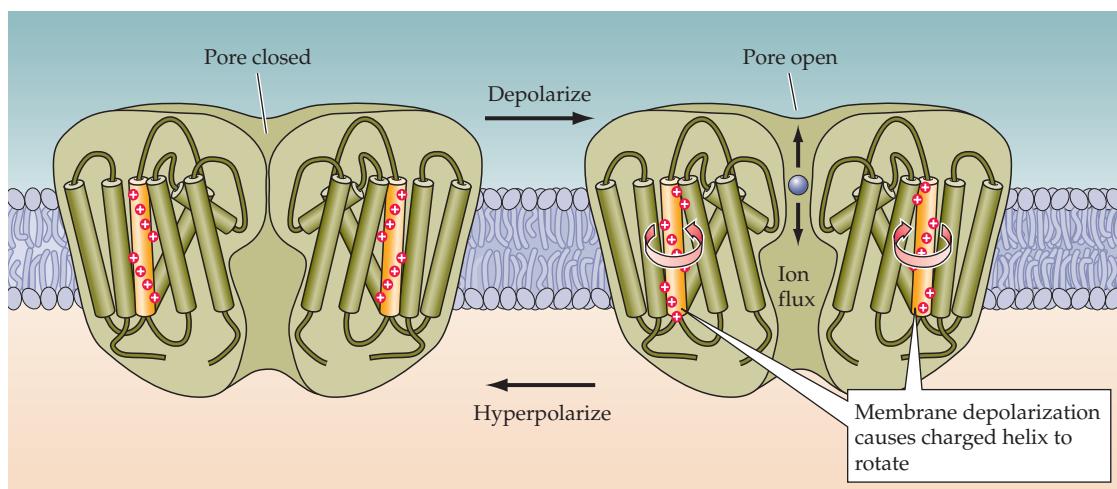
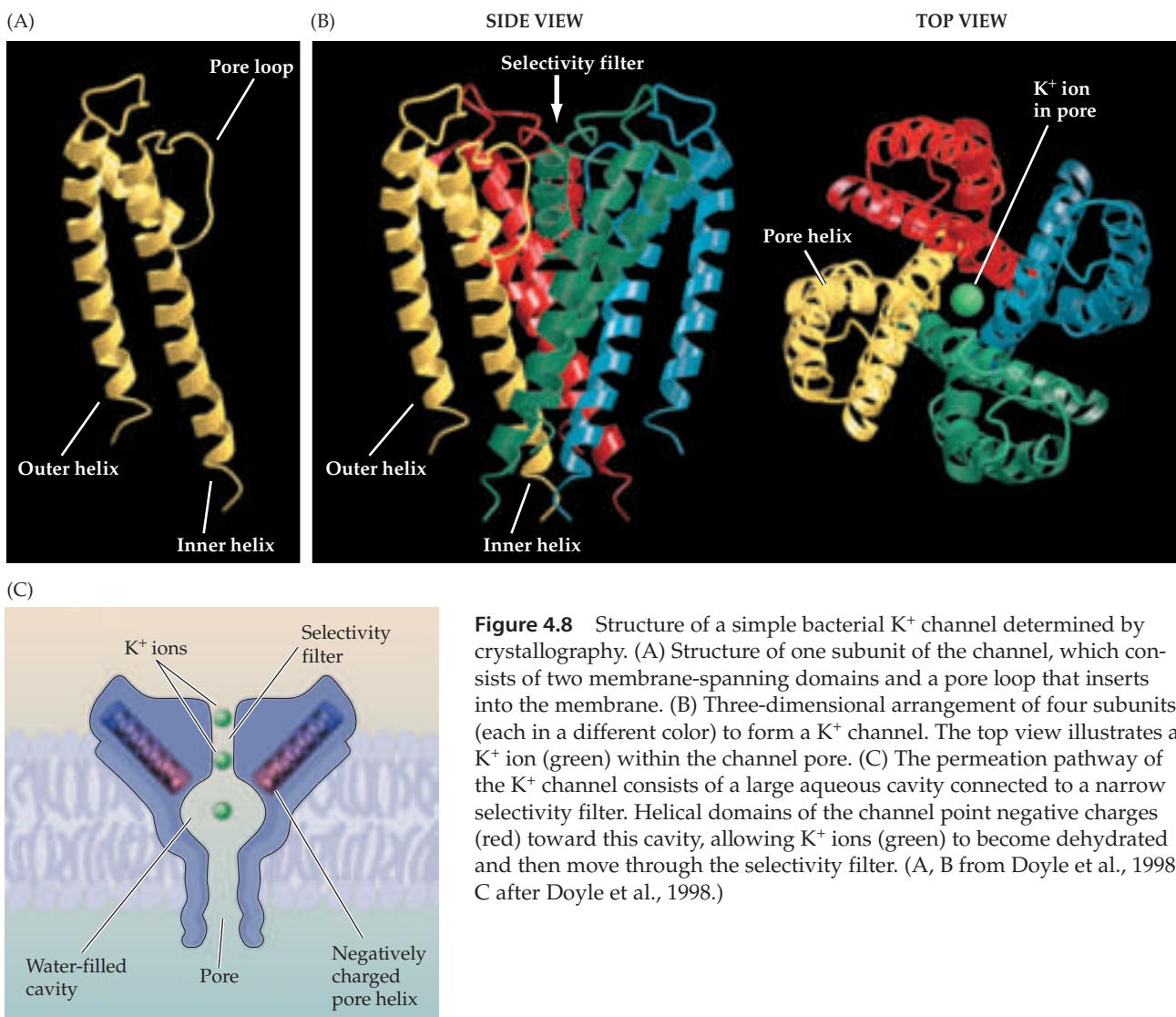


Figure 4.7 A charged voltage sensor permits voltage-dependent gating of ion channels. The process of voltage activation may involve the rotation of a positively charged transmembrane domain. This movement causes a change in the conformation of the pore loop, enabling the channel to conduct specific ions.

though there are some K^+ channels, such as a bacterial channel and some mammalian channels, that span the membrane only twice (Figure 4.6D), and others that span the membrane four times (Figure 4.6F) or seven times (Figure 4.6E). Each of these K^+ channel proteins serves as a channel subunit, with 4 of these subunits typically aggregating to form a single functional ion channel.

Other imaginative mutagenesis experiments have provided information about how these proteins function. Two membrane-spanning domains of all ion channels appear to form a central **pore** through which ions can diffuse, and one of these domains contains a protein loop that confers an ability to selectively allow certain ions to diffuse through the channel pore (Figure 4.7). As might be expected, the amino acid composition of the pore loop differs among channels that conduct different ions. These distinct structural features of channel proteins also provide unique binding sites for drugs and for various neurotoxins known to block specific subclasses of ion channels (Box C). Furthermore, many voltage gated ion channels contain a distinct type of transmembrane helix containing a number of positively charged amino acids along one face of the helix (Figures 4.6 and 4.7). This structure evidently serves as a sensor that detects changes in the electrical potential across the membrane. Membrane depolarization influences the charged amino acids such that the helix undergoes a conformational change, which in turn allows the channel pore to open. One suggestion is that the helix rotates to cause the pore to open (Figure 4.7). Other types of mutagenesis experiments have demonstrated that one end of certain K^+ channels plays a key role in channel inactivation. This intracellular structure (labeled "N" in Figure 4.6C) can plug the channel pore during prolonged depolarization.

More recently, very direct information about the structural underpinnings of ion channel function has come from **X-ray crystallography** studies of bacterial K^+ channels (Figure 4.8). This molecule was chosen for analysis because the large quantity of channel protein needed for crystallography could be obtained by growing large numbers of bacteria expressing this molecule. The results of such studies showed that the channel is formed by subunits that each cross the plasma membrane twice; between these two membrane-spanning structures is a loop that inserts into the plasma membrane (Figure 4.8A). Four of these subunits are assembled together to form a chan-



nel (Figure 4.8B). In the center of the assembled channel is a narrow opening through the protein that allows K^+ to flow across the membrane. This opening is the channel pore and is formed by the protein loop, as well as by the membrane-spanning domains. The structure of the pore is well suited for conducting K^+ ions (Figure 4.8C). The narrowest part is near the outside mouth of the channel and is so constricted that only a non-hydrated K^+ ion can fit through the bottleneck. Larger cations, such as Cs^+ , cannot traverse this region of the pore, and smaller cations such as Na^+ cannot enter the pore because the “walls” of the pore are too far apart to stabilize a dehydrated Na^+ ion. This part of the channel complex is responsible for the selective permeability to K^+ and is therefore called the **selectivity filter**. The sequence of amino acids making up part of this selectivity filter is often referred to as the K^+ channel “signature sequence”. Deeper within the channel is a water-filled cavity that connects to the interior of the cell. This cavity evidently collects K^+ from the cytoplasm and, utilizing negative charges from the protein,

Box C

Toxins That Poison Ion Channels

Given the importance of Na^+ and K^+ channels for neuronal excitation, it is not surprising that a number of organisms have evolved channel-specific toxins as mechanisms for self-defense or for capturing prey. A rich collection of natural toxins selectively target the ion channels of neurons and other cells. These toxins are valuable not only for survival, but for studying the function of cellular ion channels. The best-known channel toxin is *tetrodotoxin*, which is produced by certain puffer fish and other animals.

Tetrodotoxin produces a potent and specific obstruction of the Na^+ channels responsible for action potential generation, thereby paralyzing the animals unfortunate enough to ingest it.

Saxitoxin, a chemical homologue of tetrodotoxin produced by dinoflagellates, has a similar action on Na^+ channels. The potentially lethal effects of eating shellfish that have ingested these “red tide” dinoflagellates are due to the potent neuronal actions of saxitoxin.

Scorpions paralyze their prey by injecting a potent mix of peptide toxins that also affect ion channels. Among these are the α -toxins, which slow the inactivation of Na^+ channels (Figure A1); exposure of neurons to these toxins prolongs the action potential (Figure A2),

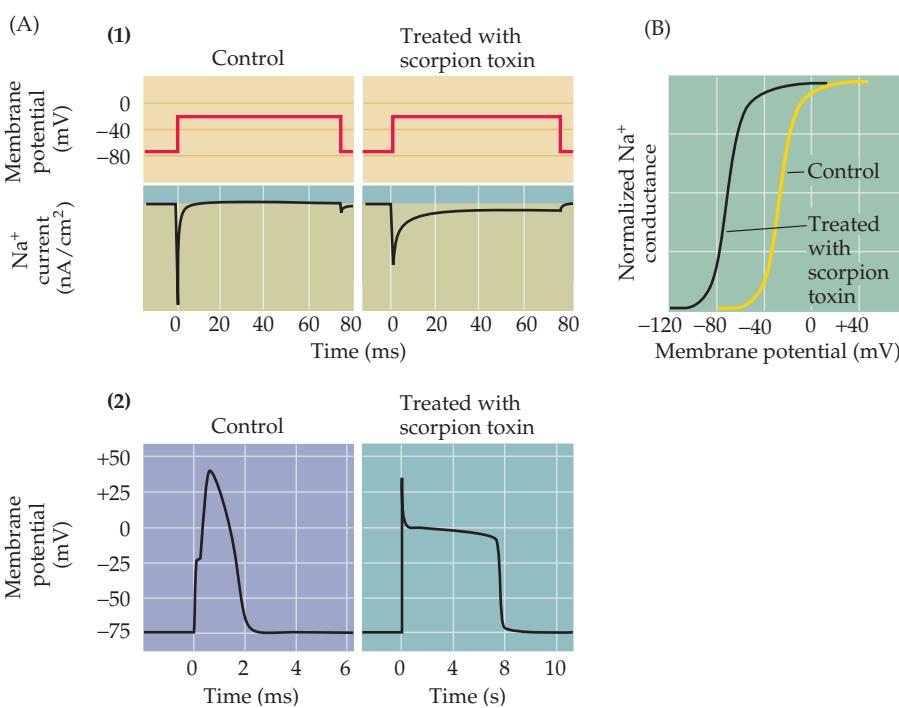
thereby scrambling information flow within the nervous system of the soon-to-be-devoured victim. Other peptides in scorpion venom, called β -toxins, shift the voltage dependence of Na^+ channel activation (Figure B). These toxins cause Na^+ channels to open at potentials much more negative than normal, disrupting action potential generation. Some alkaloid toxins combine these actions, both removing inactivation and shifting activation of Na^+ channels. One such toxin is *batrachotoxin*, produced by a species of frog; some tribes of South American Indians use this poison on their arrow tips. A number of plants produce similar toxins, including *aconitine*, from buttercups; *veratridine*, from lilies; and a number of insecticidal toxins produced by plants such as chrysanthemums and rhododendrons.

Potassium channels have also been targeted by toxin-producing organisms.

Peptide toxins affecting K^+ channels include *dendrotoxin*, from wasps; *apamin*, from bees; and *charybdotoxin*, yet another toxin produced by scorpions. All of these toxins block K^+ channels as their primary action; no toxin is known to affect the activation or inactivation of these channels, although such agents may simply be awaiting discovery.

References

- CAHALAN, M. (1975) Modification of sodium channel gating in frog myelinated nerve fibers by *Centruroides sculpturatus* scorpion venom. *J. Physiol. (Lond.)* 244: 511–534.
- NARAHASHI, T. (2000) Neuroreceptors and ion channels as the basis for drug action: Present and future. *J. Pharmacol. Exptl. Therapeutics* 294: 1–26.
- SCHMIDT, O. AND H. SCHMIDT (1972) Influence of calcium ions on the ionic currents of nodes of Ranvier treated with scorpion venom. *Pflügers Arch. 333: 51–61.*



(A) Effects of toxin treatment on frog axons. (1) α -Toxin from the scorpion *Leiurus quinquestriatus* prolongs Na^+ currents recorded with the voltage clamp method. (2) As a result of the increased Na^+ current, α -toxin greatly prolongs the duration of the axonal action potential. Note the change in timescale after treating with toxin. (B) Treatment of a frog axon with β -toxin from another scorpion, *Centruroides sculpturatus*, shifts the activation of Na^+ channels, so that Na^+ conductance begins to increase at potentials much more negative than usual. (A after Schmidt and Schmidt, 1972; B after Cahalan, 1975.)

allows K^+ ions to become dehydrated so they can enter the selectivity filter. These “naked” ions are then able to move through four K^+ binding sites within the selectivity filter to eventually reach the extracellular space (recall that the normal concentration gradient drives K^+ out of cells). On average, two K^+ ions reside within the selectivity filter at any moment, with electrostatic repulsion between the two ions helping to speed their transit through the selectivity filter, thereby permitting rapid ion flux through the channel.

Crystallographic studies have also determined the structure of the **voltage sensor** in another type of bacterial K^+ channel. Such studies indicate that the sensor is at the interface between proteins and lipid on the cytoplasmic surface of the channel, leading to the suggestion that the sensor is a paddle-like structure that moves through the membrane to gate the opening of the channel pore (Figure 4.9A), rather than being a rotating helix buried within the ion channel protein (as in Figure 4.7). Crystallographic work has also revealed the molecular basis of the rapid transitions between the closed and the open state of the channel during channel gating. By comparing data from K^+ channels crystallized in what is believed to be closed and open conformations (Figure 4.9B), it appears that channels gate by a conformational change in one of the transmembrane helices lining the channel pore. Producing a “kink” in one of these helices increases the opening from the central water-filled pore to the intracellular space, thereby permitting ion fluxes.

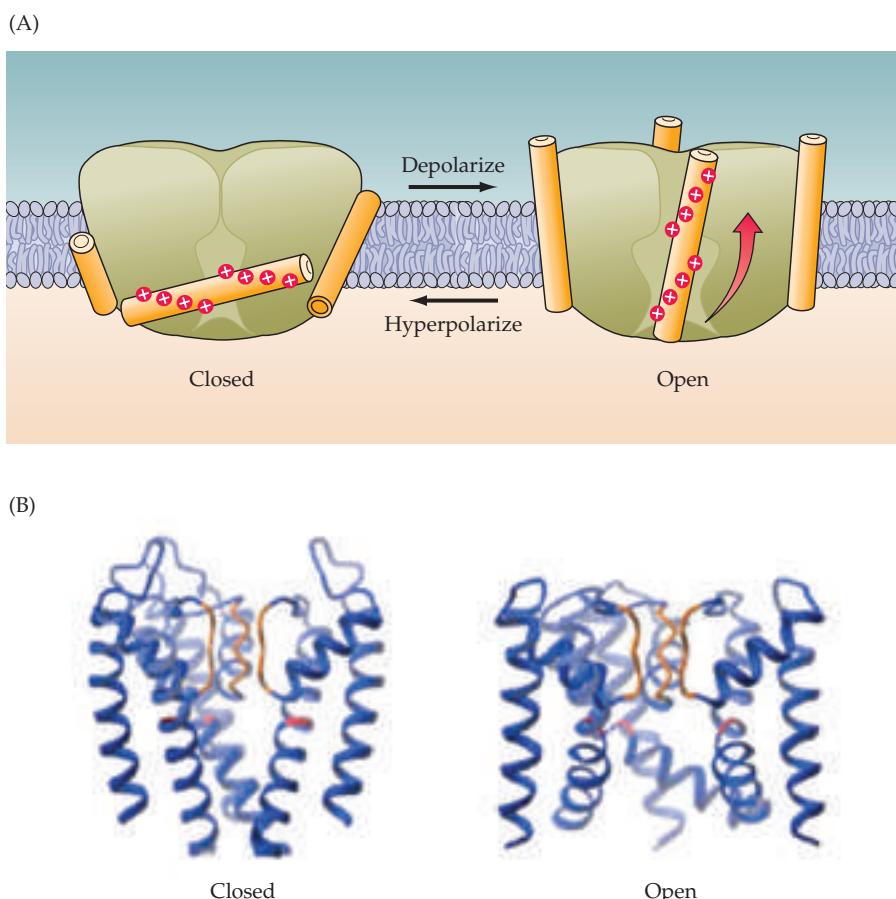


Figure 4.9 Structural features of K^+ channel gating. (A) Voltage sensing may involve paddle-like structures of the channel. These paddles reside within the lipid bilayer of the plasma membrane and may respond to changes in membrane potential by moving through the membrane. The gating charges that sense membrane potential are indicated by red “plus” signs. (B) Structure of K^+ channels in closed (left) and open (right) conformations. Three of the four channel subunits are shown. Opening of the pore of the channel involves kinking of a transmembrane domain at the point indicated in red, which then dilates the pore. (A after Jiang et al., 2003; B after MacKinnon, 2003).

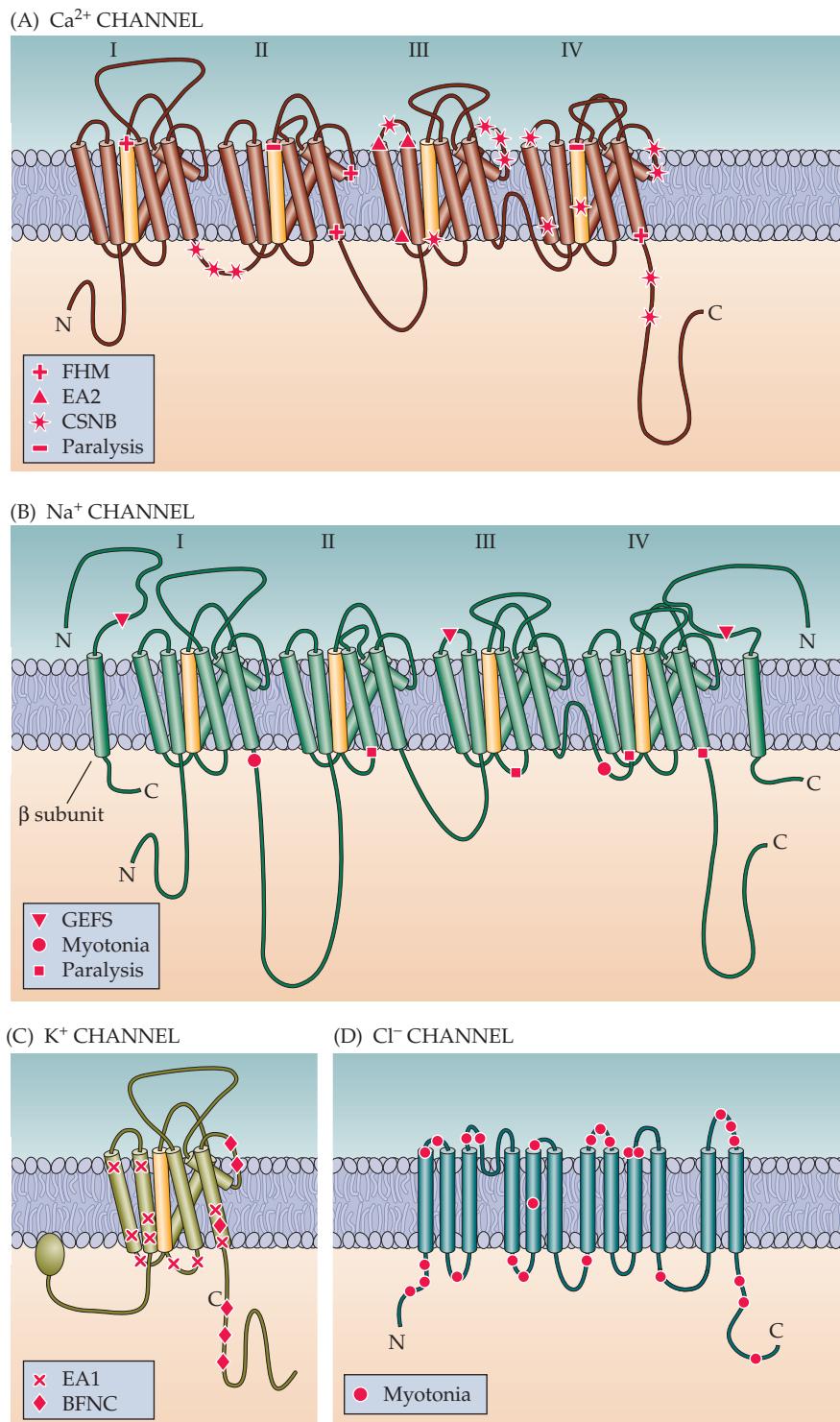
Box D

Diseases Caused by Altered Ion Channels

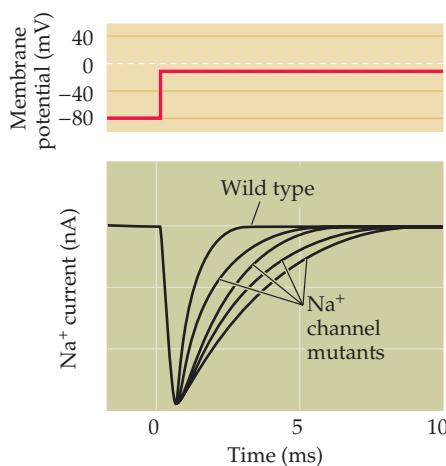
Several genetic diseases, collectively called *channelopathies*, result from small but critical alterations in ion channel genes. The best-characterized of these diseases are those that affect skeletal muscle cells. In these disorders, alterations in ion channel proteins produce either myotonia (muscle stiffness due to excessive electrical excitability) or paralysis (due to insufficient muscle excitability). Other disorders arise from ion channel defects in heart, kidney, and the inner ear.

Channelopathies associated with ion channels localized in brain are much more difficult to study. Nonetheless, voltage-gated Ca^{2+} channels have recently been implicated in a range of neurological diseases. These include episodic ataxia, spinocerebellar degeneration, night blindness, and migraine headaches. *Familial hemiplegic migraine* (FHM) is characterized by migraine attacks that typically last one to three days. During such episodes, patients experience severe headaches and vomiting. Several mutations in a human Ca^{2+} channel have been identified in families with FHM, each having different clinical symptoms. For example, a mutation in the pore-forming region of the channel produces hemiplegic migraine with progressive cerebellar ataxia, whereas other mutations cause only the usual FHM symptoms. How these altered Ca^{2+} channel properties lead to migraine attacks is not known.

Episodic ataxia type 2 (EA2) is a neurological disorder in which affected individuals suffer recurrent attacks of abnormal limb movements and severe ataxia. These problems are sometimes accompa-



Genetic mutations in (A) Ca^{2+} channels, (B) Na^{+} channels, (C) K^{+} channels, and (D) Cl^{-} channels that result in diseases. Red regions indicate the sites of these mutations; the red circles indicate mutations. (After Lehmann-Horn and Jurkat-Kott, 1999.)



Mutations in Na^+ channels slow the rate of inactivation of Na^+ currents. (After Barchi, 1995.)

nied by vertigo, nausea, and headache. Usually, attacks are precipitated by emotional stress, exercise, or alcohol and last for a few hours. The mutations in EA2 cause Ca^{2+} channels to be truncated at various sites, which may cause the clinical manifestations of the disease by preventing the normal assembly of Ca^{2+} channels in the membrane.

X-linked *congenital stationary night blindness* (CSNB) is a recessive retinal disorder that causes night blindness, decreased visual acuity, myopia, nystagmus, and strabismus. Complete CSNB causes retinal rod photoreceptors to be nonfunctional. Incomplete CSNB causes subnormal (but measurable) functioning

of both rod and cone photoreceptors. Like EA2, the incomplete type of CSNB is caused by mutations producing truncated Ca^{2+} channels. Abnormal retinal function may arise from decreased Ca^{2+} currents and neurotransmitter release from photoreceptors (see Chapter 11).

A defect in brain Na^+ channels causes *generalized epilepsy with febrile seizures* (GEFS) that begins in infancy and usually continues through early puberty. This defect has been mapped to two mutations: one on chromosome 2 that encodes an α subunit for a voltage-gated Na^+ channel, and the other on chromosome 19 that encodes a Na^+ channel β subunit. These mutations cause a slowing of Na^+ channel inactivation (see figure above), which may explain the neuronal hyperexcitability underlying GEFS.

Another type of seizure, *benign familial neonatal convulsion* (BFNC), is due to K^+ channel mutations. This disease is characterized by frequent brief seizures commencing within the first week of life and disappearing spontaneously within a few months. The mutation has been mapped to at least two voltage-gated K^+ channel genes. A reduction in K^+ current flow through the mutated channels probably accounts for the hyperexcitability associated with this defect. A related disease, episodic ataxia type 1 (EA1), has been linked to a defect in another type of voltage-gated K^+ channel. EA1 is characterized by brief episodes of ataxia. Mu-

tant channels inhibit the function of other, non-mutant K^+ channels and may produce clinical symptoms by impairing action potential repolarization. Mutations in the K^+ channels of cardiac muscle are responsible for the irregular heartbeat of patients with long Q-T syndrome. Numerous genetic disorders affect the voltage-gated channels of skeletal muscle and are responsible for a host of muscle diseases that either cause muscle weakness (*paralysis*) or muscle contraction (*myotonia*).

References

- BARCHI, R. L. (1995) Molecular pathology of the skeletal muscle sodium channel. *Ann. Rev. Physiol.* 57: 355–385.
- BERKOVIC, S. F. AND I. E. SCHEFFER (1997) Epilepsies with single gene inheritance. *Brain Develop.* 19: 13–28.
- COOPER, E. C. AND L. Y. JAN (1999) Ion channel genes and human neurological disease: Recent progress, prospects, and challenges. *Proc. Natl. Acad. Sci. USA* 96: 4759–4766.
- DAVIES, N. P. AND M. G. HANNA (1999) Neurological channelopathies: Diagnosis and therapy in the new millennium. *Ann. Med.* 31: 406–420.
- JEN, J. (1999) Calcium channelopathies in the central nervous system. *Curr. Op. Neurobiol.* 9: 274–280.
- LEHMANN-HORN, F. AND K. JURKAT-ROTT (1999) Voltage-gated ion channels and hereditary disease. *Physiol. Rev.* 79: 1317–1372.
- OPHOFF, R. A., G. M. TERWINDT, R. R. FRANTS AND M. D. FERRARI (1998) P/Q-type Ca^{2+} channel defects in migraine, ataxia and epilepsy. *Trends Pharm. Sci.* 19: 121–127.

In short, ion channels are integral membrane proteins with characteristic features that allow them to assemble into multimolecular aggregates. Collectively, these structures allow channels to conduct ions, sense the transmembrane potential, to inactivate, and to bind to various neurotoxins. A combination of physiological, molecular biological and crystallographic studies has begun to provide a detailed physical picture of K^+ channels. This work has now provided considerable insight into how ions are conducted from one side of the plasma membrane to the other, how a channel can be selectively permeable to a single type of ion, how they are able to sense changes in membrane voltage, and how they gate the opening of their pores. It is likely that other types of ion channels will be similar in their functional architecture. Finally, this sort of work has illuminated how mutations in ion channel genes can lead to a variety of neurological disorders (Box D).

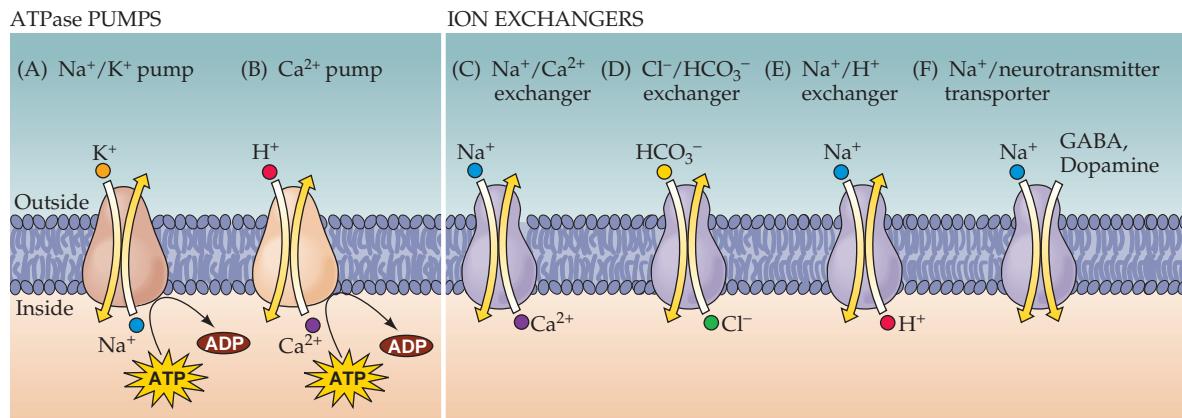
Active Transporters Create and Maintain Ion Gradients

Up to this point, the discussion of the molecular basis of electrical signaling has taken for granted the fact that nerve cells maintain ion concentration gradients across their surface membranes. However, none of the ions of physiological importance (Na^+ , K^+ , Cl^- , and Ca^{2+}) are in electrochemical equilibrium. Because channels produce electrical effects by allowing one or more of these ions to diffuse down their electrochemical gradients, there would be a gradual dissipation of these concentration gradients unless nerve cells could restore ions displaced during the current flow that occurs as a result of both neural signaling and the continual ionic leakage that occurs at rest. The work of generating and maintaining ionic concentration gradients for particular ions is carried out by a group of plasma membrane proteins known as **active transporters**.

Active transporters carry out this task by forming complexes with the ions that they are translocating. The process of ion binding and unbinding for transport typically requires several milliseconds. As a result, ion translocation by active transporters is much slower than ion movement through channels: Recall that ion channels can conduct thousands of ions across a membrane each millisecond. In short, active transporters gradually store energy in the form of ion concentration gradients, whereas the opening of ion channels rapidly dissipates this stored energy during relatively brief electrical signaling events.

Several types of active transporter have now been identified (Figure 4.10). Although the specific jobs of these transporters differ, all must translocate ions against their electrochemical gradients. Moving ions uphill requires the consumption of energy, and neuronal transporters fall into two classes based on their energy sources. Some transporters acquire energy directly from the hydrolysis of ATP and are called **ATPase pumps** (Figure 4.10, left). The most prominent example of an ATPase pump is the Na^+ pump (or, more properly, the Na^+/K^+ ATPase pump), which is responsible for maintaining transmembrane concentration gradients for both Na^+ and K^+ (Figure 4.10A). Another is the Ca^{2+} pump, which provides one of the main mechanisms for removing Ca^{2+} from cells (Figure 4.10B). The second class of active transporter does not use ATP directly, but depends instead on the electrochemical gradients of other ions as an energy source. This type of transporter carries one or more ions *up* its electrochemical gradient while simultaneously taking another ion (most often Na^+) *down* its gradient. Because at least two species of ions are

Figure 4.10 Examples of ion transporters found in cell membranes. (A,B) Some transporters are powered by the hydrolysis of ATP (ATPase pumps), whereas others (C–F) use the electrochemical gradients of co-transported ions as a source of energy (ion exchangers).



involved in such transactions, these transporters are usually called **ion exchangers** (Figure 4.10, right). An example of such a transporter is the $\text{Na}^+/\text{Ca}^{2+}$ exchanger, which shares with the Ca^{2+} pump the important job of keeping intracellular Ca^{2+} concentrations low (Figure 4.10C). Another exchanger in this category regulates both intracellular Cl^- concentration and pH by swapping intracellular Cl^- for another extracellular anion, bicarbonate (Figure 4.10D). Other ion exchangers, such as the Na^+/H^+ exchanger (Figure 4.10E), also regulate intracellular pH, in this case by acting directly on the concentration of H^+ . Yet other ion exchangers are involved in transporting neurotransmitters into synaptic terminals (Figure 4.10F), as described in Chapter 6. Although the electrochemical gradient of Na^+ (or other counter ions) is the proximate source of energy for ion exchangers, these gradients ultimately depend on the hydrolysis of ATP by ATPase pumps, such as the Na^+/K^+ ATPase pump.

Functional Properties of the Na^+/K^+ Pump

Of these various transporters, the best understood is the Na^+/K^+ pump. The activity of this pump is estimated to account for 20–40% of the brain's energy consumption, indicating its importance for brain function. The Na^+ pump was first discovered in neurons in the 1950s, when Richard Keynes at Cambridge University used radioactive Na^+ to demonstrate the energy-dependent efflux of Na^+ from squid giant axons. Keynes and his collaborators found that this efflux ceased when the supply of ATP in the axon was interrupted by treatment with metabolic poisons (Figure 4.11A, point 4). Other conditions that lower intracellular ATP also prevent Na^+ efflux. These experiments showed that removing intracellular Na^+ requires cellular metabolism. Further studies with radioactive K^+ demonstrated that Na^+ efflux is associated with simultaneous, ATP-dependent influx of K^+ . These opposing fluxes of Na^+ and K^+ are operationally inseparable: Removal of external K^+ greatly reduces Na^+ efflux (Figure 4.11, point 2), and vice versa. These energy-dependent movements of Na^+ and K^+ implicated an ATP-hydrolyzing Na^+/K^+ pump in the generation of the transmembrane gradients of both Na^+ and K^+ . The exact mechanism responsible for these fluxes of Na^+ and K^+ is still not entirely clear, but the pump is thought to alternately shuttle these ions across the membranes in a cycle fueled by the transfer of a phosphate group from ATP to the pump protein (Figure 4.11B).

Additional quantitative studies of the movements of Na^+ and K^+ indicate that the two ions are not pumped at identical rates: The K^+ influx is only about two-thirds the Na^+ efflux. Thus, the pump apparently transports two K^+ into the cell for every three Na^+ that are removed (see Figure 4.11B). This stoichiometry causes a net loss of one positively charged ion from inside of the cell during each round of pumping, meaning that the pump generates an electrical current that can hyperpolarize the membrane potential. For this reason, the Na^+/K^+ pump is said to be **electrogenic**. Because pumps act much more slowly than ion channels, the current produced by the Na^+/K^+ pump is quite small. For example, in the squid axon, the net current generated by the pump is less than 1% of the current flowing through voltage-gated Na^+ channels and affects the resting membrane potential by only a millivolt or less.

Although the electrical current generated by the activity of the Na^+/K^+ pump is small, under special circumstances the pump can significantly influence the membrane potential. For instance, prolonged stimulation of

Figure 4.11 Ionic movements due to the Na^+/K^+ pump. (A) Measurement of radioactive Na^+ efflux from a squid giant axon. This efflux depends on external K^+ and intracellular ATP. (B) A model for the movement of ions by the Na^+/K^+ pump. Uphill movements of Na^+ and K^+ are driven by ATP, which phosphorylates the pump. These fluxes are asymmetrical, with three Na^+ carried out for every two K^+ brought in. (A after Hodgkin and Keynes, 1955; B after Lingrel et al., 1994.)

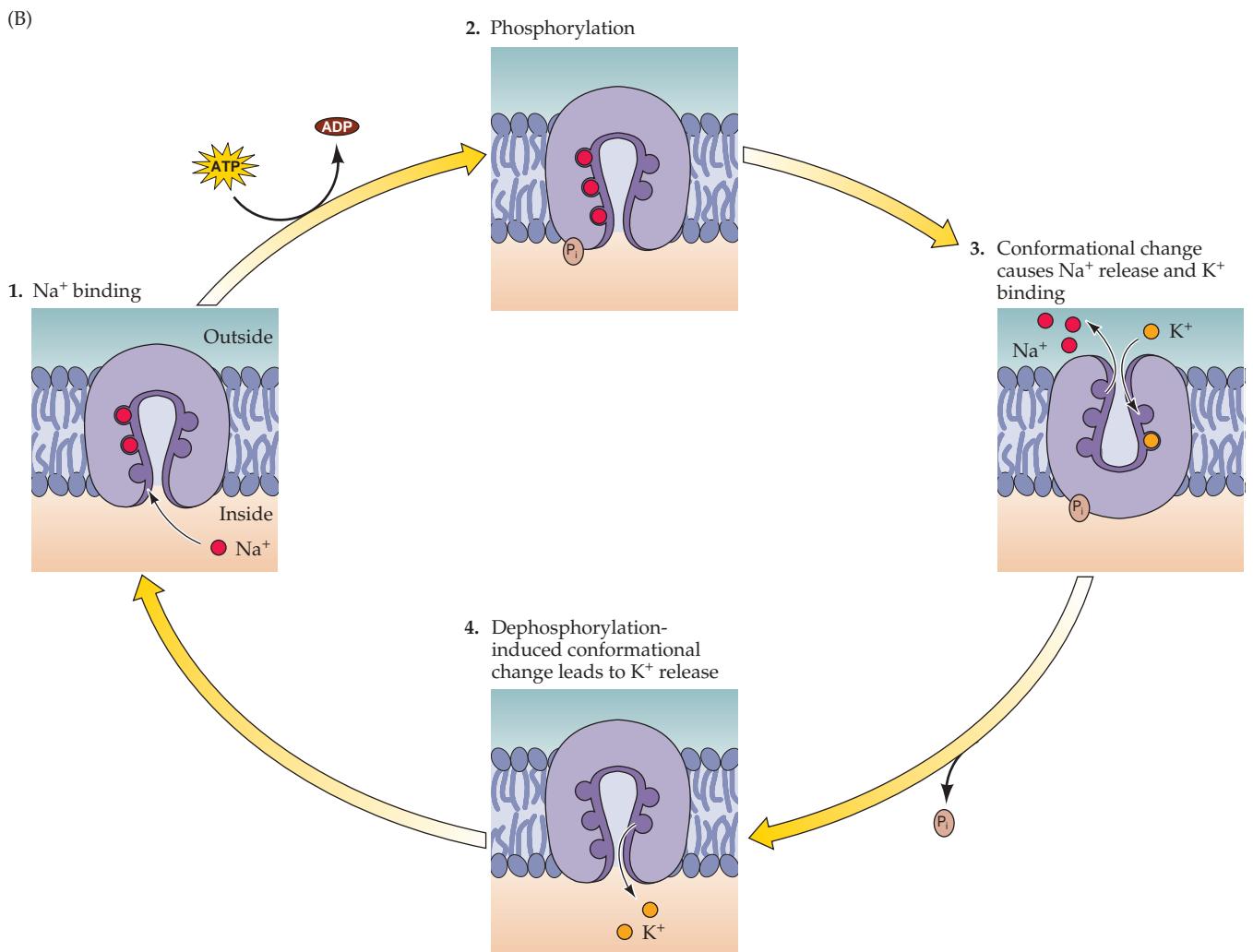
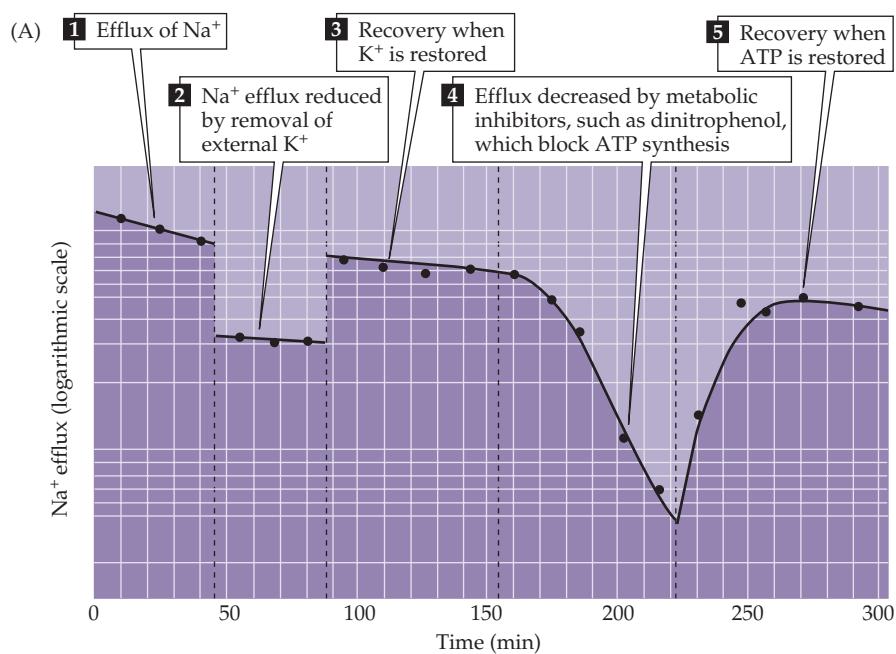


Figure 4.12 The electrogenic transport of ions by the Na^+/K^+ pump can influence membrane potential. Measurements of the membrane potential of a small unmyelinated axon show that a train of action potentials is followed by a long-lasting hyperpolarization. This hyperpolarization is blocked by ouabain, indicating that it results from the activity of the Na^+/K^+ pump. (After Rang and Ritchie, 1968.)

small unmyelinated axons produces a substantial hyperpolarization (Figure 4.12). During the period of stimulation, Na^+ enters through voltage-gated channels and accumulates within the axons. As the pump removes this extra Na^+ , the resulting current generates a long-lasting hyperpolarization. Support for this interpretation comes from the observation that conditions that block the Na^+/K^+ pump—for example, treatment with ouabain, a plant glycoside that specifically inhibits the pump—prevent the hyperpolarization. The electrical contribution of the Na^+/K^+ pump is particularly significant in these small-diameter axons because their large surface-to-volume ratio causes intracellular Na^+ concentration to rise to higher levels than it would in other cells. Nonetheless, it is important to emphasize that, in most circumstances, the Na^+/K^+ pump plays no part in generating the action potential and has very little *direct* effect on the resting potential.

The Molecular Structure of the Na^+/K^+ Pump

These observations imply that the Na^+ and K^+ pump must exhibit several molecular properties: (1) It must bind both Na^+ and K^+ ; (2) it must possess sites that bind ATP and receive a phosphate group from this ATP; and (3) it must bind ouabain, the toxin that blocks this pump (Figure 4.13A). A variety of studies have now identified the aspects of the protein that account for these properties of the Na^+/K^+ pump. This pump is a large, integral membrane protein made up of at least two subunits, called α and β . The primary sequence shows that the α subunit spans the membrane 10 times, with most of the molecule found on the cytoplasmic side, whereas the β subunit spans the membrane once and is predominantly extracellular. Although a detailed account of the functional domains of the Na^+/K^+ pump is not yet available, some parts of the amino acid sequence have identified functions (Figure 4.13B). One intracellular domain of the protein is required for ATP binding

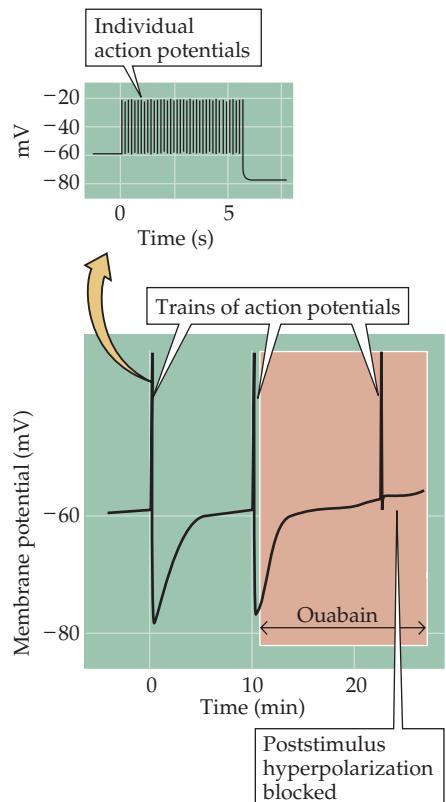
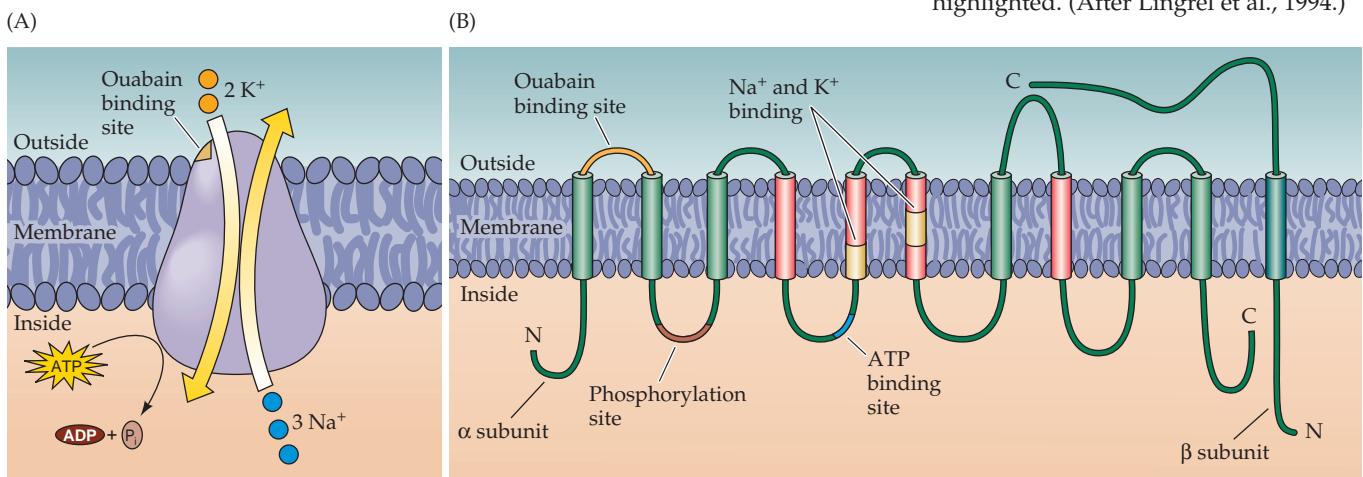


Figure 4.13 Molecular structure of the Na^+/K^+ pump. (A) General features of the pump. (B) The molecule spans the membrane 10 times. Amino acid residues thought to be important for binding of ATP, K^+ , and ouabain are highlighted. (After Lingrel et al., 1994.)



and hydrolysis, and the amino acid phosphorylated by ATP has been identified. Another extracellular domain may represent the binding site for ouabain. However, the sites involved in the most critical function of the pump—the movement of Na^+ and K^+ —have not yet been defined. Nonetheless, altering certain membrane-spanning domains (red in Figure 4.13B) impairs ion translocation; moreover, kinetic studies indicate that both ions bind to the pump at the same site. Because these ions move across the membrane, it is likely that this site traverses the plasma membrane; it is also likely that the site has a negative charge, since both Na^+ and K^+ are positively charged. The observation that removing negatively charged residues in a membrane-spanning domain of the protein (pale yellow in Figure 4.13B) greatly reduces Na^+ and K^+ binding provides at least a hint about the ion-translocating domain of the transporter molecule.

Summary

Ion transporters and channels have complementary functions. The primary purpose of transporters is to generate transmembrane concentration gradients, which are then exploited by ion channels to generate electrical signals. Ion channels are responsible for the voltage-dependent conductances of nerve cell membranes. The channels underlying the action potential are integral membrane proteins that open or close ion-selective pores in response to the membrane potential, allowing specific ions to diffuse across the membrane. The flow of ions through single open channels can be detected as tiny electrical currents, and the synchronous opening of many such channels generates the macroscopic currents that produce action potentials. Molecular studies show that such voltage-gated channels have highly conserved structures that are responsible for features such as ion permeation and voltage sensing, as well as the features that specify ion selectivity and toxin sensitivity. Other types of channels are sensitive to chemical signals, such as neurotransmitters or second messengers, or to heat or membrane deformation. A large number of ion channel genes create channels with a correspondingly wide range of functional characteristics, thus allowing different types of neurons to have a remarkable spectrum of electrical properties. Ion transporter proteins are quite different in both structure and function. The energy needed for ion movement against a concentration gradient (e.g., in maintaining the resting potential) is provided either by the hydrolysis of ATP or by the electrochemical gradient of co-transported ions. The Na^+/K^+ pump produces and maintains the transmembrane gradients of Na^+ and K^+ , while other transporters are responsible for the electrochemical gradients for other physiologically important ions, such as Cl^- , Ca^{2+} , and H^+ . Together, transporters and channels provide a reasonably comprehensive molecular explanation for the ability of neurons to generate electrical signals.

Additional Reading

Reviews

ARMSTRONG, C. M. AND B. HILLE (1998) Voltage-gated ion channels and electrical excitability. *Neuron* 20: 371–380.

BEZANILLA, F. AND A. M. CORREA (1995) Single-channel properties and gating of Na^+ and K^+ channels in the squid giant axon. In *Cephalopod Neurobiology*, N. J. Abbott, R. Williamson and L. Maddock (eds.). New York: Oxford University Press, pp. 131–151.

CATTERALL, W. A. (1988) Structure and function of voltage-sensitive ion channels. *Science* 242: 50–61.

ISOM, L. L., K. S. DE JONGH AND W. A. CATTERALL (1994) Auxiliary subunits of voltage-gated ion channels. *Neuron* 12: 1183–1194.

JAN, L. Y. AND Y. N. JAN (1997) Voltage-gated and inwardly rectifying potassium channels. *J. Physiol.* 505: 267–282.

JENTSCH, T. J., T. FRIEDRICH, A. SCHRIEVER AND H. YAMADA (1999) The CLC chloride channel family. *Pflügers Archiv* 437: 783–795.

KAPLAN, J. H. (2002) Biochemistry of Na,K -ATPase. *Annu. Rev. Biochem.* 71: 511–535.

KRISHNAL, O. (2003). The ASICs: Signaling molecules? Modulators? *Trends Neurosci.* 26: 477–483.

LINGREL, J. B., J. VAN HUYSSE, W. O'BRIEN, E. JEWELL-MOTZ, R. ASKEW AND P. SCHULTHEIS (1994) Structure-function studies of the Na,K -ATPase. *Kidney Internat.* 45: S32–S39.

MACKINNON, R. (2003) Potassium channels. *FEBS Lett.* 555: 62–65.

NEHER, E. (1992) Nobel lecture: Ion channels for communication between and within cells. *Neuron* 8: 605–612.

PATAPOUTIAN, A., A. M. PEIER, G. M. STORY AND V. VISWANATH (2003). ThermoTRP channels and beyond: Mechanisms of temperature sensation. *Nat. Rev. Neurosci.* 4: 529–539.

SEEBURG, P. H. (2002). A-to-I editing: New and old sites, functions and speculations. *Neuron* 35: 17–20.

SKOU, J. C. (1988) Overview: The Na,K pump. *Meth. Enzymol.* 156: 1–25.

Important Original Papers

ANTZ, C. AND 7 OTHERS (1997) NMR structure of inactivation gates from mammalian voltage-dependent potassium channels. *Nature* 385: 272–275.

BEZANILLA, F., E. PEROZO, D. M. PAPAZIAN AND

E. STEFANI (1991) Molecular basis of gating charge immobilization in Shaker potassium channels. *Science* 254: 679–683.

BOULTER, J. AND 6 OTHERS (1990) Molecular cloning and functional expression of glutamate receptor subunit genes. *Science* 249: 1033–1037.

CATERINA, M. J., M. A. SCHUMACHER, M. TOMINAGA, T. A. ROSEN, J. D. LEVINE AND D. JULIUS (1997) The capsaicin receptor: A heat-activated ion channel in the pain pathway. *Nature* 389: 816–824.

CHA, A., G. E. SNYDER, P. R. SELVIN AND F. BEZANILLA (1999) Atomic scale movement of the voltage-sensing region in a potassium channel measured via spectroscopy. *Nature* 402: 809–813.

DOYLE, D. A. AND 7 OTHERS (1998) The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science* 280: 69–77.

FAHLKE, C., H. T. YU, C. L. BECK, T. H. RHODES AND A. L. GEORGE JR. (1997) Pore-forming segments in voltage-gated chloride channels. *Nature* 390: 529–532.

HO, K. AND 6 OTHERS (1993) Cloning and expression of an inwardly rectifying ATP-regulated potassium channel. *Nature* 362: 31–38.

HODGKIN, A. L. AND R. D. KEYNES (1955) Active transport of cations in giant axons from *Sepia* and *Loligo*. *J. Physiol.* 128: 28–60.

HOSHI, T., W. N. ZAGOTTA AND R. W. ALDRICH (1990) Biophysical and molecular mechanisms of *Shaker* potassium channel inactivation. *Science* 250: 533–538.

JIANG, Y. AND 6 OTHERS (2003) X-ray structure of a voltage-dependent K^+ channel. *Nature* 423: 33–41.

LLANO, I., C. K. WEBB AND F. BEZANILLA (1988) Potassium conductance of squid giant axon. Single-channel studies. *J. Gen. Physiol.* 92: 179–196.

MIKAMI, A. AND 7 OTHERS (1989) Primary structure and functional expression of the cardiac dihydropyridine-sensitive calcium channel. *Nature* 340: 230–233.

NODA, M. AND 6 OTHERS (1986) Expression of functional sodium channels from cloned cDNA. *Nature* 322: 826–828.

NOWICKY, M. C., A. P. FOX AND R. W. TSIEN (1985) Three types of neuronal calcium channel with different calcium agonist sensitivity. *Nature* 316: 440–443.

PAPAZIAN, D. M., T. L. SCHWARZ, B. L. TEMPET, Y. N. JAN AND L. Y. JAN (1987) Cloning of

genomic and complementary DNA from *Shaker*, a putative potassium channel gene from *Drosophila*. *Science* 237: 749–753.

RANG, H. P. AND J. M. RITCHIE (1968) On the electrogenic sodium pump in mammalian non-myelinated nerve fibres and its activation by various external cations. *J. Physiol.* 196: 183–221.

SIGWORTH, F. J. AND E. NEHER (1980) Single Na^+ channel currents observed in cultured rat muscle cells. *Nature* 287: 447–449.

THOMAS, R. C. (1969) Membrane current and intracellular sodium changes in a snail neurone during extrusion of injected sodium. *J. Physiol.* 201: 495–514.

TOYOSHIMA, C., M. NAKASAKO, H. NOMURA AND H. OGAWA (2000) Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* 405: 647–655.

VANDERBERG, C. A. AND F. BEZANILLA (1991) A sodium channel model based on single channel, macroscopic ionic, and gating currents in the squid giant axon. *Biophys. J.* 60: 1511–1533.

WALDMANN, R., G. CHAMPIGNY, F. BASSILANA, C. HEURTEAUX AND M. LAZDUNSKI (1997) A proton-gated cation channel involved in acid-sensing. *Nature* 386: 173–177.

WEI, A. M., A. COVARUBIAS, A. BUTLER, K. BAKER, M. PAK AND L. SALKOFF (1990) K^+ current diversity is produced by an extended gene family conserved in *Drosophila* and mouse. *Science* 248: 599–603.

YANG, N., A. L. GEORGE JR. AND R. HORN (1996) Molecular basis of charge movement in voltage-gated sodium channels. *Neuron* 16: 113–22.

Books

AIDLEY, D. J. AND P. R. STANFIELD (1996) *Ion Channels: Molecules in Action*. Cambridge: Cambridge University Press.

ASHCROFT, F. M. (2000) *Ion Channels and Disease*. Boston: Academic Press.

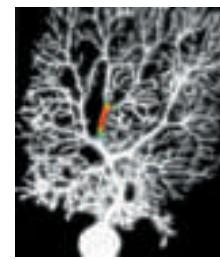
HILLE, B. (2001) *Ion Channels of Excitable Membranes*, 3rd Ed. Sunderland, MA: Sinauer Associates.

JUNGE, D. (1992) *Nerve and Muscle Excitation*, 3rd Ed. Sunderland, MA: Sinauer Associates.

NICHOLLS, D. G. (1994) *Proteins, Transmitters and Synapses*. Oxford: Blackwell Scientific

SIEGEL, G. J., B. W. AGRANOFF, R. W. ALBERS, S. K. FISHER AND M. D. UHLER (1999) *Basic Neurochemistry*. Philadelphia: Lippincott-Raven.

Chapter 5



Synaptic Transmission

Overview

The human brain contains at least 100 billion neurons, each with the ability to influence many other cells. Clearly, sophisticated and highly efficient mechanisms are needed to enable communication among this astronomical number of elements. Such communication is made possible by synapses, the functional contacts between neurons. Two different types of synapse—electrical and chemical—can be distinguished on the basis of their mechanism of transmission. At electrical synapses, current flows through gap junctions, which are specialized membrane channels that connect two cells. In contrast, chemical synapses enable cell-to-cell communication via the secretion of neurotransmitters; these chemical agents released by the presynaptic neurons produce secondary current flow in postsynaptic neurons by activating specific receptor molecules. The total number of neurotransmitters is not known, but is well over 100. Virtually all neurotransmitters undergo a similar cycle of use: synthesis and packaging into synaptic vesicles; release from the presynaptic cell; binding to postsynaptic receptors; and, finally, rapid removal and/or degradation. The secretion of neurotransmitters is triggered by the influx of Ca^{2+} through voltage-gated channels, which gives rise to a transient increase in Ca^{2+} concentration within the presynaptic terminal. The rise in Ca^{2+} concentration causes synaptic vesicles to fuse with the presynaptic plasma membrane and release their contents into the space between the pre- and postsynaptic cells. Although it is not yet understood exactly how Ca^{2+} triggers exocytosis, specific proteins on the surface of the synaptic vesicle and elsewhere in the presynaptic terminal mediate this process. Neurotransmitters evoke postsynaptic electrical responses by binding to members of a diverse group of neurotransmitter receptors. There are two major classes of receptors: those in which the receptor molecule is also an ion channel, and those in which the receptor and ion channel are separate molecules. These receptors give rise to electrical signals by transmitter-induced opening or closing of the ion channels. Whether the postsynaptic actions of a particular neurotransmitter are excitatory or inhibitory is determined by the ionic permeability of the ion channel affected by the transmitter, and by the concentration of permeant ions inside and outside the cell.

Electrical Synapses

Although there are many kinds of synapses within the human brain, they can be divided into two general classes: electrical synapses and chemical synapses. Although they are a distinct minority, electrical synapses are found in all nervous systems, permitting direct, passive flow of electrical current from one neuron to another.

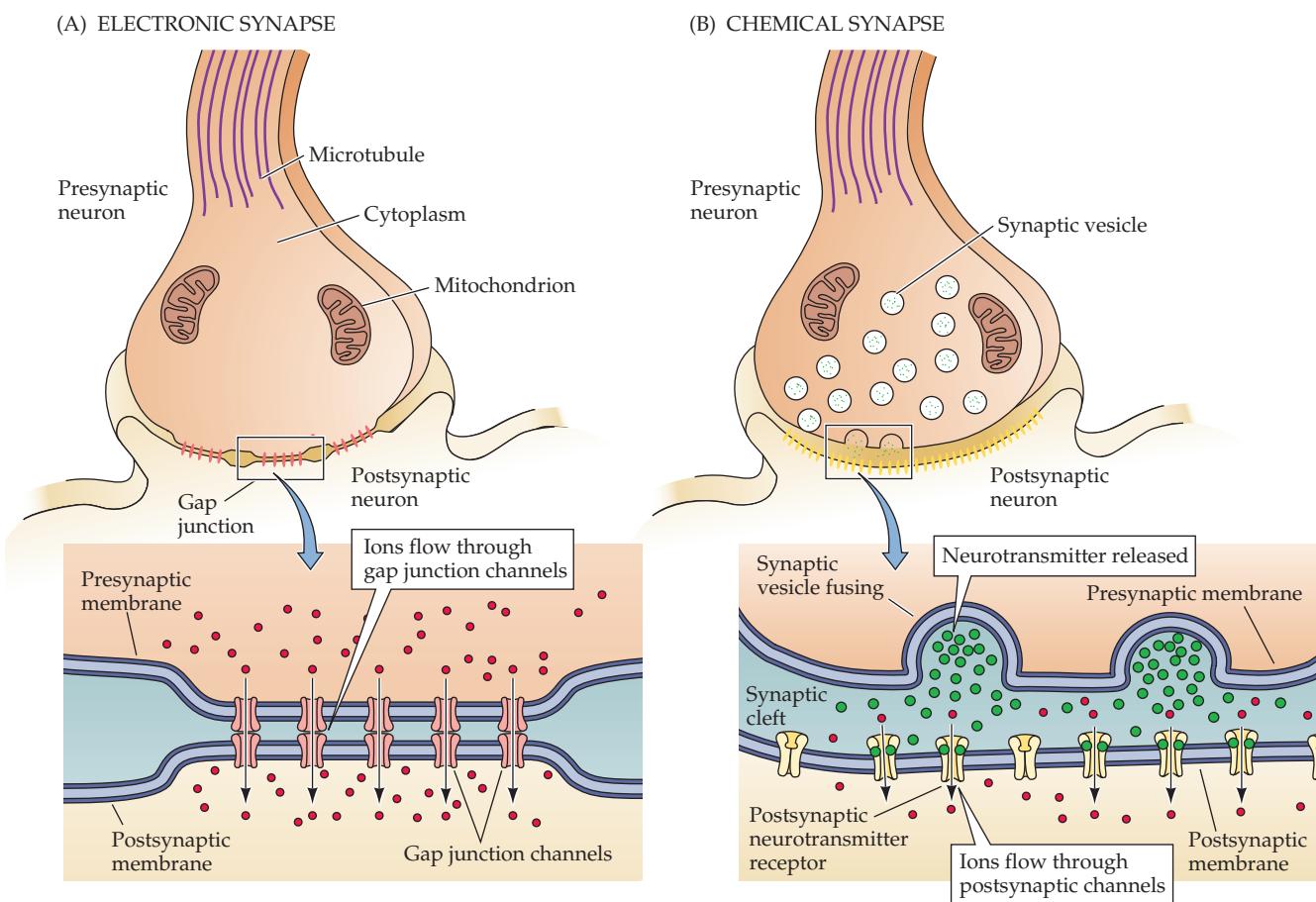


Figure 5.1 Electrical and chemical synapses differ fundamentally in their transmission mechanisms. (A) At electrical synapses, gap junctions between pre- and postsynaptic membranes permit current to flow passively through intercellular channels (blowup). This current flow changes the postsynaptic membrane potential, initiating (or in some instances inhibiting) the generation of postsynaptic action potentials. (B) At chemical synapses, there is no intercellular continuity, and thus no direct flow of current from pre- to postsynaptic cell. Synaptic current flows across the postsynaptic membrane only in response to the secretion of neurotransmitters, which open or close postsynaptic ion channels after binding to receptor molecules (blowup).

The structure of an electrical synapse is shown schematically in Figure 5.1A. The “upstream” neuron, which is the source of current, is called the **presynaptic** element, and the “downstream” neuron into which this current flows is termed **postsynaptic**. The membranes of the two communicating neurons come extremely close at the synapse and are actually linked together by an intercellular specialization called a **gap junction**. Gap junctions contain precisely aligned, paired channels in the membrane of the pre- and postsynaptic neurons, such that each channel pair forms a pore (see Figure 5.2A). The pore of a gap junction channel is much larger than the pores of the voltage-gated ion channels described in the previous chapter. As a result, a variety of substances can simply diffuse between the cytoplasm of the pre- and postsynaptic neurons. In addition to ions, substances that diffuse through gap junction pores include molecules with molecular weights as great as several hundred daltons. This permits ATP and other important intracellular metabolites, such as second messengers (see Chapter 7), to be transferred between neurons.

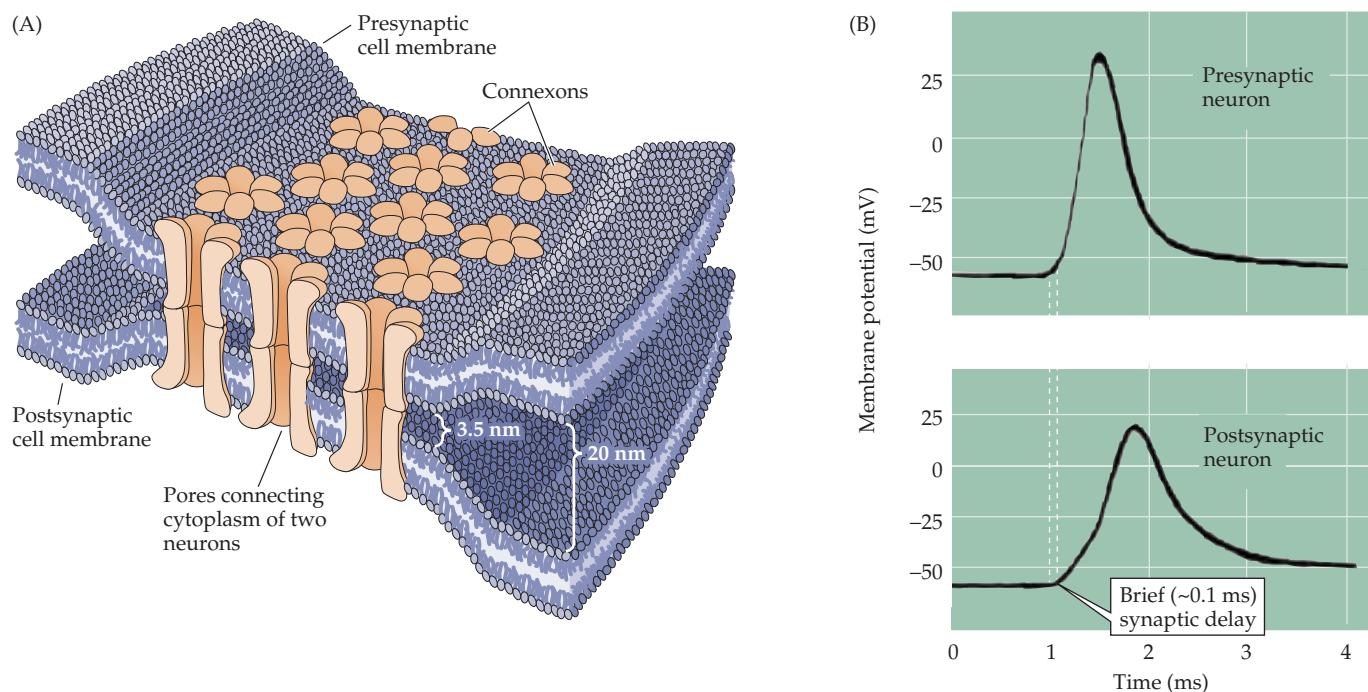
Electrical synapses thus work by allowing ionic current to flow passively through the gap junction pores from one neuron to another. The usual source of this current is the potential difference generated locally by the action potential (see Chapter 3). This arrangement has a number of interesting consequences. One is that transmission can be bidirectional; that is, current can flow in either direction across the gap junction, depending on which member of the coupled pair is invaded by an action potential (although

some types of gap junctions have special features that render their transmission unidirectional). Another important feature of the electrical synapse is that transmission is extraordinarily fast: because passive current flow across the gap junction is virtually instantaneous, communication can occur without the delay that is characteristic of chemical synapses.

These features are apparent in the operation of the first electrical synapse to be discovered, which resides in the crayfish nervous system. A postsynaptic electrical signal is observed at this synapse within a fraction of a millisecond after the generation of a presynaptic action potential (Figure 5.2). In fact, at least part of this brief synaptic delay is caused by propagation of the action potential into the presynaptic terminal, so that there may be essentially no delay at all in the transmission of electrical signals across the synapse. Such synapses interconnect many of the neurons within the circuit that allows the crayfish to escape from its predators, thus minimizing the time between the presence of a threatening stimulus and a potentially life-saving motor response.

A more general purpose of electrical synapses is to synchronize electrical activity among populations of neurons. For example, the brainstem neurons that generate rhythmic electrical activity underlying breathing are synchronized by electrical synapses, as are populations of interneurons in the cerebral cortex, thalamus, cerebellum, and other brain regions. Electrical transmission between certain hormone-secreting neurons within the mammalian hypothalamus ensures that all cells fire action potentials at about the same time, thus facilitating a burst of hormone secretion into the circulation. The fact that gap junction pores are large enough to allow molecules such as ATP and second messengers to diffuse intercellularly also permits electrical synapses to coordinate the intracellular signaling and metabolism of coupled cells. This property may be particularly important for glial cells, which form large intracellular signaling networks via their gap junctions.

Figure 5.2 Structure and function of gap junctions at electrical synapses. (A) Gap junctions consist of hexameric complexes formed by the coming together of subunits called connexons, which are present in both the pre- and postsynaptic membranes. The pores of the channels connect to one another, creating electrical continuity between the two cells. (B) Rapid transmission of signals at an electrical synapse in the crayfish. An action potential in the presynaptic neuron causes the postsynaptic neuron to be depolarized within a fraction of a millisecond. (B after Furshpan and Potter, 1959.)



Signal Transmission at Chemical Synapses

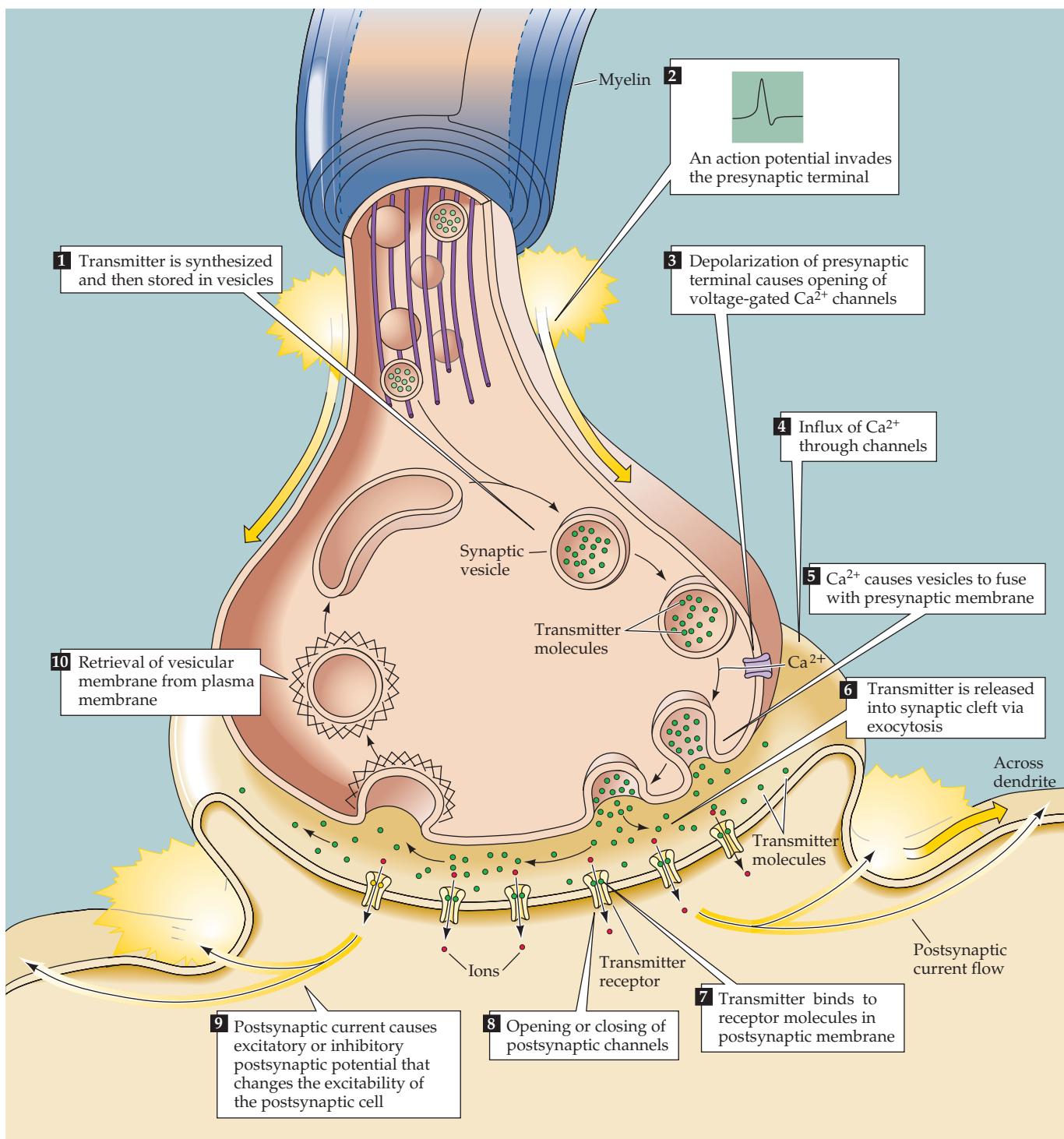
The general structure of a chemical synapse is shown schematically in Figure 5.1B. The space between the pre- and postsynaptic neurons is substantially greater at chemical synapses than at electrical synapses and is called the **synaptic cleft**. However, the key feature of all chemical synapses is the presence of small, membrane-bounded organelles called **synaptic vesicles** within the presynaptic terminal. These spherical organelles are filled with one or more **neurotransmitters**, the chemical signals secreted from the presynaptic neuron, and it is these chemical agents acting as messengers between the communicating neurons that gives this type of synapse its name.

Transmission at chemical synapses is based on the elaborate sequence of events depicted in Figure 5.3. The process is initiated when an action potential invades the terminal of the presynaptic neuron. The change in membrane potential caused by the arrival of the action potential leads to the opening of voltage-gated calcium channels in the presynaptic membrane. Because of the steep concentration gradient of Ca^{2+} across the presynaptic membrane (the external Ca^{2+} concentration is approximately 10^{-3} M , whereas the internal Ca^{2+} concentration is approximately 10^{-7} M), the opening of these channels causes a rapid influx of Ca^{2+} into the presynaptic terminal, with the result that the Ca^{2+} concentration of the cytoplasm in the terminal transiently rises to a much higher value. Elevation of the presynaptic Ca^{2+} concentration, in turn, allows synaptic vesicles to fuse with the plasma membrane of the presynaptic neuron. The Ca^{2+} -dependent fusion of synaptic vesicles with the terminal membrane causes their contents, most importantly neurotransmitters, to be released into the synaptic cleft.

Following exocytosis, transmitters diffuse across the synaptic cleft and bind to specific receptors on the membrane of the postsynaptic neuron. The binding of neurotransmitter to the receptors causes channels in the postsynaptic membrane to open (or sometimes to close), thus changing the ability of ions to flow into (or out of) the postsynaptic cells. The resulting neurotransmitter-induced current flow alters the conductance and (usually) the membrane potential of the postsynaptic neuron, increasing or decreasing the probability that the neuron will fire an action potential. In this way, information is transmitted from one neuron to another.

Properties of Neurotransmitters

The notion that electrical information can be transferred from one neuron to the next by means of chemical signaling was the subject of intense debate through the first half of the twentieth century. A key experiment that supported this idea was performed in 1926 by German physiologist Otto Loewi. Acting on an idea that allegedly came to him in the middle of the night, Loewi proved that electrical stimulation of the vagus nerve slows the heartbeat by releasing a chemical signal. He isolated and perfused the hearts of two frogs, monitoring the rates at which they were beating (Figure 5.4). His experiment collected the perfusate flowing through the stimulated heart and transferred this solution to the second heart. When the vagus nerve to the first heart was stimulated, the beat of this heart slowed. Remarkably, even though the vagus nerve of the second heart had not been stimulated, its beat also slowed when exposed to the perfusate from the first heart. This result showed that the vagus nerve regulates the heart rate by releasing a chemical that accumulates in the perfusate. Originally referred to as "vagus substance," the agent was later shown to be **acetylcholine (ACh)**. ACh is now known to be a neurotransmitter that acts not only in the heart but at a vari-



ety of postsynaptic targets in the central and peripheral nervous systems, preeminently at the neuromuscular junction of striated muscles and in the visceral motor system (see Chapters 6 and 20).

Over the years, a number of formal criteria have emerged that definitively identify a substance as a neurotransmitter (Box A). These have led to the identification of more than 100 different neurotransmitters, which can be

Figure 5.3 Sequence of events involved in transmission at a typical chemical synapse.

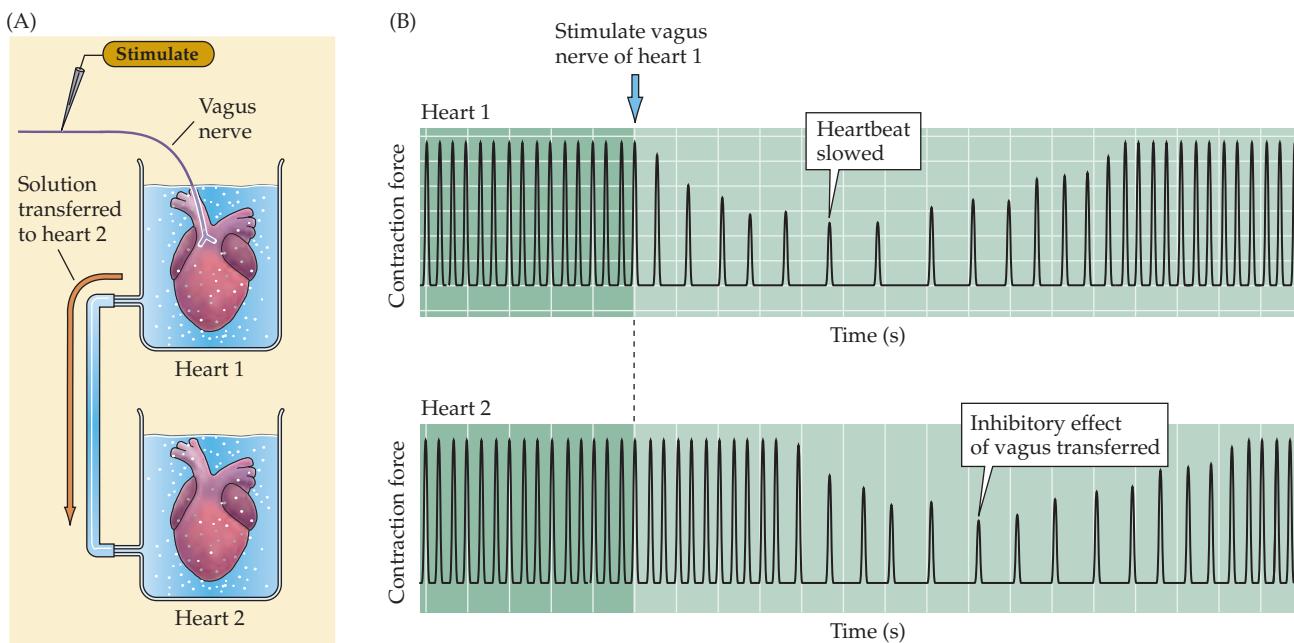


Figure 5.4 Loewi’s experiment demonstrating chemical neurotransmission. (A) Diagram of experimental setup. (B) When the vagus nerve of an isolated frog’s heart was stimulated, the heart rate decreased (upper panel). If the perfusion fluid from the stimulated heart was transferred to a second heart, its rate decreased as well (lower panel).

classified into two broad categories: small-molecule neurotransmitters and neuropeptides (Chapter 6). Having more than one transmitter diversifies the physiological repertoire of synapses. Multiple neurotransmitters can produce different types of responses on individual postsynaptic cells. For example, a neuron can be excited by one type of neurotransmitter and inhibited by another type of neurotransmitter. The speed of postsynaptic responses produced by different transmitters also differs, allowing control of electrical signaling over different time scales. In general, small-molecule neurotransmitters mediate rapid synaptic actions, whereas neuropeptides tend to modulate slower, ongoing synaptic functions.

Until relatively recently, it was believed that a given neuron produced only a single type of neurotransmitter. It is now clear, however, that many types of neurons synthesize and release two or more different neurotransmitters. When more than one transmitter is present within a nerve terminal, the molecules are called **co-transmitters**. Because different types of transmitters can be packaged in different populations of synaptic vesicles, co-transmitters need not be released simultaneously. When peptide and small-molecule neurotransmitters act as co-transmitters at the same synapse, they are differentially released according to the pattern of synaptic activity: low-frequency activity often releases only small neurotransmitters, whereas high-frequency activity is required to release neuropeptides from the same pre-synaptic terminals. As a result, the chemical signaling properties of such synapses change according to the rate of activity.

Effective synaptic transmission requires close control of the concentration of neurotransmitters within the synaptic cleft. Neurons have therefore developed a sophisticated ability to regulate the synthesis, packaging, release, and

Box A

Criteria That Define a Neurotransmitter

Three primary criteria have been used to confirm that a molecule acts as a neurotransmitter at a given chemical synapse.

1. The substance must be present within the presynaptic neuron. Clearly, a chemical cannot be secreted from a presynaptic neuron unless it is present there. Because elaborate biochemical pathways are required to produce neurotransmitters, showing that the enzymes and precursors required to synthesize the substance are present in presynaptic neurons provides additional evidence that the substance is used as a transmitter. Note, however, that since the transmitters glutamate, glycine, and aspartate are also needed for protein synthesis and other metabolic reactions in all neurons, their presence is *not* sufficient evidence to establish them as neurotransmitters.

2. The substance must be released in response to presynaptic depolarization, and the release must be Ca^{2+} -dependent.

Another essential criterion for identifying a neurotransmitter is to demonstrate that it is released from the presynaptic neuron in response to presynaptic electrical activity, and that this release requires Ca^{2+} influx into the presynaptic terminal. Meeting this criterion is technically challenging, not only because it may be difficult to selectively stimulate the presynaptic neurons, but also because enzymes and transporters efficiently remove the secreted neurotransmitters.

3. Specific receptors for the substance must be present on the postsynaptic cell. A neurotransmitter cannot act on its target unless specific receptors for the transmitter are present in the postsynaptic membrane. One way to demonstrate receptors is to show that application of exogenous transmitter mimics the post-

synaptic effect of presynaptic stimulation. A more rigorous demonstration is to show that agonists and antagonists that alter the normal postsynaptic response have the same effect when the substance in question is applied exogenously. High-resolution histological methods can also be used to show that specific receptors are present in the postsynaptic membrane (by detection of radioactively labeled receptor antibodies, for example).

Fulfilling these criteria establishes unambiguously that a substance is used as a transmitter at a given synapse. Practical difficulties, however, have prevented these standards from being applied at many types of synapses. It is for this reason that so many substances must be referred to as "putative" neurotransmitters.

Demonstrating the identity of a neurotransmitter at a synapse requires showing (1) its presence, (2) its release, and (3) the postsynaptic presence of specific receptors.

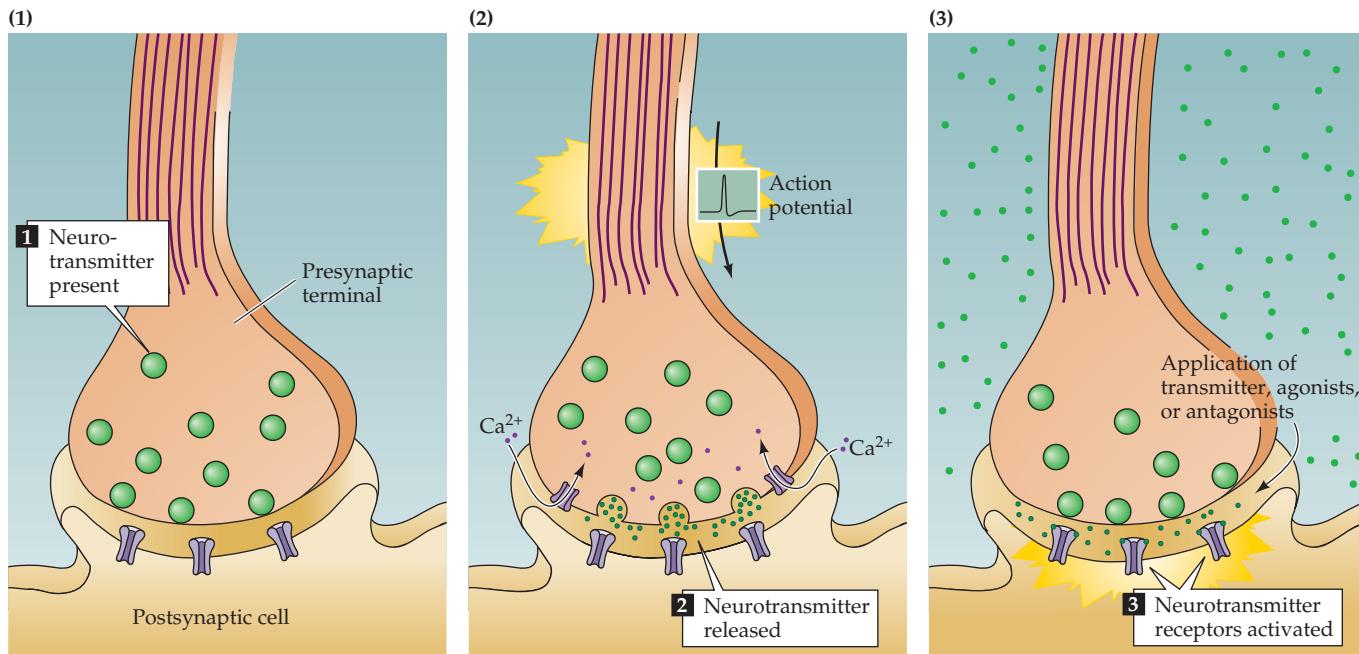
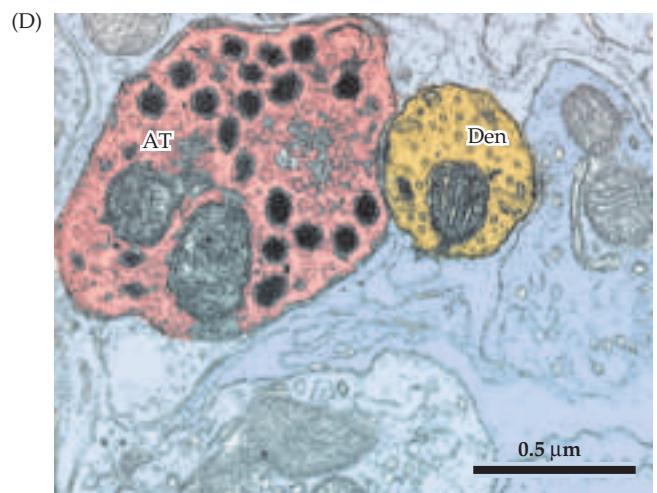
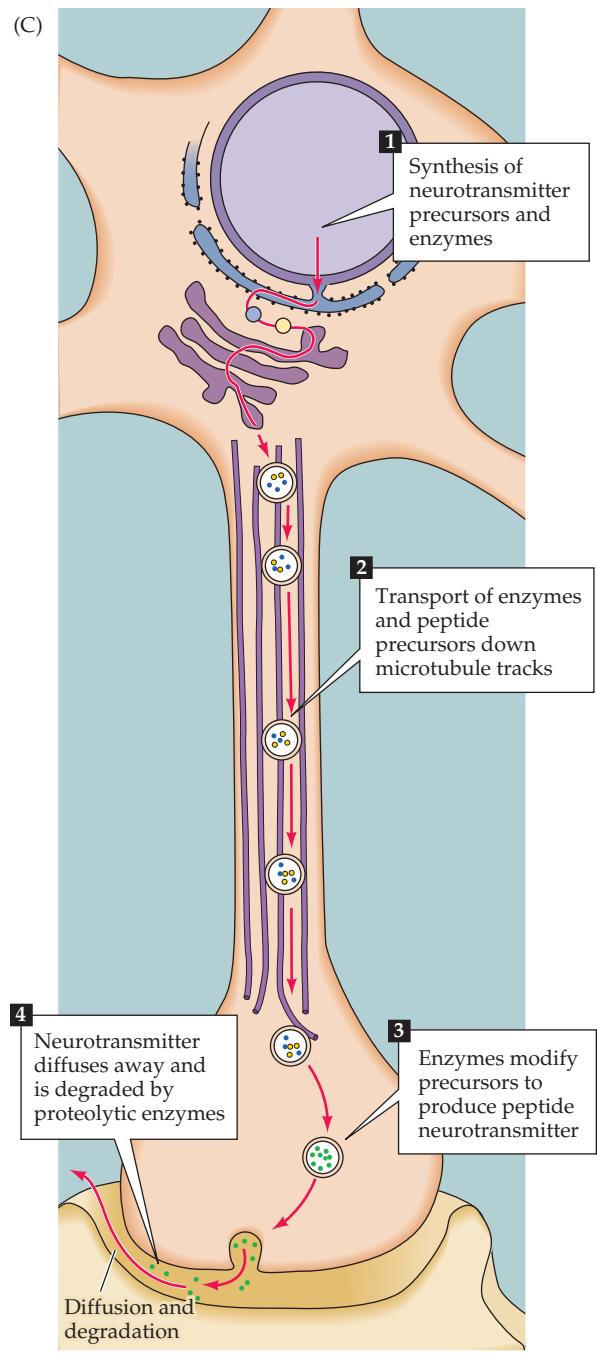
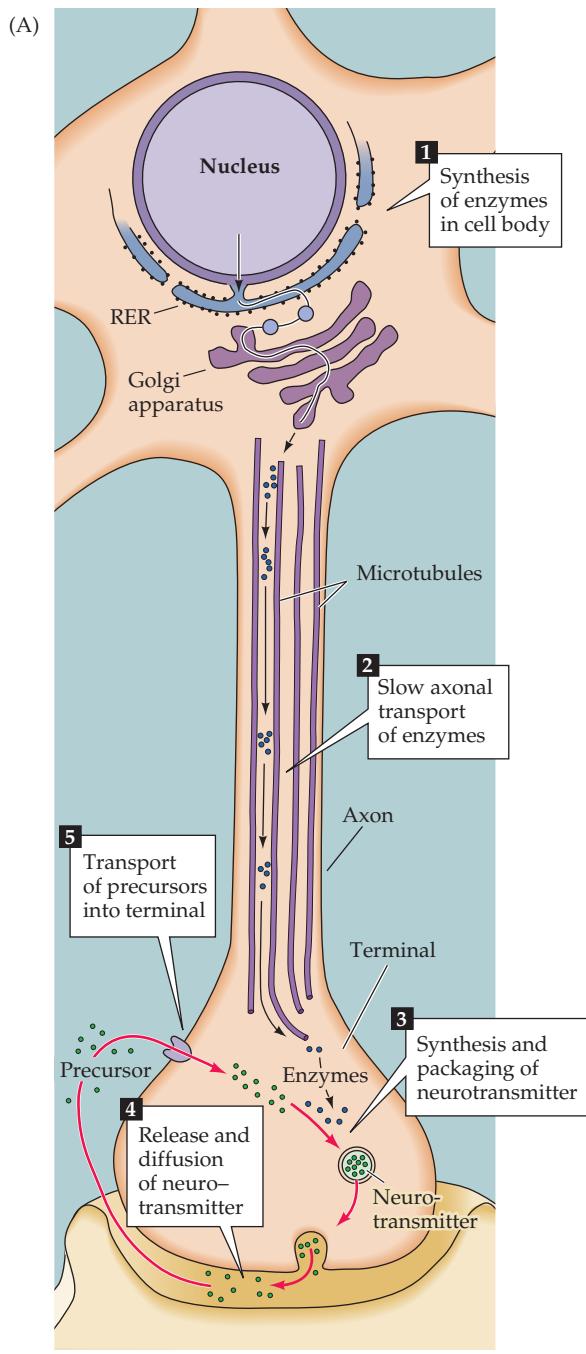
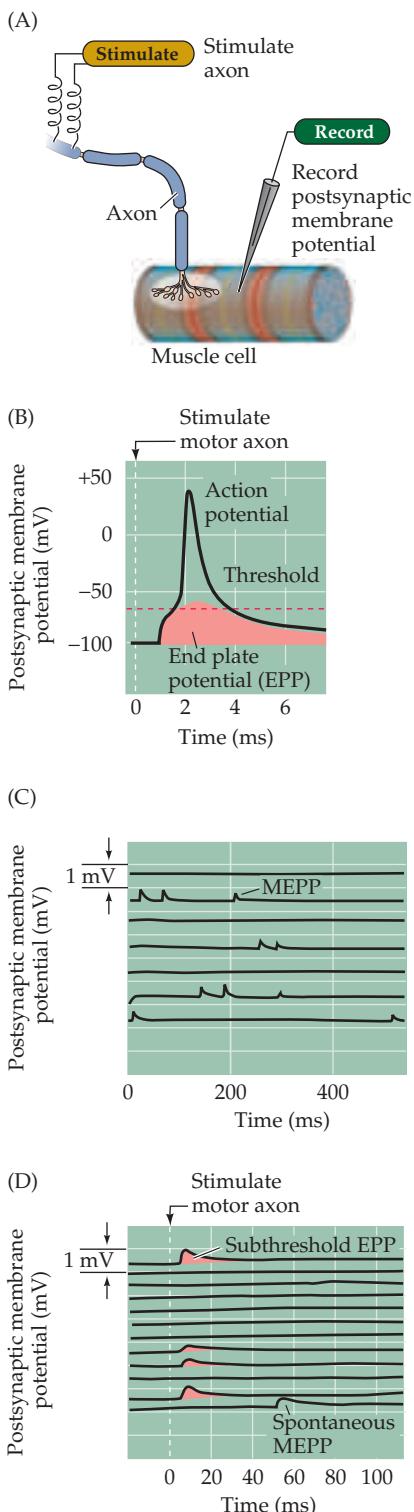


Figure 5.5 Metabolism of small-molecule and peptide transmitters. (A) Small-molecule neurotransmitters are synthesized at nerve terminals. The enzymes necessary for neurotransmitter synthesis are made in the cell body of the presynaptic cell (1) and are transported down the axon by slow axonal transport (2). Precursors are taken up into the terminals by specific transporters, and neurotransmitter synthesis and packaging take place within the nerve endings (3). After vesicle fusion and release (4), the neurotransmitter may be enzymatically degraded. The reuptake of the neurotransmitter (or its metabolites) starts another cycle of synthesis, packaging, release, and removal (5). (B) Small clear-core vesicles at a synapse between an axon terminal (AT) and a dendritic spine (Den) in the central nervous system. Such vesicles typically contain small-molecule neurotransmitters. (C) Peptide neurotransmitters, as well as the enzymes that modify their precursors, are synthesized in the cell body (1). Enzymes and propeptides are packaged into vesicles in the Golgi apparatus. During fast axonal transport of these vesicles to the nerve terminals (2), the enzymes modify the propeptides to produce one or more neurotransmitter peptides (3). After vesicle fusion and exocytosis, the peptides diffuse away and are degraded by proteolytic enzymes (4). (D) Large dense-core vesicles in a central axon terminal (AT) synapsing onto a dendrite (Den). Such vesicles typically contain neuropeptides or, in some cases, biogenic amines. (B and D from Peters, Palay, and Webster, 1991.)

degradation (or removal) of neurotransmitters to achieve the desired levels of transmitter molecules. The synthesis of small-molecule neurotransmitters occurs locally within presynaptic terminals (Figure 5.5A). The enzymes needed to synthesize these transmitters are produced in the neuronal cell body and transported to the nerve terminal cytoplasm at 0.5–5 millimeters a day by a mechanism called **slow axonal transport**. The precursor molecules required to make new molecules of neurotransmitter are usually taken into the nerve terminal by transporters found in the plasma membrane of the terminal. The enzymes synthesize neurotransmitters in the cytoplasm of the presynaptic terminal and the transmitters are then loaded into synaptic vesicles via transporters in the vesicular membrane (see Chapter 4). For some small-molecule neurotransmitters, the final steps of synthesis occur inside the synaptic vesicles. Most small-molecule neurotransmitters are packaged in vesicles 40 to 60 nm in diameter, the centers of which appear clear in electron micrographs; accordingly, these vesicles are referred to as **small clear-core vesicles** (Figure 5.5B). Neuropeptides are synthesized in the cell body of a neuron, meaning that the peptide is produced a long distance away from its site of secretion (Figure 5.5C). To solve this problem, peptide-filled vesicles are transported along an axon and down to the synaptic terminal via **fast axonal transport**. This process carries vesicles at rates up to 400 mm/day along cytoskeletal elements called microtubules (in contrast to the slow axonal transport of the enzymes that synthesize small-molecule transmitters). Microtubules are long, cylindrical filaments, 25 nm in diameter, present throughout neurons and other cells. Peptide-containing vesicles are moved along these microtubule “tracks” by ATP-requiring “motor” proteins such as kinesin. Neuropeptides are packaged into synaptic vesicles that range from 90 to 250 nm in diameter. These vesicles are electron-dense in electron micrographs—hence they are referred to as **large dense-core vesicles** (Figure 5.5D).

After a neurotransmitter has been secreted into the synaptic cleft, it must be removed to enable the postsynaptic cell to engage in another cycle of syn-





aptic transmission. The removal of neurotransmitters involves diffusion away from the postsynaptic receptors, in combination with reuptake into nerve terminals or surrounding glial cells, degradation by specific enzymes, or a combination of these mechanisms. Specific transporter proteins remove most small-molecule neurotransmitters (or their metabolites) from the synaptic cleft, ultimately delivering them back to the presynaptic terminal for reuse.

Quantal Release of Neurotransmitters

Much of the evidence leading to the present understanding of chemical synaptic transmission was obtained from experiments examining the release of ACh at neuromuscular junctions. These synapses between spinal motor neurons and skeletal muscle cells are simple, large, and peripherally located, making them particularly amenable to experimental analysis. Such synapses occur at specializations called **end plates** because of the saucer-like appearance of the site on the muscle fiber where the presynaptic axon elaborates its terminals (Figure 5.6A). Most of the pioneering work on neuromuscular transmission was performed by Bernard Katz and his collaborators at University College London during the 1950s and 1960s, and Katz has been widely recognized for his remarkable contributions to understanding synaptic transmission. Though he worked primarily on the frog neuromuscular junction, numerous subsequent experiments have confirmed the applicability of his observations to transmission at chemical synapses throughout the nervous system.

When an intracellular microelectrode is used to record the membrane potential of a muscle cell, an action potential in the presynaptic motor neuron can be seen to elicit a transient depolarization of the postsynaptic muscle fiber. This change in membrane potential, called an **end plate potential (EPP)**, is normally large enough to bring the membrane potential of the muscle cell well above the threshold for producing a postsynaptic action potential (Figure 5.6B). The postsynaptic action potential triggered by the EPP causes the muscle fiber to contract. Unlike the case for electrical synapses, there is a pronounced delay between the time that the presynaptic motor neuron is stimulated and when the EPP occurs in the postsynaptic muscle cell. Such a delay is characteristic of all chemical synapses.

One of Katz's seminal findings, in studies carried out with Paul Fatt in 1951, was that spontaneous changes in muscle cell membrane potential occur even in the absence of stimulation of the presynaptic motor neuron (Figure 5.6C). These changes have the same shape as EPPs but are much

Figure 5.6 Synaptic transmission at the neuromuscular junction. (A) Experimental arrangement, typically using the muscle of a frog or rat. The axon of the motor neuron innervating the muscle fiber is stimulated with an extracellular electrode, while an intracellular microelectrode is inserted into the postsynaptic muscle cell to record its electrical responses. (B) End plate potentials (EPPs) evoked by stimulation of a motor neuron are normally above threshold and therefore produce an action potential in the postsynaptic muscle cell. (C) Spontaneous miniature EPPs (MEPPs) occur in the absence of presynaptic stimulation. (D) When the neuromuscular junction is bathed in a solution that has a low concentration of Ca^{2+} , stimulating the motor neuron evokes EPPs whose amplitudes are reduced to about the size of MEPPs. (After Fatt and Katz, 1952.)

smaller (typically less than 1 mV in amplitude, compared to an EPP of perhaps 40 or 50 mV). Both EPPs and these small, spontaneous events are sensitive to pharmacological agents that block postsynaptic acetylcholine receptors, such as curare (see Box B in Chapter 6). These and other parallels between EPPs and the spontaneously occurring depolarizations led Katz and his colleagues to call these spontaneous events **miniature end plate potentials**, or MEPPs.

The relationship between the full-blown end plate potential and MEPPs was clarified by careful analysis of the EPPs. The magnitude of the EPP provides a convenient electrical assay of neurotransmitter secretion from a motor neuron terminal; however, measuring it is complicated by the need to prevent muscle contraction from dislodging the microelectrode. The usual means of eliminating muscle contractions is either to lower Ca^{2+} concentration in the extracellular medium or to partially block the postsynaptic ACh receptors with the drug curare. As expected from the scheme illustrated in Figure 5.3, lowering the Ca^{2+} concentration reduces neurotransmitter secretion, thus reducing the magnitude of the EPP below the threshold for postsynaptic action potential production and allowing it to be measured more precisely. Under such conditions, stimulation of the motor neuron produces very small EPPs that fluctuate in amplitude from trial to trial (Figure 5.6D). These fluctuations give considerable insight into the mechanisms responsible for neurotransmitter release. In particular, the variable evoked response in low Ca^{2+} is now known to result from the release of unit amounts of ACh by the presynaptic nerve terminal. Indeed, the amplitude of the smallest evoked response is strikingly similar to the size of single MEPPs (compare Figure 5.6C and D). Further supporting this similarity, increments in the EPP response (Figure 5.7A) occur in units about the size of single MEPPs (Figure 5.7B). These “quantal” fluctuations in the amplitude of EPPs indicated to Katz and colleagues that EPPs are made up of individual units, each equivalent to a MEPP.

The idea that EPPs represent the simultaneous release of many MEPP-like units can be tested statistically. A method of statistical analysis based on the independent occurrence of unitary events (called Poisson statistics) predicts what the distribution of EPP amplitudes would look like during a large number of trials of motor neuron stimulation, under the assumption that EPPs are built up from unitary events like MEPPs (see Figure 5.7B). The distribution of EPP amplitudes determined experimentally was found to be just that expected if transmitter release from the motor neuron is indeed quantal (the red curve in Figure 5.7A). Such analyses confirmed the idea that release of acetylcholine does indeed occur in discrete packets, each equivalent to a MEPP. In short, a presynaptic action potential causes a postsynaptic EPP because it synchronizes the release of many transmitter quanta.

Release of Transmitters from Synaptic Vesicles

The discovery of the quantal release of packets of neurotransmitter immediately raised the question of how such quanta are formed and discharged into the synaptic cleft. At about the time Katz and his colleagues were using physiological methods to discover quantal release of neurotransmitter, electron microscopy revealed, for the first time, the presence of synaptic vesicles in presynaptic terminals. Putting these two discoveries together, Katz and others proposed that synaptic vesicles loaded with transmitter are the source of the quanta. Subsequent biochemical studies confirmed that synaptic ves-

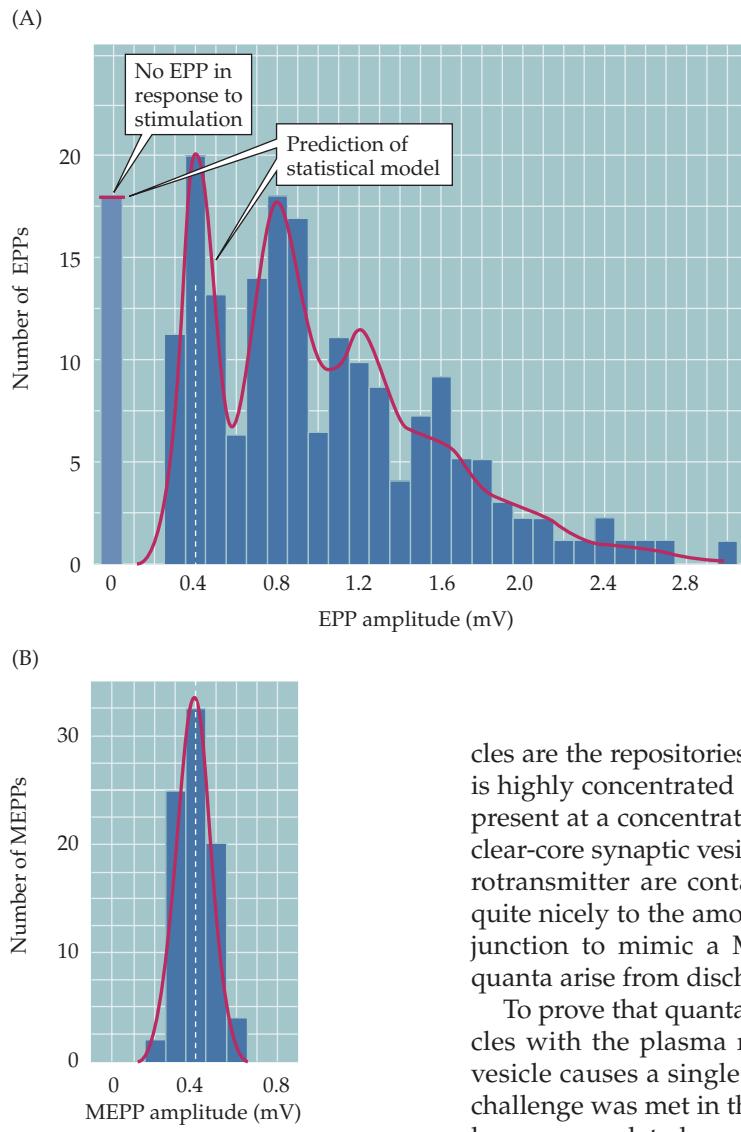


Figure 5.7 Quantized distribution of EPP amplitudes evoked in a low Ca^{2+} solution. Peaks of EPP amplitudes (A) tend to occur in integer multiples of the mean amplitude of MEPPs, whose amplitude distribution is shown in (B). The leftmost bar in the EPP amplitude distribution shows trials in which presynaptic stimulation failed to elicit an EPP in the muscle cell. The red curve indicates the prediction of a statistical model based on the assumption that the EPPs result from the independent release of multiple MEPP-like quanta. The observed match, including the predicted number of failures, supports this interpretation. (After Boyd and Martin, 1955.)

cles are the repositories of transmitters. These studies have shown that ACh is highly concentrated in the synaptic vesicles of motor neurons, where it is present at a concentration of about 100 mM. Given the diameter of a small, clear-core synaptic vesicle (~50 nm), approximately 10,000 molecules of neurotransmitter are contained in a single vesicle. This number corresponds quite nicely to the amount of ACh that must be applied to a neuromuscular junction to mimic a MEPP, providing further support for the idea that quanta arise from discharge of the contents of single synaptic vesicles.

To prove that quanta are caused by the fusion of individual synaptic vesicles with the plasma membrane, it is necessary to show that each fused vesicle causes a single quantal event to be recorded postsynaptically. This challenge was met in the late 1970s, when John Heuser, Tom Reese, and colleagues correlated measurements of vesicle fusion with the quantal content of EPPs at the neuromuscular junction. In their experiments, the number of vesicles that fused with the presynaptic plasma membrane was measured by electron microscopy in terminals that had been treated with a drug (4-aminopyridine, or 4-AP) that enhances the number of vesicle fusion events produced by single action potentials (Figure 5.8A). Parallel electrical measurements were made of the quantal content of the EPPs elicited in this way. A comparison of the number of synaptic vesicle fusions observed with the electron microscope and the number of quanta released at the synapse showed a good correlation between these two measures (Figure 5.8B). These results remain one of the strongest lines of support for the idea that a quantum of transmitter release is due to a synaptic vesicle fusing with the presynaptic membrane. Subsequent evidence, based on other means of measuring vesicle fusion, has left no doubt about the validity of this general interpretation of chemical synaptic transmission. Very recent work has identified structures within the presynaptic terminal that connect vesicles to the plasma membrane and may be involved in membrane fusion (Figure 5.8C).

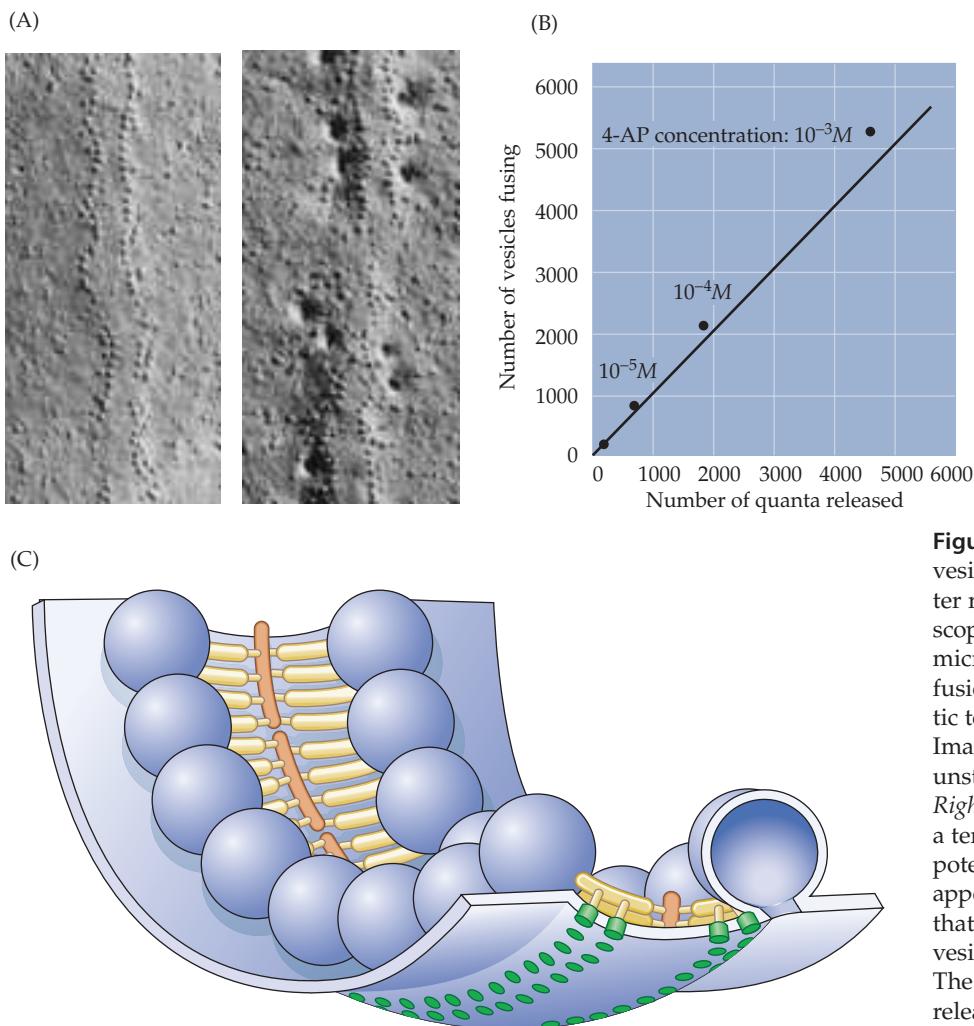


Figure 5.8 Relationship of synaptic vesicle exocytosis and quantal transmitter release. (A) A special electron microscopical technique called freeze-fracture microscopy was used to visualize the fusion of synaptic vesicles in presynaptic terminals of frog motor neurons. *Left:* Image of the plasma membrane of an unstimulated presynaptic terminal. *Right:* Image of the plasma membrane of a terminal stimulated by an action potential. Stimulation causes the appearance of dimple-like structures that represent the fusion of synaptic vesicles with the presynaptic membrane. The view is as if looking down on the release sites from outside the presynaptic terminal. (B) Comparison of the number of observed vesicle fusions to the number of quanta released by a pre-synaptic action potential. Transmitter release was varied by using a drug (4-AP) that affects the duration of the pre-synaptic action potential, thus changing the amount of calcium that enters during the action potential. The diagonal line is the 1:1 relationship expected if each vesicle that opened released a single quantum of transmitter. (C) Fine structure of vesicle fusion sites of frog presynaptic terminals. Synaptic vesicles are arranged in rows and are connected to each other and to the plasma membrane by a variety of proteinaceous structures (yellow). Green structures in the presynaptic membrane, corresponding to the rows of particles seen in (A), are thought to be Ca^{2+} channels. (A and B from Heuser et al., 1979; C after Harlow et al., 2001)

Local Recycling of Synaptic Vesicles

The fusion of synaptic vesicles causes new membrane to be added to the plasma membrane of the presynaptic terminal, but the addition is not permanent. Although a bout of exocytosis can dramatically increase the surface area of presynaptic terminals, this extra membrane is removed within a few minutes. Heuser and Reese performed another important set of experiments showing that the fused vesicle membrane is actually retrieved and taken back into the cytoplasm of the nerve terminal (a process called endocytosis). The experiments, again carried out at the frog neuromuscular junction, were based on filling the synaptic cleft with horseradish peroxidase (HRP), an enzyme that can be made to produce a dense reaction product that is visible in an electron microscope. Under appropriate experimental conditions, endocytosis could then be visualized by the uptake of HRP into the nerve terminal (Figure 5.9). To activate endocytosis, the presynaptic terminal was stimulated with a train of action potentials, and the subsequent fate of the HRP was followed by electron microscopy. Immediately follow-

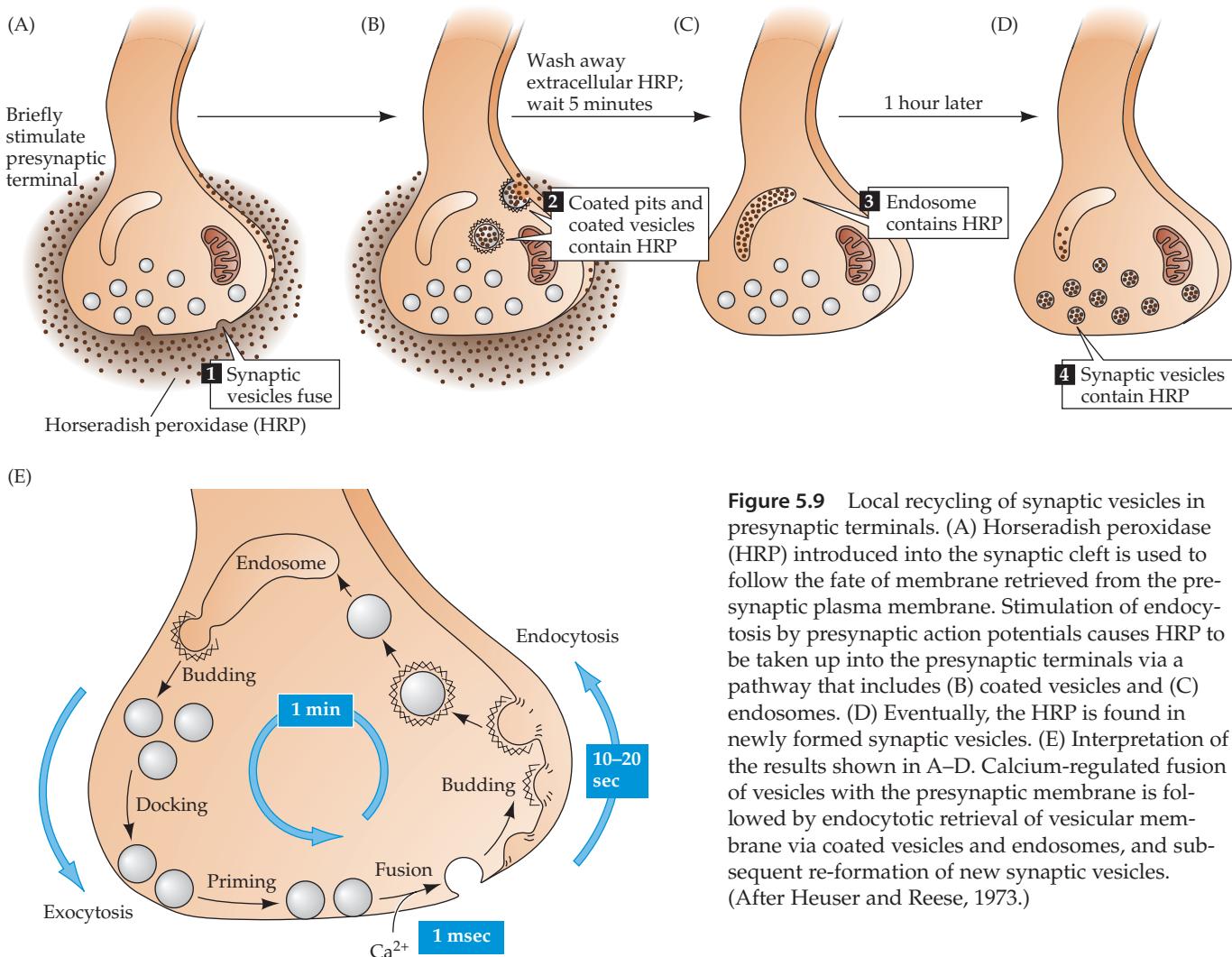


Figure 5.9 Local recycling of synaptic vesicles in presynaptic terminals. (A) Horseradish peroxidase (HRP) introduced into the synaptic cleft is used to follow the fate of membrane retrieved from the pre-synaptic plasma membrane. Stimulation of endocytosis by presynaptic action potentials causes HRP to be taken up into the presynaptic terminals via a pathway that includes (B) coated vesicles and (C) endosomes. (D) Eventually, the HRP is found in newly formed synaptic vesicles. (E) Interpretation of the results shown in A–D. Calcium-regulated fusion of vesicles with the presynaptic membrane is followed by endocytotic retrieval of vesicular membrane via coated vesicles and endosomes, and subsequent re-formation of new synaptic vesicles. (After Heuser and Reese, 1973.)

ing stimulation, the HRP was found within special endocytotic organelles called coated vesicles (Figure 5.9A,B). A few minutes later, however, the coated vesicles had disappeared and the HRP was found in a different organelle, the endosome (Figure 5.9C). Finally, within an hour after stimulating the terminal, the HRP reaction product appeared inside synaptic vesicles (Figure 5.9D).

These observations indicate that synaptic vesicle membrane is recycled within the presynaptic terminal via the sequence summarized in Figure 5.9E. In this process, called the **synaptic vesicle cycle**, the retrieved vesicular membrane passes through a number of intracellular compartments—such as coated vesicles and endosomes—and is eventually used to make new synaptic vesicles. After synaptic vesicles are re-formed, they are stored in a reserve pool within the cytoplasm until they need to participate again in neurotransmitter release. These vesicles are mobilized from the reserve pool, docked at the presynaptic plasma membrane, and primed to participate in exocytosis once again. More recent experiments, employing a fluorescent label rather than HRP, have determined the time course of synaptic vesicle recycling. These studies indicate that the entire vesicle cycle requires approximately 1 minute, with membrane budding during endocytosis requiring 10–20 sec-

onds of this time. As can be seen from the 1-millisecond delay in transmission following excitation of the presynaptic terminal (see Figure 5.6B), membrane fusion during exocytosis is much more rapid than budding during endocytosis. Thus, all of the recycling steps interspersed between membrane budding and subsequent refusion of a vesicle are completed in less than a minute.

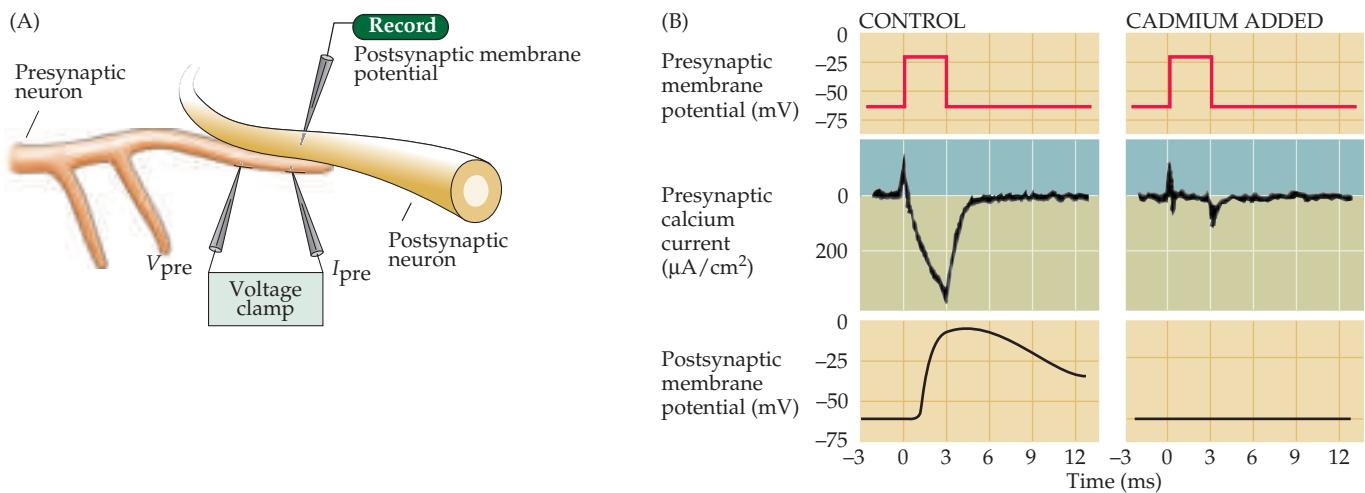
The precursors to synaptic vesicles *originally* are produced in the endoplasmic reticulum and Golgi apparatus in the neuronal cell body. Because of the long distance between the cell body and the presynaptic terminal in most neurons, transport of vesicles from the soma would not permit rapid replenishment of synaptic vesicles during continuous neural activity. Thus, local recycling is well suited to the peculiar anatomy of neurons, giving nerve terminals the means to provide a continual supply of synaptic vesicles. As might be expected, defects in synaptic vesicle recycling can cause severe neurological disorders, some of which are described in Box B.

The Role of Calcium in Transmitter Secretion

As was apparent in the experiments of Katz and others described in the preceding sections, lowering the concentration of Ca^{2+} outside a presynaptic motor nerve terminal reduces the size of the EPP (compare Figure 5.6B and D). Moreover, measurement of the number of transmitter quanta released under such conditions shows that the reason the EPP gets smaller is that lowering Ca^{2+} concentration decreases the number of vesicles that fuse with the plasma membrane of the terminal. An important insight into how Ca^{2+} regulates the fusion of synaptic vesicles was the discovery that presynaptic terminals have voltage-sensitive Ca^{2+} channels in their plasma membranes (see Chapter 4).

The first indication of presynaptic Ca^{2+} channels was provided by Katz and Ricardo Miledi. They observed that presynaptic terminals treated with tetrodotoxin (which blocks Na^+ channels; see Chapter 3) could still produce a peculiarly prolonged type of action potential. The explanation for this surprising finding was that current was still flowing through Ca^{2+} channels, substituting for the current ordinarily carried by the blocked Na^+ channels. Subsequent voltage clamp experiments, performed by Rodolfo Llinás and others at a giant presynaptic terminal of the squid (Figure 5.10A), confirmed

Figure 5.10 The entry of Ca^{2+} through the specific voltage-dependent calcium channels in the presynaptic terminals causes transmitter release. (A) Experimental setup using an extraordinarily large synapse in the squid. The voltage clamp method detects currents flowing across the presynaptic membrane when the membrane potential is depolarized. (B) Pharmacological agents that block currents flowing through Na^+ and K^+ channels reveal a remaining inward current flowing through Ca^{2+} channels. This influx of calcium triggers transmitter secretion, as indicated by a change in the postsynaptic membrane potential. Treatment of the same presynaptic terminal with cadmium, a calcium channel blocker, eliminates both the presynaptic calcium current and the postsynaptic response. (After Augustine and Eckert, 1984.)



Box B

Diseases That Affect the Presynaptic Terminal

Various steps in the exocytosis and endocytosis of synaptic vesicles are targets of a number of rare but debilitating neurological diseases. Many of these are myasthenic syndromes, in which abnormal transmission at neuromuscular synapses leads to weakness and fatigability of skeletal muscles (see Box B in Chapter 7). One of the best-understood examples of such disorders is the Lambert-Eaton myasthenic syndrome (LEMS), an occasional complication in patients with certain kinds of cancers. Biopsies of muscle tissue removed from LEMS patients allow intracellular recordings identical to those shown in Figure 5.6. Such recordings have shown that when a motor neuron is stimulated, the number of quanta contained in individual EPPs is greatly reduced, although the amplitude of spontaneous MEPPs is normal. Thus, LEMS impairs evoked neurotransmitter release, but does not affect the size of individual quanta.

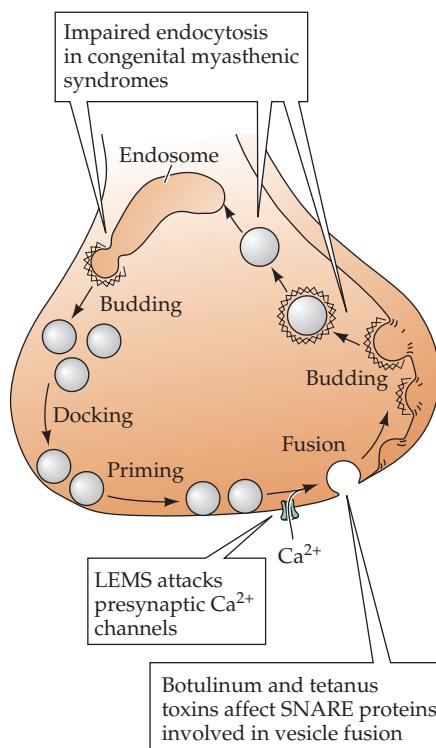
Several lines of evidence indicate that this reduction in neurotransmitter release is due to a loss of voltage-gated Ca^{2+} channels in the presynaptic terminal of motor neurons (see figure). Thus, the defect in neuromuscular transmission can be overcome by increasing the extracellular concentration of Ca^{2+} , and anatomical studies indicate a lower density of Ca^{2+} channel proteins in the presynaptic plasma membrane. The loss of presynaptic Ca^{2+} channels in LEMS apparently arises from a defect in the immune system. The blood of LEMS patients has a very high concentration of antibodies that bind to Ca^{2+} channels, and it seems likely that these antibodies are the primary cause of LEMS. For example, removal of Ca^{2+} channel antibodies from the blood of LEMS patients by plasma exchange reduces muscle weakness. Similarly, immunosuppressive drugs also can alleviate LEMS

symptoms. Perhaps most telling, injecting these antibodies into experimental animals elicits muscle weakness and abnormal neuromuscular transmission. Why the immune system generates antibodies against Ca^{2+} channels is not clear. Most LEMS patients have small-cell carcinoma, a form of lung cancer that may somehow initiate the immune response to Ca^{2+} channels. Whatever the origin, the binding of antibodies to Ca^{2+} channels causes a reduction in Ca^{2+} channel currents. It is this antibody-induced defect in presynaptic Ca^{2+} entry that accounts for the muscle weakness associated with LEMS.

Congenital myasthenic syndromes are genetic disorders that also cause muscle weakness by affecting neuromuscular transmission. Some of these syndromes affect the acetylcholinesterase that degrades acetylcholine in the synaptic cleft, whereas others arise from autoimmune attack of acetylcholine receptors (see Box C in Chapter 6). However, a number of congenital myasthenic syndromes arise from defects in acetylcholine release due to altered synaptic vesicle traffic within the motor neuron terminal. Neuromuscular synapses in some of these patients have EPPs with reduced quantal content, a deficit that is especially prominent when the synapse is activated repeatedly. Electron microscopy shows that presynaptic motor nerve terminals have a greatly reduced number of synaptic vesicles. The defect in neurotransmitter release evidently results from an inadequate number of synaptic vesicles available for release during sustained presynaptic activity. The origins of this shortage of synaptic vesicles is not clear, but could result either from an impairment in endocytosis in the nerve terminal (see figure) or from a reduced supply of vesicles from the motor neuron cell body.

Still other patients suffering from familial infantile myasthenia appear to have neuromuscular weakness that arises from reductions in the size of individual quanta, rather than the number of quanta released. Motor nerve terminals from these patients have synaptic vesicles that are normal in number, but smaller in diameter. This finding suggests a different type of genetic lesion that somehow alters formation of new synaptic vesicles following endocytosis, thereby leading to less acetylcholine in each vesicle.

Another disorder of synaptic transmitter release results from poisoning by anaerobic *Clostridium* bacteria. This genus of microorganisms produces some



Presynaptic targets of several neurological disorders.

of the most potent toxins known, including several botulinum toxins and tetanus toxin. Both botulism and tetanus are potentially deadly disorders.

Botulism can occur by consuming food containing *Clostridium* bacteria or by infection of wounds with the spores of these ubiquitous organisms. In either case, the presence of the toxin can cause paralysis of peripheral neuromuscular synapses due to abolition of neurotransmitter release. This interference with neuromuscular transmission causes skeletal muscle weakness, in extreme cases producing respiratory failure due to paralysis of the diaphragm and other muscles required for breathing. Botulinum toxins also block synapses innervating the smooth muscles of several organs, giving rise to visceral motor dysfunction.

Tetanus typically results from the contamination of puncture wounds by

Clostridium bacteria that produce tetanus toxin. In contrast to botulism, tetanus poisoning blocks the release of inhibitory transmitters from interneurons in the spinal cord. This effect causes a loss of synaptic inhibition on spinal motor neurons, producing hyperexcitation of skeletal muscle and tetanic contractions in affected muscles (hence the name of the disease).

Although their clinical consequences are dramatically different, clostridial toxins have a common mechanism of action (see figure). Tetanus toxin and botulinum toxins work by cleaving the SNARE proteins involved in fusion of synaptic vesicles with the presynaptic plasma membrane (see Box C). This proteolytic action presumably accounts for the block of transmitter release at the afflicted synapses. The different actions of these toxins on synaptic transmission at excitatory motor versus inhibitory synapses appar-

ently results from the fact that these toxins are taken up by different types of neurons: Whereas the botulinum toxins are taken up by motor neurons, tetanus toxin is preferentially targeted to interneurons. The basis for this differential uptake of toxins is not known, but presumably arises from the presence of different types of toxin receptors on the two types of neurons.

References

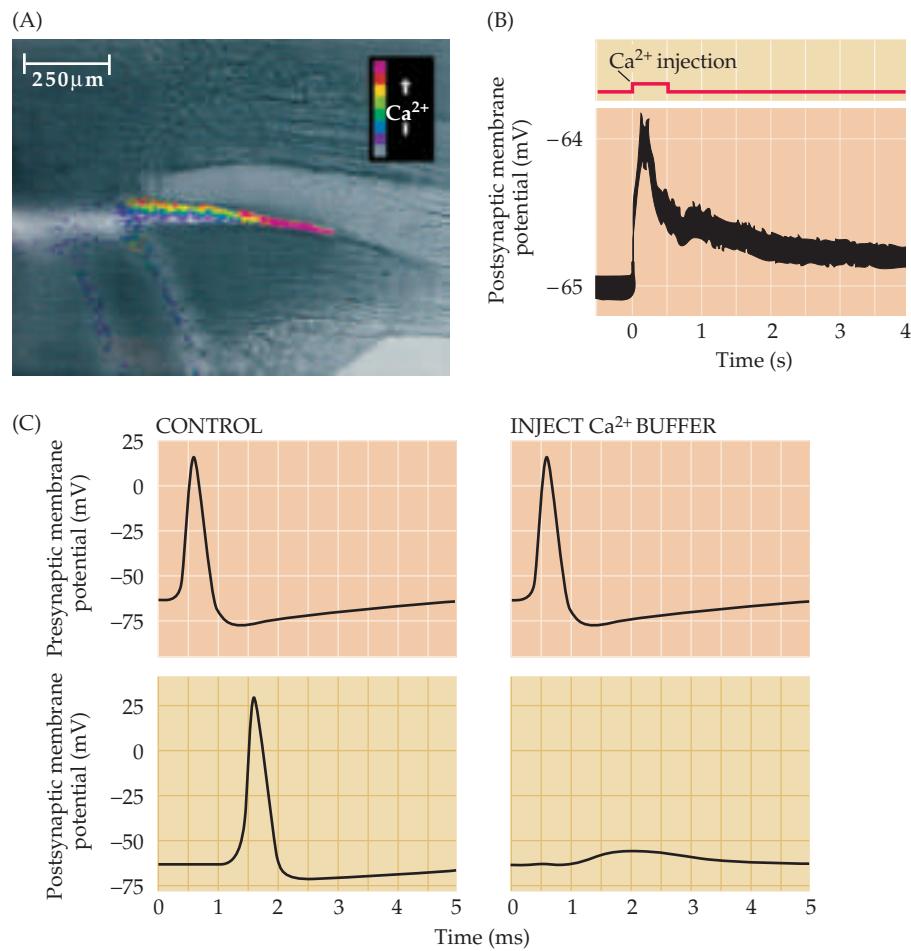
- ENGEL, A. G. (1991) Review of evidence for loss of motor nerve terminal calcium channels in Lambert-Eaton myasthenic syndrome. Ann. N.Y. Acad. Sci. 635: 246–258.
- ENGEL, A. G. (1994) Congenital myasthenic syndromes. Neurol. Clin. 12: 401–437.
- LANG, B. AND A. VINCENT (2003) Autoantibodies to ion channels at the neuromuscular junction. Autoimmun. Rev. 2: 94–100.
- MASELLI, R. A. (1998) Pathogenesis of human botulism. Ann. N.Y. Acad. Sci. 841: 122–139.

the presence of voltage-gated Ca^{2+} channels in the presynaptic terminal (Figure 5.10B). Such experiments showed that the amount of neurotransmitter released is very sensitive to the exact amount of Ca^{2+} that enters. Further, blockade of these Ca^{2+} channels with drugs also inhibits transmitter release (Figure 5.10B, right). These observations all confirm that the voltage-gated Ca^{2+} channels are directly involved in neurotransmission. Thus, presynaptic action potentials open voltage-gated Ca^{2+} channels, with a resulting influx of Ca^{2+} .

That Ca^{2+} entry into presynaptic terminals causes a rise in the concentration of Ca^{2+} within the terminal has been documented by microscopic imaging of terminals filled with Ca^{2+} -sensitive fluorescent dyes (Figure 5.11A). The consequences of the rise in presynaptic Ca^{2+} concentration for neurotransmitter release has been directly shown in two ways. First, microinjection of Ca^{2+} into presynaptic terminals triggers transmitter release in the absence of presynaptic action potentials (Figure 5.11B). Second, presynaptic microinjection of calcium chelators (chemicals that bind Ca^{2+} and keep its concentration buffered at low levels) prevents presynaptic action potentials from causing transmitter secretion (Figure 5.11C). These results prove beyond any doubt that a rise in presynaptic Ca^{2+} concentration is both necessary and sufficient for neurotransmitter release. Thus, as is the case for many other forms of neuronal signaling (see Chapter 7), Ca^{2+} serves as a second messenger during transmitter release.

While Ca^{2+} is a universal trigger for transmitter release, not all transmitters are released with the same speed. For example, while secretion of ACh

Figure 5.11 Evidence that a rise in presynaptic Ca^{2+} concentration triggers transmitter release from presynaptic terminals. (A) Fluorescence microscopy measurements of presynaptic Ca^{2+} concentration at the squid giant synapse (see Figure 5.8A). A train of presynaptic action potentials causes a rise in Ca^{2+} concentration, as revealed by a dye (called fura-2) that fluoresces more strongly when the Ca^{2+} concentration increases. (B) Microinjection of Ca^{2+} into a squid giant presynaptic terminal triggers transmitter release, measured as a depolarization of the postsynaptic membrane potential. (C) Microinjection of BAPTA, a Ca^{2+} chelator, into a squid giant presynaptic terminal prevents transmitter release. (A from Smith et al., 1993; B after Miledi, 1971; C after Adler et al., 1991.)



from motor neurons requires only a fraction of a millisecond (see Figure 5.6), release of neuropeptides require high-frequency bursts of action potentials for many seconds. These differences in the rate of release probably arise from differences in the spatial arrangement of vesicles relative to presynaptic Ca^{2+} channels. This perhaps is most evident in cases where small molecules and peptides serve as co-transmitters (Figure 5.12). Whereas the small, clear-core vesicles containing small-molecule transmitters are typically docked at the plasma membrane in advance of Ca^{2+} entry, large dense core vesicles containing peptide transmitters are farther away from the plasma membrane (see Figure 5.5D). At low firing frequencies, the concentration of Ca^{2+} may increase only locally at the presynaptic plasma membrane, in the vicinity of open Ca^{2+} channels, limiting release to small-molecule transmitters from the docked small, clear-core vesicles. Prolonged high-frequency stimulation increases the Ca^{2+} concentration throughout the presynaptic terminal, thereby inducing the slower release of neuropeptides.

Molecular Mechanisms of Transmitter Secretion

Precisely how an increase in presynaptic Ca^{2+} concentration goes on to trigger vesicle fusion and neurotransmitter release is not understood. However, many important clues have come from molecular studies that have identified and characterized the proteins found on synaptic vesicles and their binding

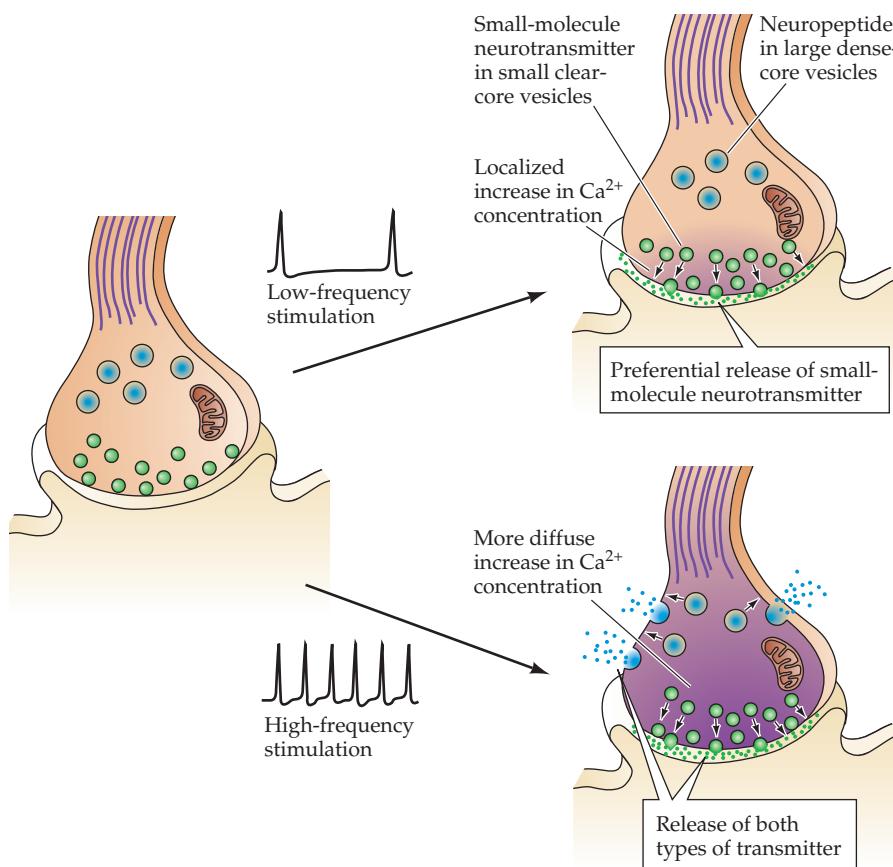


Figure 5.12 Differential release of neuropeptide and small-molecule co-transmitters. Low-frequency stimulation preferentially raises the Ca^{2+} concentration close to the membrane, favoring the release of transmitter from small clear-core vesicles docked at presynaptic specializations. High-frequency stimulation leads to a more general increase in Ca^{2+} , causing the release of peptide neurotransmitters from large dense-core vesicles, as well as small-molecule neurotransmitters from small clear-core vesicles.

partners on the presynaptic plasma membrane and cytoplasm (Figure 5.13). Most, if not all, of these proteins act at one or more steps in the synaptic vesicle cycle. Although a complete molecular picture of neurotransmitter release is still lacking, the roles of several proteins involved in vesicle fusion have been deduced.

Several of the proteins important for neurotransmitter release are also involved in other types of membrane fusion events common to all cells. For example, two proteins originally found to be important for the fusion of vesicles with membranes of the Golgi apparatus, the ATPase **NSF** (**NEM-sensitive fusion protein**) and **SNAPs** (**soluble NSF-attachment proteins**), are also involved in priming synaptic vesicles for fusion. These two proteins work by regulating the assembly of other proteins that are called **SNAREs** (**SNAP receptors**). One of these SNARE proteins, **synaptobrevin**, is in the membrane of synaptic vesicles, while two other SNARE proteins called **syntaxin** and **SNAP-25** are found primarily on the plasma membrane. These SNARE proteins can form a macromolecular complex that spans the two membranes, thus bringing them into close apposition (Figure 5.14A). Such an arrangement is well suited to promote the fusion of the two membranes, and several lines of evidence suggest that this is what actually occurs. One important observation is that toxins that cleave the SNARE proteins block neurotransmitter release (Box C). In addition, putting SNARE proteins into artificial lipid membranes and allowing these proteins to form complexes with each other causes the membranes to fuse. Many other proteins, such as

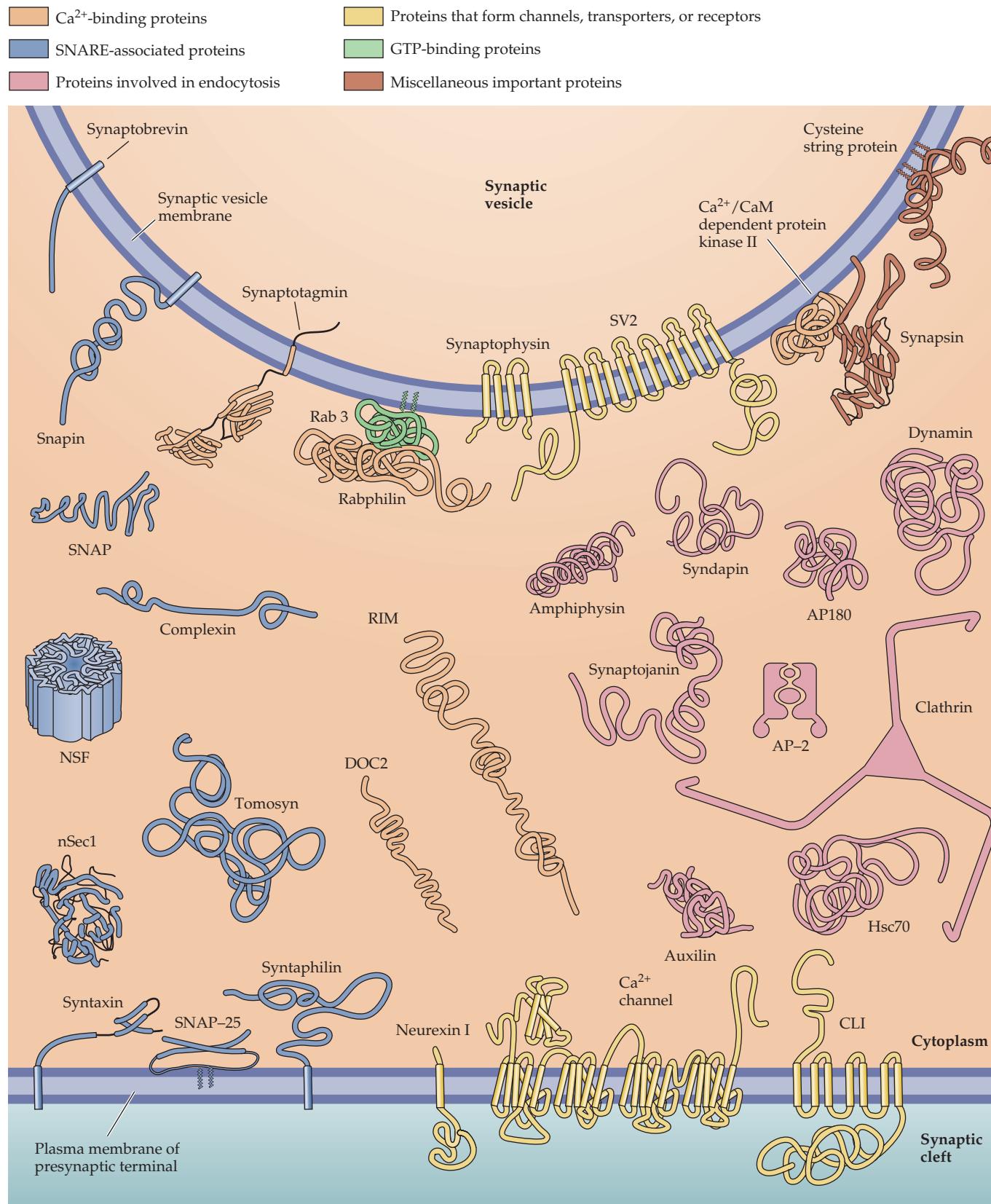


Figure 5.13 Presynaptic proteins implicated in neurotransmitter release. Structures adapted from Brunger (2001) and Brodsky et al. (2001).

complexin, nSec-1, snapin, syntaphilin, and tomosyn, bind to the SNAREs and presumably regulate the formation or disassembly of this complex.

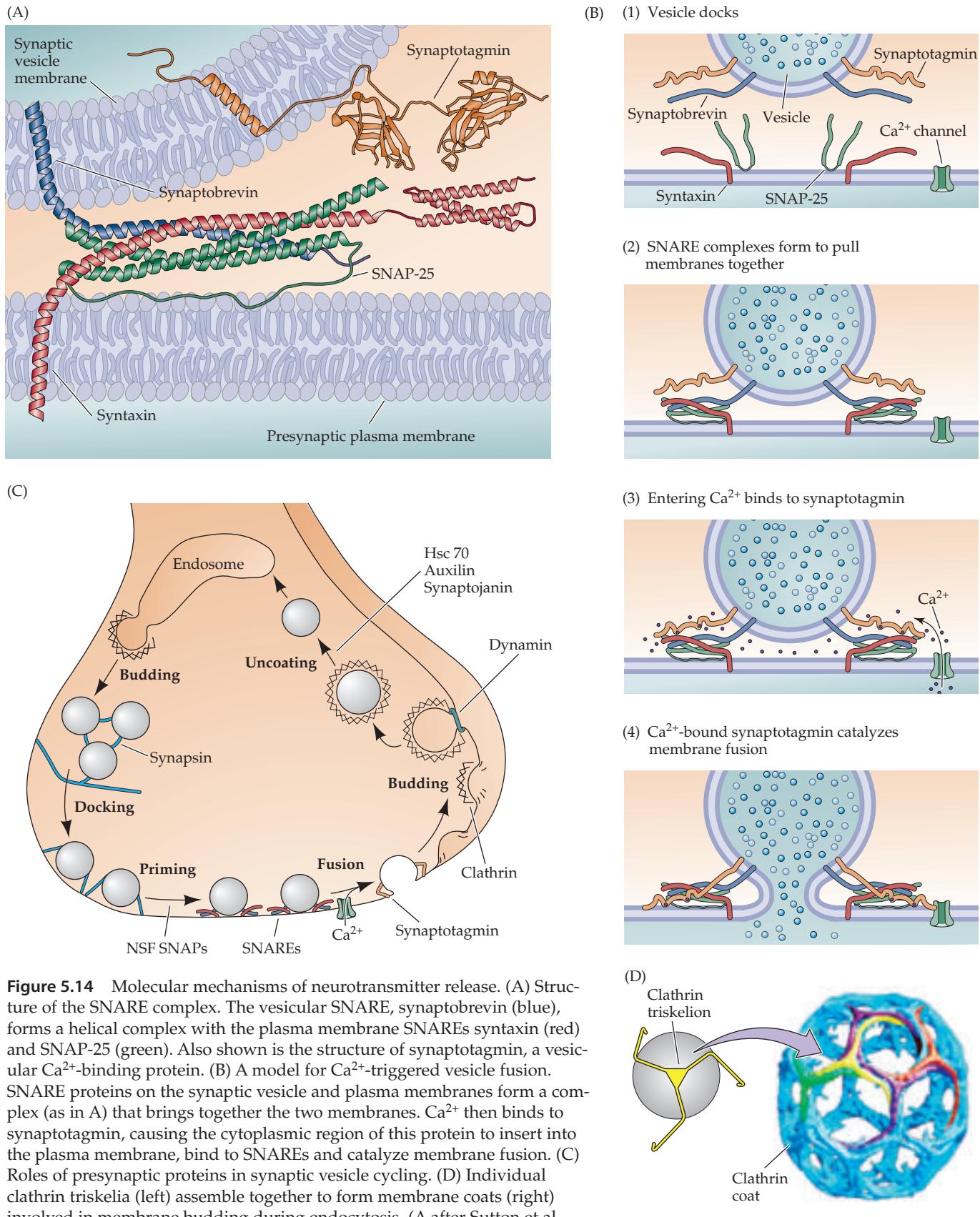
Because the SNARE proteins do not bind Ca^{2+} , still other molecules must be responsible for Ca^{2+} regulation of neurotransmitter release. Several presynaptic proteins, including calmodulin, CAPS, and munc-13, are capable of binding Ca^{2+} . However, the leading candidate for Ca^{2+} regulation of neurotransmitter release is **synaptotagmin**, a protein found in the membrane of synaptic vesicles. Synaptotagmin binds Ca^{2+} at concentrations similar to those required to trigger vesicle fusion within the presynaptic terminal. It may act as a Ca^{2+} sensor, signaling the elevation of Ca^{2+} within the terminal and thus triggering vesicle fusion. In support of this idea, alterations of the properties of synaptotagmin in the presynaptic terminals of mice, fruit flies, squid, and other experimental animals impair Ca^{2+} -dependent neurotransmitter release. In fact, deletion of only one of the 19 synaptotagmin genes of mice is a lethal mutation, causing the mice to die soon after birth. How Ca^{2+} binding to synaptotagmin could lead to exocytosis is not yet clear. It is known that Ca^{2+} changes the chemical properties of synaptotagmin, allowing it to insert into membranes and to bind to other proteins, including the SNAREs. A plausible model is that the SNARE proteins bring the two membranes close together, and that Ca^{2+} -induced changes in synaptotagmin then produce the final fusion of these membranes (Figure 5.14B).

Still other proteins appear to be involved at subsequent steps of the synaptic vesicle cycle (Figure 5.14C). For example, the protein **clathrin** is involved in endocytotic budding of vesicles from the plasma membrane. Clathrin forms structures that resemble geodesic domes (Figure 5.14D); these structures form coated pits that initiate membrane budding. Assembly of individual clathrin triskelia (so named because of their 3-legged appearance) into coats is aided by several other accessory proteins, such as AP2, AP180 and amphiphysin. The coats increase the curvature of the budding membrane until it forms a coated vesicle-like structure. Another protein, called **dynamin**, is at least partly responsible for the final pinching-off of membrane to convert the coated pits into coated vesicles. The coats are then removed by an ATPase, **Hsc70**, with another protein called **auxilin** serving as a co-factor. Other proteins, such as **synaptojanin**, are also important for vesicle uncoating. Several lines of evidence indicate that the protein **synapsin**, which reversibly binds to synaptic vesicles, may cross-link newly formed vesicles to the cytoskeleton to keep the vesicles tethered within the reserve pool. Mobilization of these reserve pool vesicles is caused by phosphorylation of synapsin by proteins kinases (Chapter 7), which allows synapsin to dissociate from the vesicles, thus freeing the vesicles to make their way to the plasma membrane.

In summary, a complex cascade of proteins, acting in a defined temporal and spatial order, allows neurons to secrete transmitters. Although the detailed mechanisms responsible for transmitter secretion are not completely clear, rapid progress is being made toward this goal.

Neurotransmitter Receptors

The generation of postsynaptic electrical signals is also understood in considerable depth. Such studies began in 1907, when the British physiologist John N. Langley introduced the concept of **receptor molecules** to explain the specific and potent actions of certain chemicals on muscle and nerve cells. Much subsequent work has shown that receptor molecules do indeed account for the ability of neurotransmitters, hormones, and drugs to alter the



Box C

Toxins That Affect Transmitter Release

Several important insights about the molecular basis of neurotransmitter secretion have come from analyzing the actions of a series of biological toxins produced by a fascinating variety of organisms. One family of such agents is the clostridial toxins responsible for botulism and tetanus (see Box B). Clever and patient biochemical work has shown that these toxins are highly specific proteases that cleave presynaptic SNARE proteins (see figure). Tetanus toxin and botulinum toxin (types B, D, F, and G) specifically cleave the vesicle SNARE protein, synaptobrevin. Other botulinum toxins are proteases that cleave syntaxin (type C) and SNAP-25 (types A and E), SNARE proteins found on the presynaptic plasma membrane. Destruction of these presynaptic proteins is the basis for the actions of the toxins on neurotransmitter release. The evidence described in the text also implies that these three syn-

aptic SNARE proteins are somehow important in the process of vesicle–plasma membrane fusion.

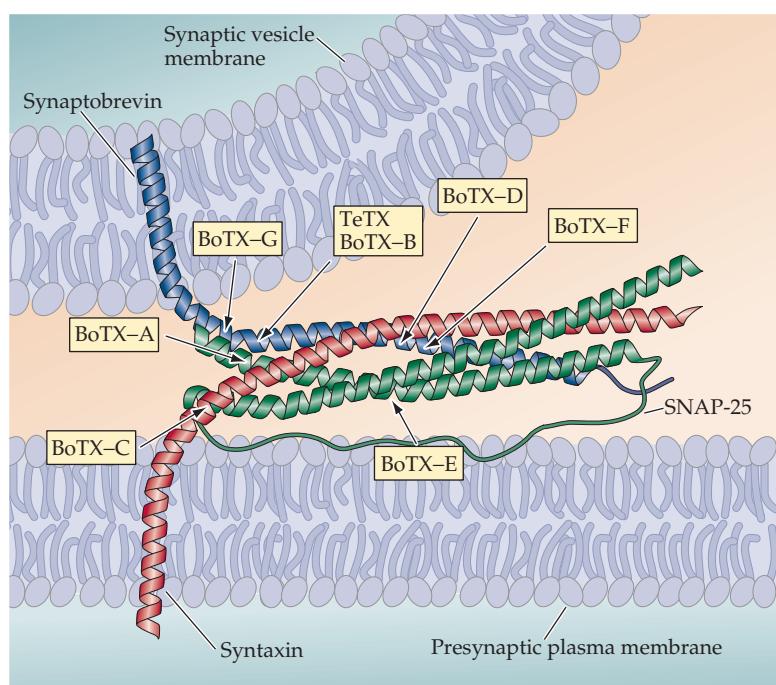
Another toxin that targets neurotransmitter release is α -latrotoxin, a protein found in the venom of the female black widow spider. Application of this molecule to neuromuscular synapses causes a massive discharge of synaptic vesicles, even when Ca^{2+} is absent from the extracellular medium. While it is not yet clear how this toxin triggers Ca^{2+} -independent exocytosis, α -latrotoxin binds to two different types of presynaptic proteins that may mediate its actions. One group of binding partners for α -latrotoxin is the neurexins, a group of integral membrane proteins found in presynaptic terminals (see Figure 5.13). Several lines of evidence implicate binding to neurexins in at least some of the actions of α -latrotoxin. Because the neurexins bind to synaptotagmin, a vesicular Ca^{2+} -binding

protein that is known to be important in exocytosis, this interaction may allow α -latrotoxin to bypass the usual Ca^{2+} requirement for triggering vesicle fusion. Another type of presynaptic protein that can bind to α -latrotoxin is called CL1 (based on its previous names, Ca^{2+} -independent receptor for latrotoxin and *latrophilin-1*). CL1 is a relative of the G-protein-coupled receptors that mediate the actions of neurotransmitters and other extracellular chemical signals (see Chapter 7). Thus, the binding of α -latrotoxin to CL1 is thought to activate an intracellular signal transduction cascade that may be involved in the Ca^{2+} -independent actions of α -latrotoxin. While more work is needed to establish the roles of neurexins and CL1 in the actions of α -latrotoxin definitively, effects on these two proteins probably account for the potent presynaptic actions of this toxin.

Still other toxins produced by snakes, snails, spiders, and other predatory animals are known to affect transmitter release, but their sites of action have yet to be identified. Based on the precedents described here, it is likely that these biological poisons will continue to provide valuable tools for elucidating the molecular basis of neurotransmitter release, just as they will continue to enable the predators to feast on their prey.

References

- KRASNOPOREV, V. G. AND 10 OTHERS (1997) α -Latrotoxin stimulates exocytosis by the interaction with a neuronal G-protein-coupled receptor. *Neuron* 18: 925–937.
- MONTECUCCO, C. AND G. SCHIAVO (1994) Mechanism of action of tetanus and botulinum neurotoxins. *Mol. Microbiol.* 13: 1–8.
- SCHIAVO, G., M. MATTEOLI AND C. MONTECUCCO (2000) Neurotoxins affecting neuro-exocytosis. *Physiol. Rev.* 80: 717–766.
- SUGITA, S., M. KHVOCHTEV AND T. C. SUDHOF (1999) Neurexins are functional α -latrotoxin receptors. *Neuron* 22: 489–496.



Cleavage of SNARE proteins by clostridial toxins. Indicated are the sites of proteolysis by tetanus toxin (TeTX) and various types of botulinum toxin (BoTX). (After Sutton et al., 1998.)

functional properties of neurons. While it has been clear since Langley's day that receptors are important for synaptic transmission, their identity and detailed mechanism of action remained a mystery until quite recently. It is now known that neurotransmitter receptors are proteins embedded in the plasma membrane of postsynaptic cells. Domains of receptor molecules that extend into the synaptic cleft bind neurotransmitters that are released into this space by the presynaptic neuron. The binding of neurotransmitters, either directly or indirectly, causes ion channels in the postsynaptic membrane to open or close. Typically, the resulting ion fluxes change the membrane potential of the postsynaptic cell, thus mediating the transfer of information across the synapse.

Postsynaptic Membrane Permeability Changes during Synaptic Transmission

Just as studies of the neuromuscular synapse paved the way for understanding neurotransmitter release mechanisms, this peripheral synapse has been equally valuable for understanding the mechanisms that allow neurotransmitter receptors to generate postsynaptic signals. The binding of ACh to postsynaptic receptors opens ion channels in the muscle fiber membrane. This effect can be demonstrated directly by using the patch clamp method (see Box A in Chapter 4) to measure the minute postsynaptic currents that flow when two molecules of individual ACh bind to receptors, as Erwin Neher and Bert Sakmann first did in 1976. Exposure of the extracellular surface of a patch of postsynaptic membrane to ACh causes single-channel currents to flow for a few milliseconds (Figure 5.15A). This shows that ACh binding to its receptors opens ligand-gated ion channels, much in the way that changes in membrane potential open voltage-gated ion channels (Chapter 4).

The electrical actions of ACh are greatly multiplied when an action potential in a presynaptic motor neuron causes the release of millions of molecules of ACh into the synaptic cleft. In this more physiological case, the transmitter molecules bind to many thousands of ACh receptors packed in a dense array on the postsynaptic membrane, transiently opening a very large number of postsynaptic ion channels. Although individual ACh receptors only open briefly, (Figure 5.15B1), the opening of a large number of channels is synchronized by the brief duration during which ACh is secreted from pre-synaptic terminals (Figure 5.15B2,3). The macroscopic current resulting from the summed opening of many ion channels is called the **end plate current**, or **EPC**. Because the current flowing during the EPC is normally inward, it causes the postsynaptic membrane potential to depolarize. This depolarizing change in potential is the EPP (Figure 5.15C), which typically triggers a postsynaptic action potential by opening voltage-gated Na^+ and K^+ channels (see Figure 5.6B).

The identity of the ions that flow during the EPC can be determined via the same approaches used to identify the roles of Na^+ and K^+ fluxes in the currents underlying action potentials (Chapter 3). Key to such an analysis is identifying the membrane potential at which no current flows during transmitter action. When the potential of the postsynaptic muscle cell is controlled by the voltage clamp method (Figure 5.16A), the magnitude of the membrane potential clearly affects the amplitude and polarity of EPCs (Figure 5.16B). Thus, when the postsynaptic membrane potential is made more negative than the resting potential, the amplitude of the EPC becomes larger, whereas this current is reduced when the membrane potential is made more positive. At approximately 0 mV, no EPC is detected, and at even more positive poten-

tials, the current reverses its polarity, becoming outward rather than inward (Figure 5.16C). The potential where the EPC reverses, about 0 mV in the case of the neuromuscular junction, is called the **reversal potential**.

As was the case for currents flowing through voltage-gated ion channels (see Chapter 3), the magnitude of the EPC at any membrane potential is given by the product of the ionic conductance activated by ACh (g_{ACh}) and the electrochemical driving force on the ions flowing through ligand-gated channels. Thus, the value of the EPC is given by the relationship

$$EPC = g_{ACh}(V_m - E_{rev})$$

where E_{rev} is the reversal potential for the EPC. This relationship predicts that the EPC will be an inward current at potentials more negative than E_{rev} because the electrochemical driving force, $V_m - E_{rev}$, is a negative number. Further, the EPC will become smaller at potentials approaching E_{rev} because the driving force is reduced. At potentials more positive than E_{rev} , the EPC is outward because the driving force is reversed in direction (that is, positive). Because the channels opened by ACh are largely insensitive to membrane voltage, g_{ACh} will depend only on the number of channels opened by ACh, which depends in turn on the concentration of ACh in the synaptic cleft.

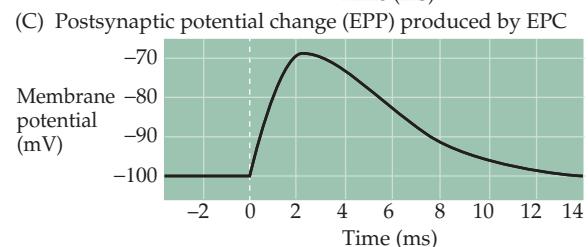
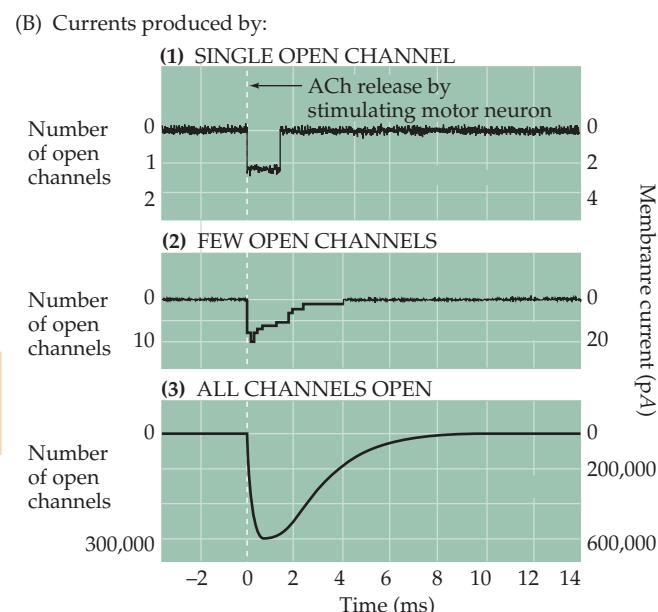
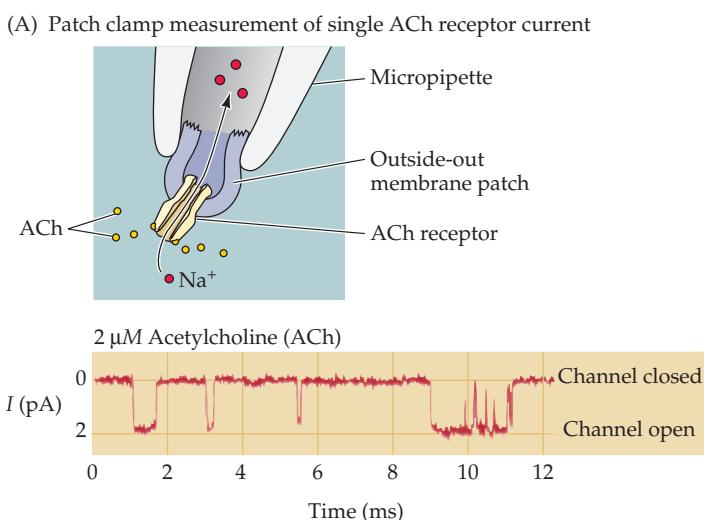
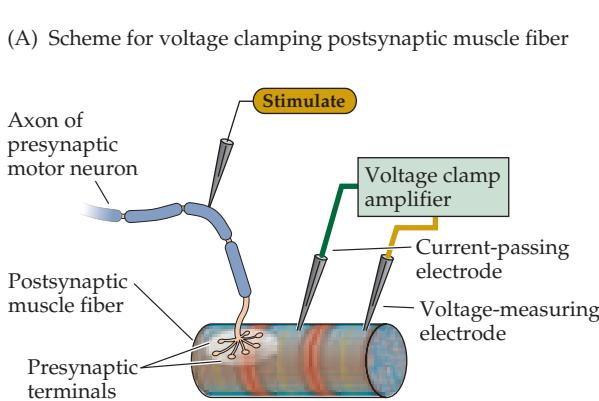


Figure 5.15 Activation of ACh receptors at neuromuscular synapses. (A) Outside-out patch clamp measurement of single ACh receptor currents from a patch of membrane removed from the postsynaptic muscle cell. When ACh is applied to the extracellular surface of the membrane clamped at negative voltages, the repeated brief opening of a single channel can be seen as downward deflections corresponding to inward current (i.e., positive ions flowing into the cell). (B) Synchronized opening of many ACh-activated channels at a synapse being voltage-clamped at negative voltages. (1) If a single channel is examined during the release of ACh from the presynaptic terminal, the channel opens transiently. (2) If a number of channels are examined together, ACh release opens the channels almost synchronously. (3) The opening of a very large number of postsynaptic channels produces a macroscopic EPC. (C) In a normal muscle cell (i.e., not being voltage-clamped), the inward EPC depolarizes the postsynaptic muscle cell, giving rise to an EPP. Typically, this depolarization generates an action potential (not shown).

Thus, the magnitude and polarity of the postsynaptic membrane potential determines the direction and amplitude of the EPC solely by altering the driving force on ions flowing through the receptor channels opened by ACh.

When V_m is at the reversal potential, $V_m - E_{rev}$ is equal to 0 and there is no net driving force on the ions that can permeate the receptor-activated channel. As a result, the identity of the ions that flow during the EPC can be deduced by observing how the reversal potential of the EPC compares to the equilibrium potential for various ion species (Figure 5.17). For example, if ACh were to open an ion channel permeable only to K^+ , then the reversal



(B) Effect of membrane voltage on postsynaptic end plate currents

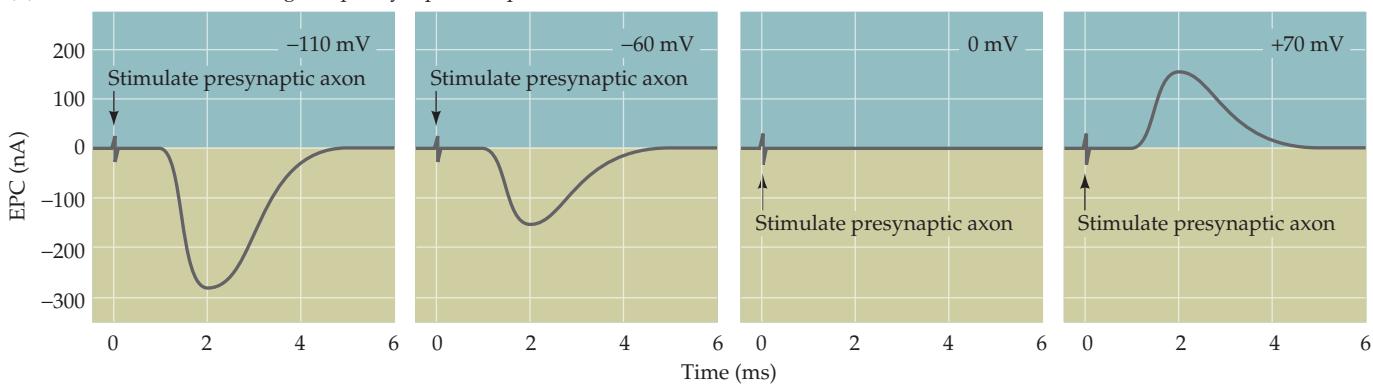


Figure 5.16 The influence of the postsynaptic membrane potential on end plate currents. (A) A postsynaptic muscle fiber is voltage clamped using two electrodes, while the presynaptic neuron is electrically stimulated to cause the release of ACh from presynaptic terminals. This experimental arrangement allows the recording of macroscopic EPCs produced by ACh. (B) Amplitude and time course of EPCs generated by stimulating the presynaptic motor neuron while the postsynaptic cell is voltage clamped at four different membrane potentials. (C) The relationship between the peak amplitude of EPCs and postsynaptic membrane potential is nearly linear, with a reversal potential (the voltage at which the direction of the current changes from inward to outward) close to 0 mV. Also indicated on this graph are the equilibrium potentials of Na^+ , K^+ , and Cl^- ions. (D) Lowering the external Na^+ concentration causes EPCs to reverse at more negative potentials. (E) Raising the external K^+ concentration makes the reversal potential more positive. (After Takeuchi and Takeuchi, 1960.)

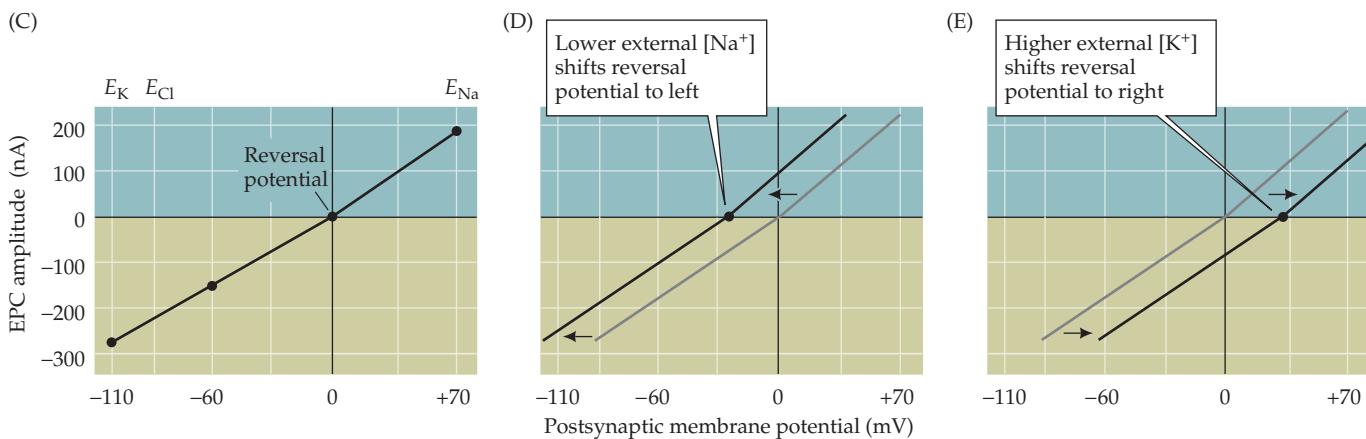


Figure 5.17 The effect of ion channel selectivity on the reversal potential. Voltage clamping a postsynaptic cell while activating presynaptic neurotransmitter release reveals the identity of the ions permeating the postsynaptic receptors being activated. (A) The activation of postsynaptic channels permeable only to K^+ results in currents reversing at E_K , near -100 mV. (B) The activation of postsynaptic Na^+ channels results in currents reversing at E_{Na} , near $+70$ mV. (C) Cl^- -selective currents reverse at E_{Cl} , near -50 mV. (D) Ligand-gated channels that are about equally permeable to both K^+ and Na^+ show a reversal potential near 0 mV.

potential of the EPC would be at the equilibrium potential for K^+ , which for a muscle cell is close to -100 mV (Figure 5.17A). If the ACh-activated channels were permeable only to Na^+ , then the reversal potential of the current would be approximately $+70$ mV, the Na^+ equilibrium potential of muscle cells (Figure 5.17B); if these channels were permeable only to Cl^- , then the reversal potential would be approximately -50 mV (Figure 5.17C). By this reasoning, ACh-activated channels cannot be permeable to only one of these ions, because the reversal potential of the EPC is not near the equilibrium potential for any of them (see Figure 5.16C). However, if these channels were permeable to both Na^+ and K^+ , then the reversal potential of the EPC would be between $+70$ mV and -100 mV (Figure 5.17D).

The fact that EPCs reverse at approximately 0 mV is therefore consistent with the idea that ACh-activated ion channels are almost equally permeable to both Na^+ and K^+ . This was tested in 1960, by the husband and wife team of Akira and Noriko Takeuchi, by altering the extracellular concentration of these two ions. As predicted, the magnitude and reversal potential of the EPC was changed by altering the concentration gradient of each ion. Lowering the external Na^+ concentration, which makes E_{Na} more negative, produces a negative shift in E_{rev} (Figure 5.16D), whereas elevating external K^+ concentration, which makes E_K more positive, causes E_{rev} to shift to a more positive potential (Figure 5.16E). Such experiments confirm that the ACh-activated ion channels are in fact permeable to both Na^+ and K^+ .

Even though the channels opened by the binding of ACh to its receptors are permeable to both Na^+ and K^+ , at the resting membrane potential the EPC is generated primarily by Na^+ influx (Figure 5.18). If the membrane potential is kept at E_K , the EPC arises entirely from an influx of Na^+ because at this potential there is no driving force on K^+ (Figure 5.18A). At the usual muscle fiber resting membrane potential of -90 mV, there is a small driving force on K^+ , but a much greater one on Na^+ . Thus, during the EPC, much more Na^+ flows into the muscle cell than K^+ flows out (Figure 5.18B); it is the net influx of positively charged Na^+ that constitutes the inward current measured as the EPC. At the reversal potential of about 0 mV, Na^+ influx and K^+ efflux are exactly balanced, so no current flows during the opening of channels by ACh binding (Figure 5.18C). At potentials more positive than E_{rev} the balance reverses; for example, at E_{Na} there is no influx of Na^+ and a large efflux of K^+ because of the large driving force on Na^+ (Figure 5.18D). Even more positive potentials cause efflux of both Na^+ and K^+ and produce an even larger outward EPC.

Were it possible to measure the EPP at the same time as the EPC (of course, the voltage clamp technique prevents this by keeping membrane potential constant), the EPP would be seen to vary in parallel with the amplitude and polarity of the EPC (Figures 5.18E,F). At the usual postsynaptic resting membrane potential of -90 mV, the large inward EPC causes the postsynaptic membrane potential to become more depolarized (see Figure

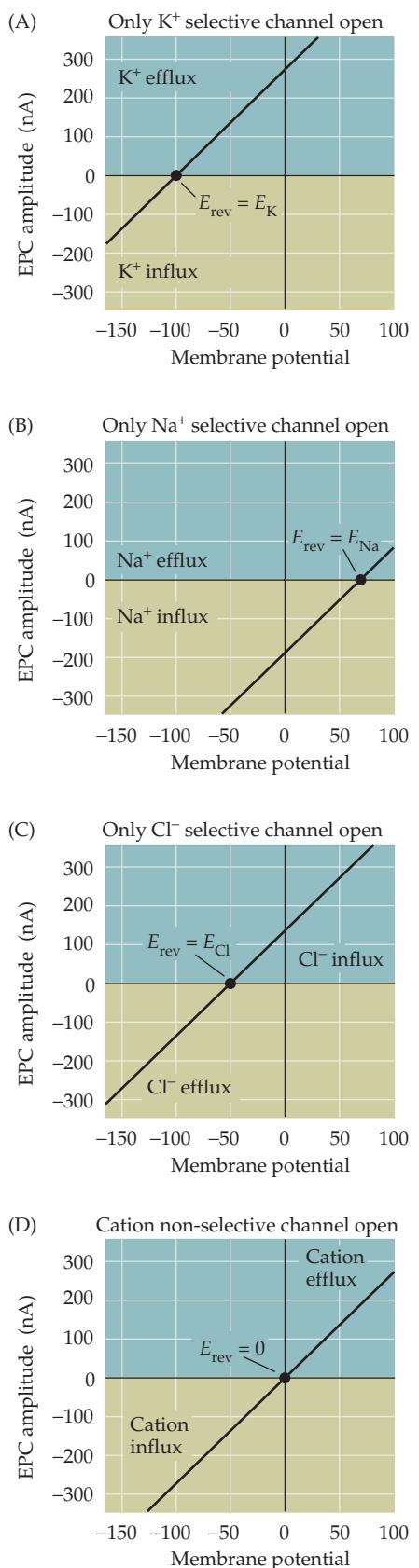
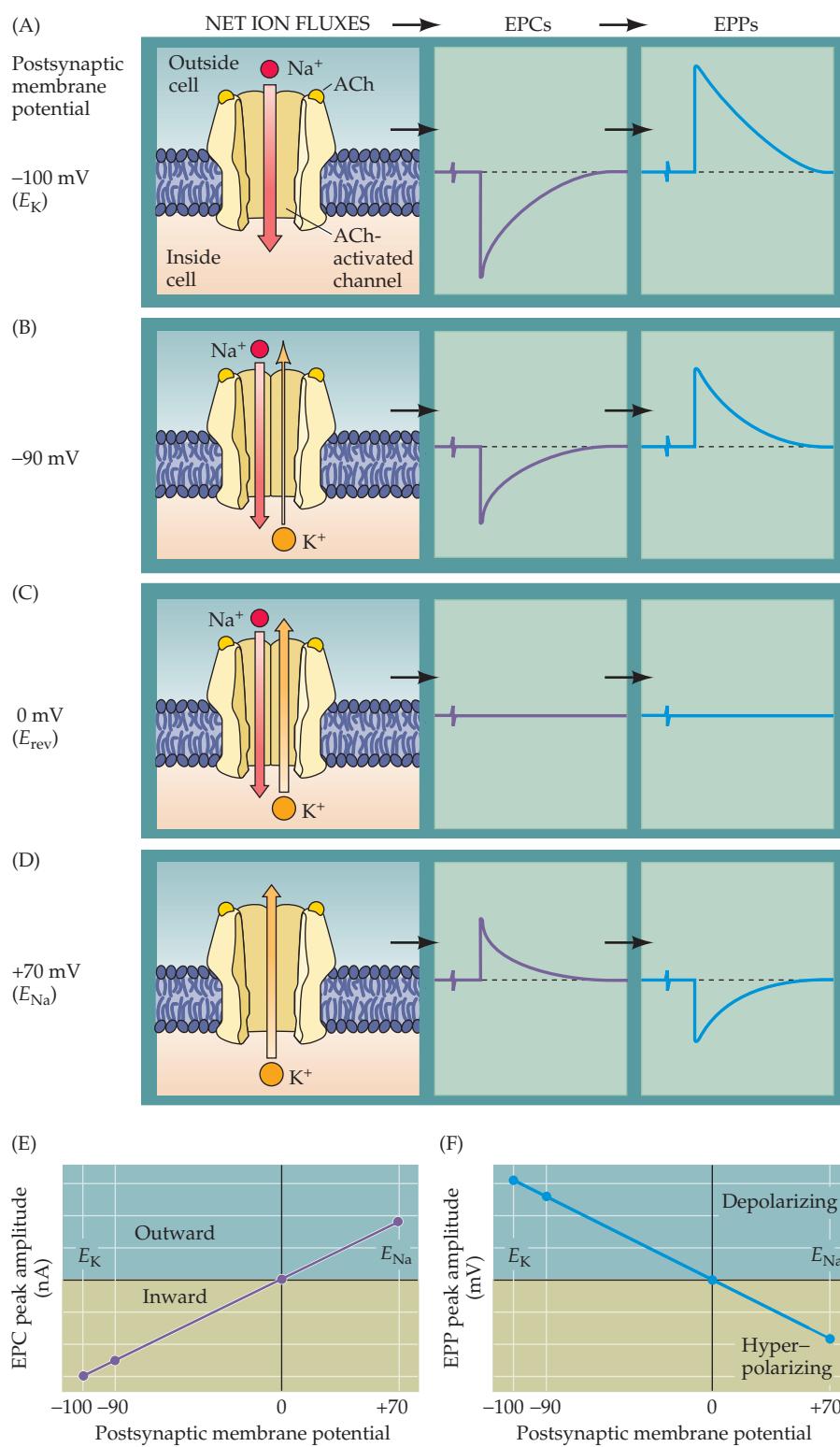


Figure 5.18 Na^+ and K^+ movements during EPCs and EPPs. (A–D) Each of the postsynaptic potentials (V_{post}) indicated at the left results in different relative fluxes of net Na^+ and K^+ (ion fluxes). These ion fluxes determine the amplitude and polarity of the EPCs, which in turn determine the EPPs. Note that at about 0 mV the Na^+ flux is exactly balanced by an opposite K^+ flux, resulting in no net current flow, and hence no change in the membrane potential. (E) EPCs are inward currents at potentials more negative than E_{rev} and outward currents at potentials more positive than E_{rev} . (F) EPPs depolarize the postsynaptic cell at potentials more negative than E_{rev} . At potentials more positive than E_{rev} , EPPs hyperpolarize the cell.



5.18F). However, at 0 mV, the EPP reverses its polarity, and at more positive potentials, the EPP is hyperpolarizing. Thus, the polarity and magnitude of the EPC depend on the electrochemical driving force, which in turn determines the polarity and magnitude of the EPP. EPPs will depolarize when the membrane potential is more negative than E_{rev} , and hyperpolarize when the membrane potential is more positive than E_{rev} . The general rule, then, is that

the action of a transmitter drives the postsynaptic membrane potential toward E_{rev} for the particular ion channels being activated.

Although this discussion has focused on the neuromuscular junction, similar mechanisms generate postsynaptic responses at all chemical synapses. The general principle is that transmitter binding to postsynaptic receptors produces a postsynaptic conductance change as ion channels are opened (or sometimes closed). The postsynaptic conductance is increased if—as at the neuromuscular junction—channels are opened, and decreased if channels are closed. This conductance change typically generates an electrical current, the **postsynaptic current (PSC)**, which in turn changes the postsynaptic membrane potential to produce a **postsynaptic potential (PSP)**. As in the specific case of the EPP at the neuromuscular junction, PSPs are depolarizing if their reversal potential is more positive than the postsynaptic membrane potential and hyperpolarizing if their reversal potential is more negative.

The conductance changes and the PSPs that typically accompany them are the ultimate outcome of most chemical synaptic transmission, concluding a sequence of electrical and chemical events that begins with the invasion of an action potential into the terminals of a presynaptic neuron. In many ways, the events that produce PSPs at synapses are similar to those that generate action potentials in axons; in both cases, conductance changes produced by ion channels lead to ionic current flow that changes the membrane potential (see Figure 5.18).

Excitatory and Inhibitory Postsynaptic Potentials

PSPs ultimately alter the probability that an action potential will be produced in the postsynaptic cell. At the neuromuscular junction, synaptic action increases the probability that an action potential will occur in the postsynaptic muscle cell; indeed, the large amplitude of the EPP ensures that an action potential always is triggered. At many other synapses, PSPs similarly increase the probability of firing a postsynaptic action potential. However, still other synapses actually *decrease* the probability that the postsynaptic cell will generate an action potential. PSPs are called **excitatory (or EPSPs)** if they increase the likelihood of a postsynaptic action potential occurring, and **inhibitory (or IPSPs)** if they decrease this likelihood. Given that most neurons receive inputs from both excitatory and inhibitory synapses, it is important to understand more precisely the mechanisms that determine whether a particular synapse excites or inhibits its postsynaptic partner.

The principles of excitation just described for the neuromuscular junction are pertinent to all excitatory synapses. The principles of postsynaptic inhibition are much the same as for excitation, and are also quite general. In both cases, neurotransmitters binding to receptors open or close ion channels in the postsynaptic cell. Whether a postsynaptic response is an EPSP or an IPSP depends on the type of channel that is coupled to the receptor, and on the concentration of permeant ions inside and outside the cell. In fact, the only distinction between postsynaptic excitation and inhibition is the reversal potential of the PSP in relation to the threshold voltage for generating action potentials in the postsynaptic cell.

Consider, for example, a neuronal synapse that uses glutamate as the transmitter. Many such synapses have receptors that, like the ACh receptors at neuromuscular synapses, open ion channels that are nonselectively permeable to cations (see Chapter 6). When these glutamate receptors are activated, both Na^+ and K^+ flow across the postsynaptic membrane, yielding an E_{rev} of approximately 0 mV for the resulting postsynaptic current. If the rest-

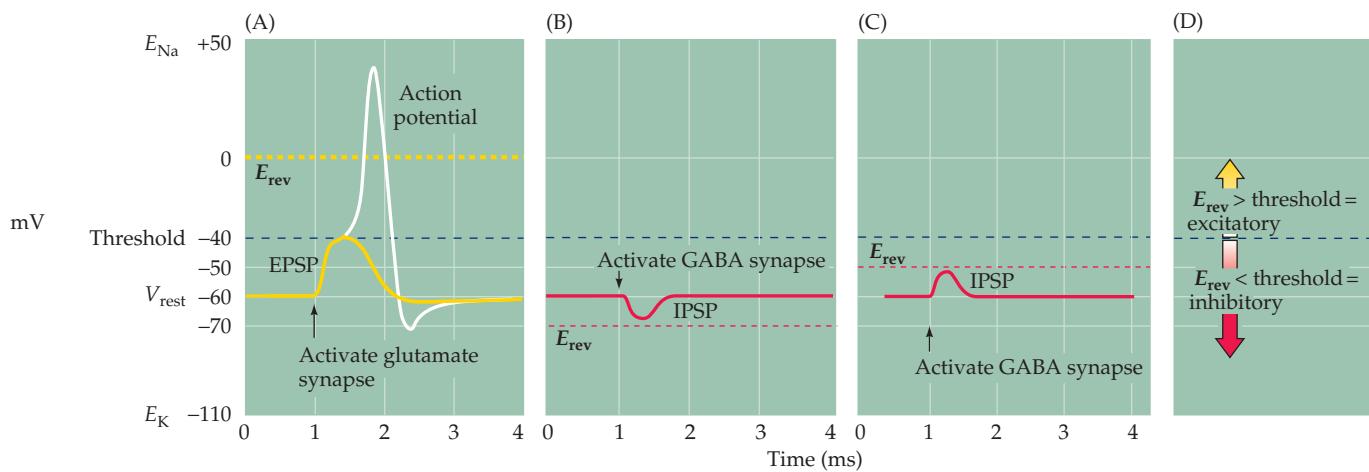


Figure 5.19 Reversal potentials and threshold potentials determine postsynaptic excitation and inhibition. (A) If the reversal potential for a PSP (0 mV) is more positive than the action potential threshold (-40 mV), the effect of a transmitter is excitatory, and it generates EPSPs. (B) If the reversal potential for a PSP is more negative than the action potential threshold, the transmitter is inhibitory and generate IPSPs. (C) IPSPs can nonetheless depolarize the postsynaptic cell if their reversal potential is between the resting potential and the action potential threshold. (D) The general rule of postsynaptic action is: If the reversal potential is more positive than threshold, excitation results; inhibition occurs if the reversal potential is more negative than threshold.

ing potential of the postsynaptic neuron is -60 mV, the resulting EPSP will depolarize by bringing the postsynaptic membrane potential toward 0 mV. For the hypothetical neuron shown in Figure 5.19A, the action potential threshold voltage is -40 mV. Thus, a glutamate-induced EPSP will increase the probability that this neuron produces an action potential, defining the synapse as excitatory.

As an example of inhibitory postsynaptic action, consider a neuronal synapse that uses GABA as its transmitter. At such synapses, the GABA receptors typically open channels that are selectively permeable to Cl^- and the action of GABA causes Cl^- to flow across the postsynaptic membrane. Consider a case where E_{Cl}^- is -70 mV, as is typical for many neurons, so that the postsynaptic resting potential of -60 mV is less negative than E_{Cl}^- . The resulting positive electrochemical driving force ($V_m - E_{\text{rev}}$) will cause negatively charged Cl^- to flow into the cell and produce a hyperpolarizing IPSP (Figure 5.19B). This hyperpolarizing IPSP will take the postsynaptic membrane away from the action potential threshold of -40 mV, clearly inhibiting the postsynaptic cell.

Surprisingly, inhibitory synapses need not produce hyperpolarizing IPSPs. For instance, if E_{Cl}^- were -50 mV instead of -70 mV, then the negative electrochemical driving force would cause Cl^- to flow out of the cell and produce a depolarizing IPSP (Figure 5.19C). However, the synapse would still be inhibitory: Given that the reversal potential of the IPSP still is more negative than the action potential threshold (-40 mV), the depolarizing IPSP would inhibit because the postsynaptic membrane potential would be kept more negative than the threshold for action potential initiation. Another way to think about this peculiar situation is that if another excitatory input onto this neuron brought the cell's membrane potential to -41 mV, just below threshold for firing an action potential, the IPSP would then hyperpolarize the membrane potential toward -50 mV, bringing the potential away from the action potential threshold. Thus, while EPSPs depolarize the postsynaptic cell, IPSPs can hyperpolarize or depolarize; indeed, an inhibitory conductance change may produce no potential change at all and still exert an inhibitory effect by making it more difficult for an EPSP to evoke an action potential in the postsynaptic cell.

Although the particulars of postsynaptic action can be complex, a simple rule distinguishes postsynaptic excitation from inhibition: An EPSP has a reversal potential more positive than the action potential threshold, whereas

an IPSP has a reversal potential more negative than threshold (Figure 5.19D). Intuitively, this rule can be understood by realizing that an EPSP will tend to depolarize the membrane potential so that it exceeds threshold, whereas an IPSP will always act to keep the membrane potential more negative than the threshold potential.

Summation of Synaptic Potentials

The PSPs produced at most synapses in the brain are much smaller than those at the neuromuscular junction; indeed, EPSPs produced by individual excitatory synapses may be only a fraction of a millivolt and are usually well below the threshold for generating postsynaptic action potentials. How, then, can such synapses transmit information if their PSPs are subthreshold? The answer is that neurons in the central nervous system are typically innervated by thousands of synapses, and the PSPs produced by each active synapse can *sum together*—in space and in time—to determine the behavior of the postsynaptic neuron.

Consider the highly simplified case of a neuron that is innervated by two excitatory synapses, each generating a subthreshold EPSP, and an inhibitory synapse that produces an IPSP (Figure 5.20A). While activation of either one of the excitatory synapses alone (E1 or E2 in Figure 5.20B) produces a sub-

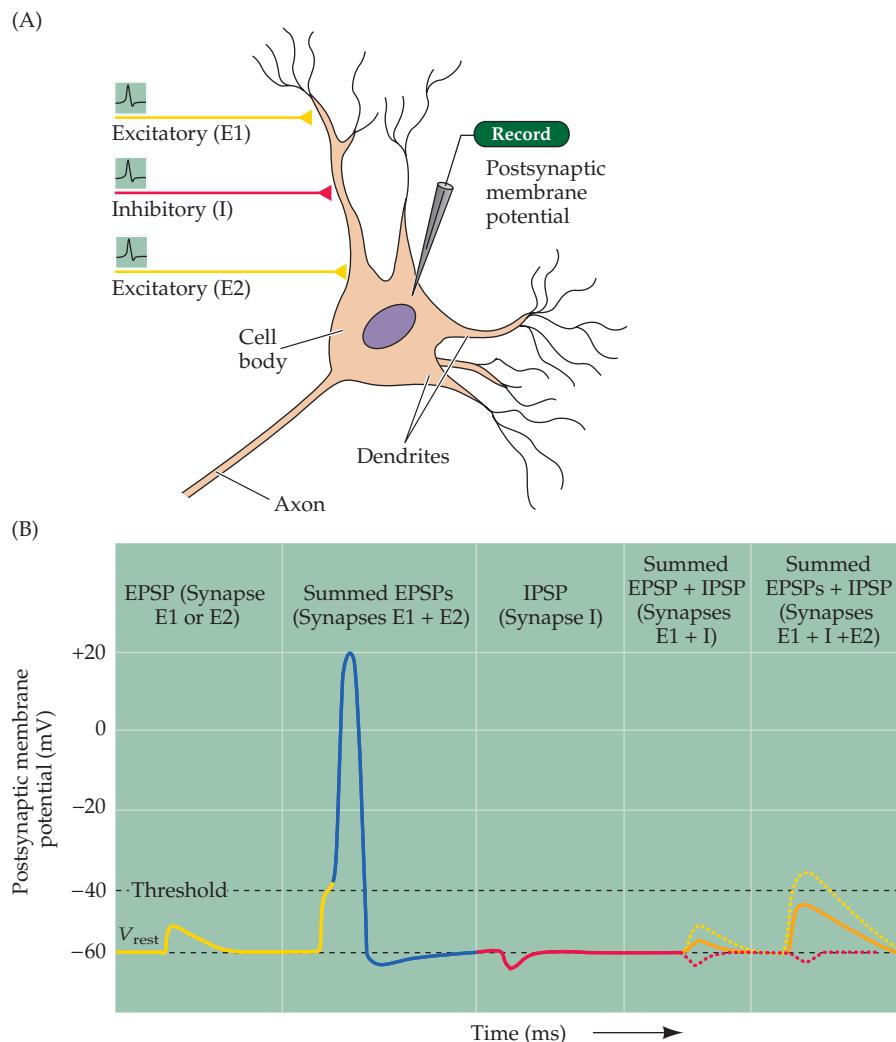


Figure 5.20 Summation of postsynaptic potentials. (A) A microelectrode records the postsynaptic potentials produced by the activity of two excitatory synapses (E1 and E2) and an inhibitory synapse (I). (B) Electrical responses to synaptic activation. Stimulating either excitatory synapse (E1 or E2) produces a subthreshold EPSP, whereas stimulating both synapses at the same time (E1 + E2) produces a suprathreshold EPSP that evokes a postsynaptic action potential (shown in blue). Activation of the inhibitory synapse alone (I) results in a hyperpolarizing IPSP. Summing this IPSP (dashed red line) with the EPSP (dashed yellow line) produced by one excitatory synapse (E1 + I) reduces the amplitude of the EPSP (orange line), while summing it with the suprathreshold EPSP produced by activating synapses E1 and E2 keeps the postsynaptic neuron below threshold, so that no action potential is evoked.

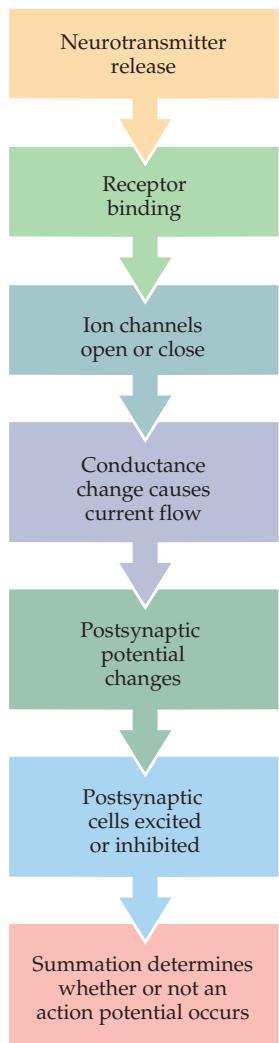


Figure 5.21 Events from neurotransmitter release to postsynaptic excitation or inhibition. Neurotransmitter release at all presynaptic terminals on a cell results in receptor binding, which causes the opening or closing of specific ion channels. The resulting conductance change causes current to flow, which may change the membrane potential. The postsynaptic cell sums (or integrates) all of the EPSPs and IPSPs, resulting in moment-to-moment control of action potential generation.

threshold EPSP, activation of both excitatory synapses at about the same time causes the two EPSPs to sum together. If the sum of the two EPSPs ($E_1 + E_2$) depolarizes the postsynaptic neuron sufficiently to reach the threshold potential, a postsynaptic action potential results. **Summation** thus allows subthreshold EPSPs to influence action potential production. Likewise, an IPSP generated by an inhibitory synapse (I) can sum (algebraically speaking) with a subthreshold EPSP to reduce its amplitude ($E_1 + I$) or can sum with suprathreshold EPSPs to prevent the postsynaptic neuron from reaching threshold ($E_1 + I + E_2$).

In short, the summation of EPSPs and IPSPs by a postsynaptic neuron permits a neuron to integrate the electrical information provided by all the inhibitory and excitatory synapses acting on it at any moment. Whether the sum of active synaptic inputs results in the production of an action potential depends on the balance between excitation and inhibition. If the sum of all EPSPs and IPSPs results in a depolarization of sufficient amplitude to raise the membrane potential above threshold, then the postsynaptic cell will produce an action potential. Conversely, if inhibition prevails, then the postsynaptic cell will remain silent. Normally, the balance between EPSPs and IPSPs changes continually over time, depending on the number of excitatory and inhibitory synapses active at a given moment and the magnitude of the current at each active synapse. Summation is therefore a neurotransmitter-induced tug-of-war between all excitatory and inhibitory postsynaptic currents; the outcome of the contest determines whether or not a postsynaptic neuron fires an action potential and, thereby, becomes an active element in the neural circuits to which it belongs (Figure 5.21).

Two Families of Postsynaptic Receptors

The opening or closing of postsynaptic ion channels is accomplished in different ways by two broad families of receptor proteins. The receptors in one family—called **ionotropic receptors**—are linked directly to ion channels (the Greek *tropos* means to move in response to a stimulus). These receptors contain two functional domains: an extracellular site that binds neurotransmitters, and a membrane-spanning domain that forms an ion channel (Figure 5.22A). Thus ionotropic receptors combine transmitter-binding and channel functions into a single molecular entity (they are also called **ligand-gated ion channels** to reflect this concatenation). Such receptors are multimers made up of at least four or five individual protein subunits, each of which contributes to the pore of the ion channel.

The second family of neurotransmitter receptors are the **metabotropic receptors**, so called because the eventual movement of ions through a channel depends on one or more metabolic steps. These receptors do not have ion channels as part of their structure; instead, they affect channels by the activation of intermediate molecules called **G-proteins** (Figure 5.22B). For this reason, metabotropic receptors are also called **G-protein-coupled receptors**. Metabotropic receptors are monomeric proteins with an extracellular domain that contains a neurotransmitter binding site and an intracellular domain that binds to G-proteins. Neurotransmitter binding to metabotropic receptors activates G-proteins, which then dissociate from the receptor and interact directly with ion channels or bind to other effector proteins, such as enzymes, that make intracellular messengers that open or close ion channels. Thus, G-proteins can be thought of as transducers that couple neurotransmitter binding to the regulation of postsynaptic ion channels. The postsynaptic signaling events initiated by metabotropic receptors are taken up in detail in Chapter 7.

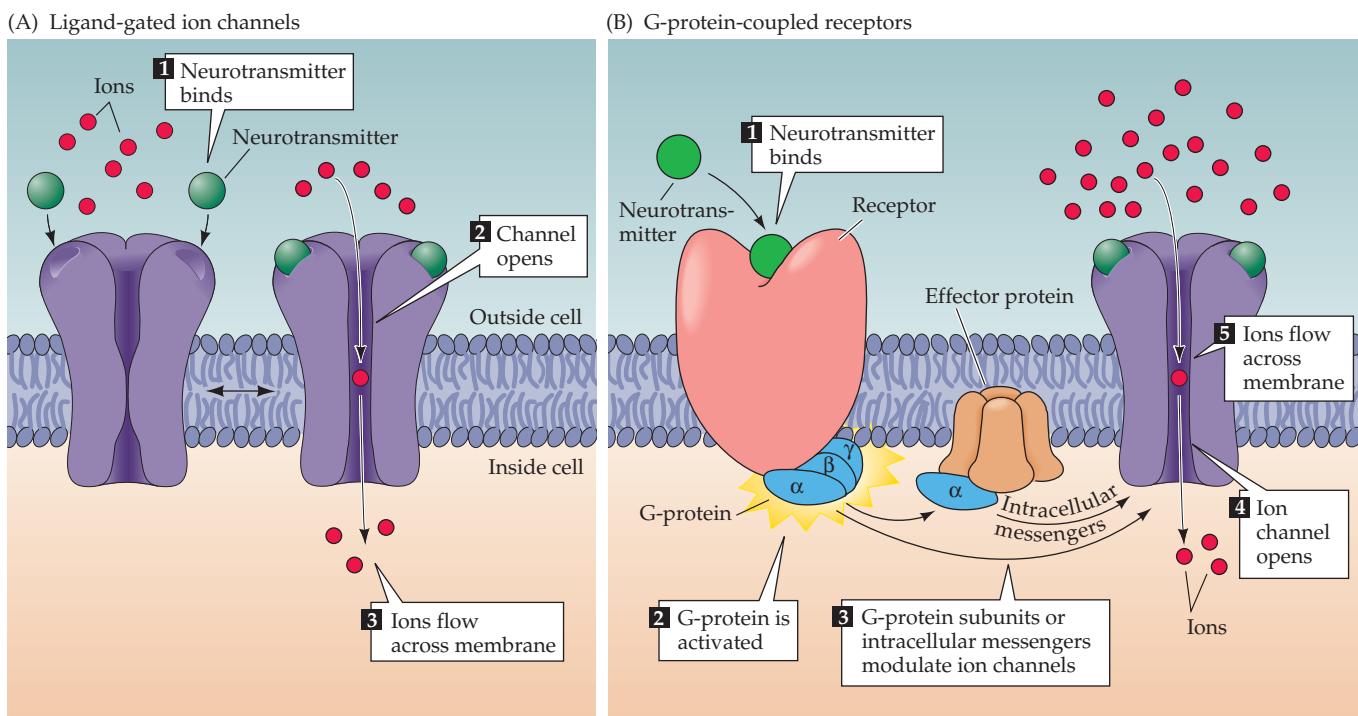


Figure 5.22 A neurotransmitter can affect the activity of a postsynaptic cell via two different types of receptor proteins: ionotropic or ligand-gated ion channels, and metabotropic or G-protein-coupled receptors. (A) Ligand-gated ion channels combine receptor and channel functions in a single protein complex. (B) Metabotropic receptors usually activate G-proteins, which modulate ion channels directly or indirectly through intracellular effector enzymes and second messengers.

These two families of postsynaptic receptors give rise to PSPs with very different time courses, producing postsynaptic actions that range from less than a millisecond to minutes, hours, or even days. Ionotropic receptors generally mediate rapid postsynaptic effects. Examples are the EPP produced at neuromuscular synapses by ACh (see Figure 5.15), EPSPs produced at certain glutamatergic synapses (Figure 5.19A), and IPSPs produced at certain GABAergic synapses (Figure 5.19B). In all three cases, the PSPs arise within a millisecond or two of an action potential invading the presynaptic terminal and last for only a few tens of milliseconds or less. In contrast, the activation of metabotropic receptors typically produces much slower responses, ranging from hundreds of milliseconds to minutes or even longer. The comparative slowness of metabotropic receptor actions reflects the fact that multiple proteins need to bind to each other sequentially in order to produce the final physiological response. Importantly, a given transmitter may activate both ionotropic and metabotropic receptors to produce both fast and slow PSPs at the same synapse.

Perhaps the most important principle to keep in mind is that the response elicited at a given synapse depends upon the neurotransmitter released and the postsynaptic complement of receptors and associated channels. The molecular mechanisms that allow neurotransmitters and their receptors to generate synaptic responses are considered in the next chapter.

Summary

Synapses communicate the information carried by action potentials from one neuron to the next in neural circuits. The cellular mechanisms that underlie synaptic transmission are closely related to the mechanisms that generate other types of neuronal electrical signals, namely ion flow through membrane channels. In the case of electrical synapses, these channels are gap junctions; direct but passive flow of current through the gap junctions is the basis for transmission. In the case of chemical synapses, channels with smaller and more selective pores are activated by the binding of neurotransmitters to postsynaptic receptors after release from the presynaptic terminal. The large number of neurotransmitters in the nervous system can be divided into two broad classes: small-molecule transmitters and neuropeptides. Neurotransmitters are synthesized from defined precursors by regulated enzymatic pathways, packaged into one of several types of synaptic vesicle, and released into the synaptic cleft in a Ca^{2+} -dependent manner. Many synapses release more than one type of neurotransmitter, and multiple transmitters can even be packaged within the same synaptic vesicle. Transmitter agents are released presynaptically in units or quanta, reflecting their storage within synaptic vesicles. Vesicles discharge their contents into the synaptic cleft when the presynaptic depolarization generated by the invasion of an action potential opens voltage-gated calcium channels, allowing Ca^{2+} to enter the presynaptic terminal. How calcium triggers neurotransmitter release is not yet established, but synaptotagmin, SNAREs, and a number of other proteins found within the presynaptic terminal are clearly involved. Postsynaptic receptors are a diverse group of proteins that transduce binding of neurotransmitters into electrical signals by opening or closing postsynaptic ion channels. The postsynaptic currents produced by the synchronous opening or closing of ion channels changes the conductance of the postsynaptic cell, thus increasing or decreasing its excitability. Conductance changes that increase the probability of firing an action potential are excitatory, whereas those that decrease the probability of generating an action potential are inhibitory. Because postsynaptic neurons are usually innervated by many different inputs, the integrated effect of the conductance changes underlying all EPSPs and IPSPs produced in a postsynaptic cell at any moment determines whether or not the cell fires an action potential. Two broadly different families of neurotransmitter receptors have evolved to carry out the postsynaptic signaling actions of neurotransmitters. The postsynaptic effects of neurotransmitters are terminated by the degradation of the transmitter in the synaptic cleft, by transport of the transmitter back into cells, or by diffusion out of the synaptic cleft.

Additional Reading

Reviews

AUGUSTINE, G. J. (2001) How does calcium trigger neurotransmitter release? *Curr. Opin. Neurobiol.* 11: 320–326.

BENNETT, M. V. L. (2000) Electrical synapses, a personal perspective (or history). *Brain Res. Rev.* 32: 16–28.

BRODSKY, F. M., C. Y. CHEN, C. KNUEHL, M. C. TOWLER AND D. E. WAKEHAM (2001) Biological basket weaving: Formation and function of clathrin-coated vesicles. *Annu. Rev. Cell. Dev. Biol.* 17: 517–568.

BRUNGER, A. T. (2001) Structure of proteins involved in synaptic vesicle fusion in neurons. *Annu. Rev. Biophys. Biomol. Struct.* 30: 157–171.

CARLSSON, A. (1987) Perspectives on the discovery of central monoaminergic neurotransmission. *Annu. Rev. Neurosci.* 10: 19–40.

CHANGEUX, J.-P. (1993) Chemical signaling in the brain. *Sci. Am.* 269 (May): 58–62.

EMSON, P. C. (1979) Peptides as neurotransmitter candidates in the CNS. *Prog. Neurobiol.* 13: 61–116.

GALARRETA, M. AND S. HESTRIN (2001) Electrical synapses between GABA-releasing interneurons. *Nature Rev. Neurosci.* 2: 425–433.

JAHN, R., T. LANG AND T. C. SÜDHOF (2003) Membrane fusion. *Cell* 112: 519–533.

KUPFERMANN, I. (1991) Functional studies of cotransmission. *Physiol. Rev.* 71: 683–732.

MARSH, M. AND H. T. McMAHON (1999) The structural era of endocytosis. *Science* 285: 215–220.

MURTHY, V. N. AND P. DE CAMILLI (2003) Cell biology of the presynaptic terminal. *Annu. Rev. Neurosci.* 26: 701–728.

ROTHMAN, J. E. (1994) Mechanisms of intracellular protein transport. *Nature* 372: 55–63.

SÜDHOF, T. C. (1995) The synaptic vesicle cycle: A cascade of protein-protein interactions. *Nature* 375: 645–653.

TUCKER, W. C. AND E. R. CHAPMAN (2002) Role of synaptotagmin in Ca^{2+} triggered exocytosis. *Biochem. J.* 366: 1–13.

Important Original Papers

ADLER, E., M. ADLER, G. J. AUGUSTINE, M. P. CHARLTON AND S. N. DUFFY (1991) Alien intracellular calcium chelators attenuate neurotransmitter release at the squid giant synapse. *J. Neurosci.* 11: 1496–1507.

AUGUSTINE, G. J. AND R. ECKERT (1984) Divalent cations differentially support transmitter release at the squid giant synapse. *J. Physiol. (Lond.)* 346: 257–271.

BOYD, I. A. AND A. R. MARTIN (1955) The end-plate potential in mammalian muscle. *J. Physiol. (Lond.)* 132: 74–91.

CURTIS, D. R., J. W. PHILLIS AND J. C. WATKINS (1959) Chemical excitation of spinal neurons. *Nature* 183: 611–612.

DALE, H. H., W. FELDBERG AND M. VOGT (1936) Release of acetylcholine at voluntary motor nerve endings. *J. Physiol.* 86: 353–380.

DEL CASTILLO, J. AND B. KATZ (1954) Quantal components of the end plate potential. *J. Physiol. (Lond.)* 124: 560–573.

FATT, P. AND B. KATZ (1951) An analysis of the end plate potential recorded with an intracellular electrode. *J. Physiol. (Lond.)* 115: 320–370.

FATT, P. AND B. KATZ (1952) Spontaneous sub-threshold activity at motor nerve endings. *J. Physiol. (Lond.)* 117: 109–128.

FURSHPAN, E. J. AND D. D. POTTER (1959) Transmission at the giant motor synapses of the crayfish. *J. Physiol. (Lond.)* 145: 289–325.

GEPPERT, M. AND 6 OTHERS (1994) Synaptotagmin I: A major Ca^{2+} sensor for transmitter release at a central synapse. *Cell* 79: 717–727.

GIBSON, J. R., M. BEIERLEIN AND B. W. CONNORS. (1999) Two networks of electrically coupled inhibitory neurons in neocortex. *Nature* 402: 75–79.

HARRIS, B. A., J. D. ROBISHAW, S. M. MUMBY AND A. G. GILMAN (1985) Molecular cloning of complementary DNA for the alpha subunit of the G protein that stimulates adenylyl cyclase. *Science* 229: 1274–1277.

HEUSER, J. E. AND 5 OTHERS (1979) Synaptic vesicle exocytosis captured by quick freezing and correlated with quantal transmitter release. *J. Cell Biol.* 81: 275–300.

HEUSER, J. E. AND T. S. REESE (1973) Evidence for recycling of synaptic vesicle membrane during transmitter release at the frog neuromuscular junction. *J. Cell Biol.* 57: 315–344.

HÖKFELT, T., O. JOHANSSON, A. LJUNGDAHL, J. M. LUNDBERG AND M. SCHULTZBERG (1980) Peptidergic neurons. *Nature* 284: 515–521.

JONAS, P., J. BISCHOFBERGER AND J. SANDKUHLER (1998) Corelease of two fast neurotransmitters at a central synapse. *Science* 281: 419–424.

LOEWI, O. (1921) Über humorale übertragbarkeit der herznervenwirkung. *Pflügers Arch.* 189: 239–242.

MILEDI, R. (1973) Transmitter release induced by injection of calcium ions into nerve terminals. *Proc. R. Soc. Lond. B* 183: 421–425.

NEHER, E. AND B. SAKMANN (1976) Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* 260:799–802.

REKLING, J. C., X. M. SHAO AND J. L. FELDMAN (2000) Electrical coupling and excitatory synaptic transmission between rhythmogenic respiratory neurons in the preBotzinger complex. *J. Neurosci.* 20: RC113: 1–5.

SMITH, S. J., J. BUCHANAN, L. R. OSSES, M. P. CHARLTON AND G. J. AUGUSTINE (1993) The spatial distribution of calcium signals in squid presynaptic terminals. *J. Physiol. (Lond.)* 472: 573–593.

SOSSIN, W. S., A. SWEET-CORDERO AND R. H. SCHELLER (1990) Dale's hypothesis revisited: Different neuropeptides derived from a common prohormone are targeted to different processes. *Proc. Natl. Acad. Sci. U.S.A.* 87: 4845–4548.

SUTTON, R. B., D. FASSHAUER, R. JAHN AND A. T. BRÜNGER (1998) Crystal structure of a SNARE complex involved in synaptic exocytosis at 2.4 Å resolution. *Nature* 395: 347–353.

TAKEUCHI, A. AND N. TAKEUCHI (1960) One the permeability of end-plate membrane during the action of transmitter. *J. Physiol. (Lond.)* 154: 52–67.

WICKMAN, K. AND 7 OTHERS (1994) Recombinant $\text{G}_{\beta\gamma}$ activates the muscarinic-gated atrial potassium channel I_{KACH} . *Nature* 368: 255–257.

Books

BRADFORD, H. F. (1986) *Chemical Neurobiology*. New York: W. H. Freeman.

COOPER, J. R., F. E. BLOOM AND R. H. ROTH (1991) *The Biochemical Basis of Neuropharmacology*. New York: Oxford University Press.

HALL, Z. (1992) *An Introduction to Molecular Neurobiology*. Sunderland, MA: Sinauer Associates.

KATZ, B. (1966) *Nerve, Muscle, and Synapse*. New York: McGraw-Hill.

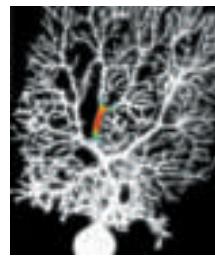
KATZ, B. (1969) *The Release of Neural Transmitter Substances*. Liverpool: Liverpool University Press.

LLINÁS, R. R. (1999) *The Squid Giant Synapse: A Model for Chemical Synaptic Transmission*. Oxford: Oxford University Press.

NICHOLLS, D. G. (1994) *Proteins, Transmitters, and Synapses*. Oxford: Blackwell.

PETERS, A., S. L. PALAY AND H. DEF. WEBSTER (1991) *The Fine Structure of the Nervous System: Neurons and their Supporting Cells*. 3rd edition. Oxford: Oxford University Press.

Chapter 6



Neurotransmitters and Their Receptors

Overview

For the most part, neurons in the human brain communicate with one another by releasing chemical messengers called neurotransmitters. A large number of neurotransmitters are now known and more remain to be discovered. Neurotransmitters evoke postsynaptic electrical responses by binding to members of a diverse group of proteins called neurotransmitter receptors. There are two major classes of receptors: those in which the receptor molecule is also an ion channel, and those in which the receptor and ion channel are separate molecules. The former are called ionotropic receptors or ligand-gated ion channels, and give rise to fast postsynaptic responses that typically last only a few milliseconds. The latter are called metabotropic receptors, and they produce slower postsynaptic effects that may endure much longer. Abnormalities in the function of neurotransmitter systems contribute to a wide range of neurological and psychiatric disorders. As a result, many neuropsychological therapies are based on drugs that affect neurotransmitter release, binding, and/or removal.

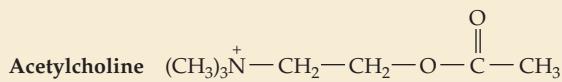
Categories of Neurotransmitters

More than 100 different agents are known to serve as neurotransmitters. This large number of transmitters allows for tremendous diversity in chemical signaling between neurons. It is useful to separate this panoply of transmitters into two broad categories based simply on size (Figure 6.1). **Neuropeptides** are relatively large transmitter molecules composed of 3 to 36 amino acids. Individual amino acids, such as glutamate and GABA, as well as the transmitters acetylcholine, serotonin, and histamine, are much smaller than neuropeptides and have therefore come to be called **small-molecule neurotransmitters**. Within the category of small-molecule neurotransmitters, the **biogenic amines** (dopamine, norepinephrine, epinephrine, serotonin, and histamine) are often discussed separately because of their similar chemical properties and postsynaptic actions. The particulars of synthesis, packaging, release, and removal differ for each neurotransmitter (Table 6.1). This chapter will describe some of the main features of these transmitters and their postsynaptic receptors.

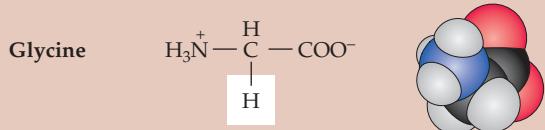
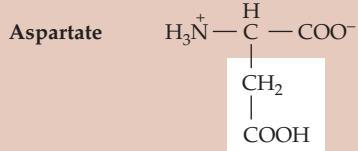
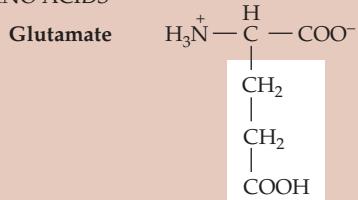
Acetylcholine

As mentioned in the previous chapter, acetylcholine (ACh) was the first substance identified as a neurotransmitter. In addition to the action of ACh as the neurotransmitter at skeletal neuromuscular junctions (see Chapter 5), as well as the neuromuscular synapse between the vagus nerve and cardiac

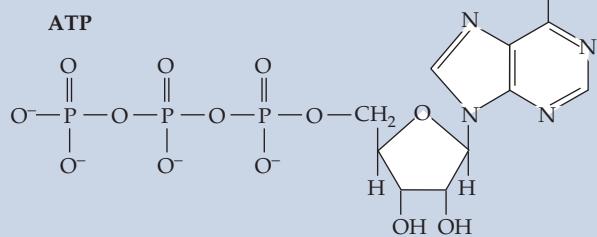
SMALL-MOLECULE NEUROTRANSMITTERS



AMINO ACIDS



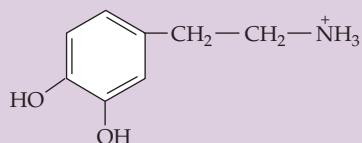
PURINES



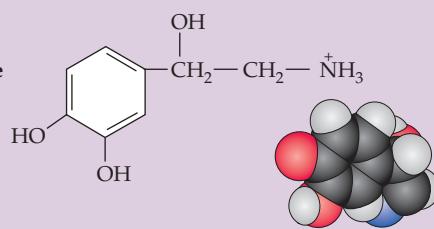
BIOGENIC AMINES

CATECHOLAMINES

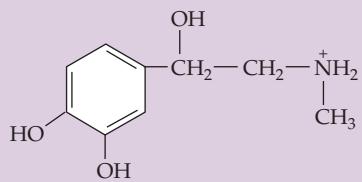
Dopamine



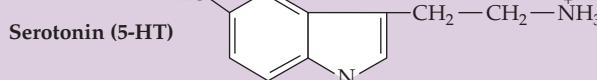
Norepinephrine



Epinephrine

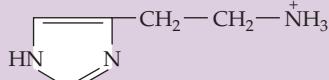


INDOLEAMINE



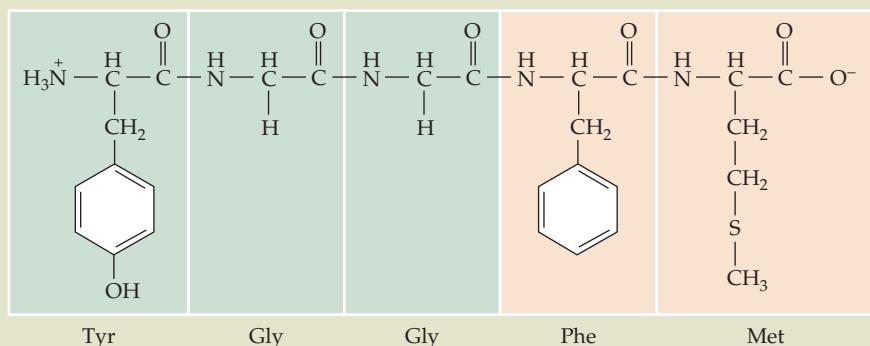
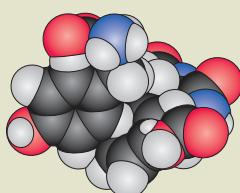
IMIDAZOLEAMINE

Histamine



PEPTIDE NEUROTRANSMITTERS (more than 100 peptides, usually 3–30 amino acids long)

Example: **Methionine enkephalin** (Tyr–Gly–Gly–Phe–Met)



◀ **Figure 6.1** Examples of small-molecule and peptide neurotransmitters. Small-molecule transmitters can be subdivided into acetylcholine, the amino acids, purines, and biogenic amines. The catecholamines, so named because they all share the catechol moiety (i.e., a hydroxylated benzene ring), make up a distinctive subgroup within the biogenic amines. Serotonin and histamine contain an indole ring and an imidazole ring, respectively. Size differences between the small-molecule neurotransmitters and the peptide neurotransmitters are indicated by the space-filling models for glycine, norepinephrine, and methionine enkephalin. (Carbon atoms are black, nitrogen atoms blue, and oxygen atoms red.)

muscle fibers, ACh serves as a transmitter at synapses in the ganglia of the visceral motor system, and at a variety of sites within the central nervous system. Whereas a great deal is known about the function of cholinergic transmission at neuromuscular junctions and ganglionic synapses, the actions of ACh in the central nervous system are not as well understood.

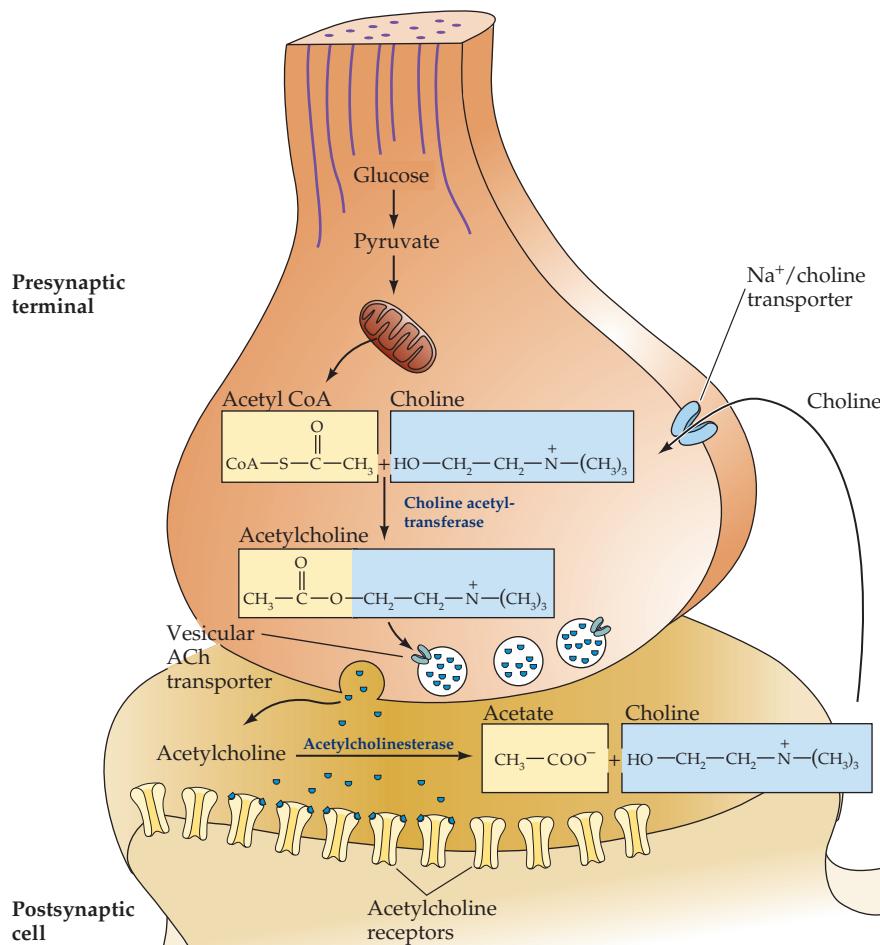
Acetylcholine is synthesized in nerve terminals from the precursors acetyl coenzyme A (acetyl CoA, which is synthesized from glucose) and choline, in a reaction catalyzed by choline acetyltransferase (CAT; Figure 6.2). Choline is present in plasma at a high concentration (about 10 mM) and is taken up into cholinergic neurons by a high-affinity $\text{Na}^+/\text{choline}$ transporter. After synthesis in the cytoplasm of the neuron, a vesicular ACh

TABLE 6.1
Functional Features of the Major Neurotransmitters

Neurotransmitter	Postsynaptic effect ^a	Precursor(s)	Rate-limiting step in synthesis	Removal mechanism	Type of vesicle
ACh	Excitatory	Choline + acetyl CoA	CAT	AChEase	Small, clear
Glutamate	Excitatory	Glutamine	Glutaminase	Transporters	Small, clear
GABA	Inhibitory	Glutamate	GAD	Transporters	Small, clear
Glycine	Inhibitory	Serine	Phosphoserine	Transporters	Small, clear
Catecholamines (epinephrine, norepinephrine, dopamine)	Excitatory	Tyrosine	Tyrosine hydroxylase	Transporters, MAO, COMT	Small dense-core, or large irregular dense-core
Serotonin (5-HT)	Excitatory	Tryptophan	Tryptophan hydroxylase	Transporters, MAO	Large, dense-core
Histamine	Excitatory	Histidine	Histidine decarboxylase	Transporters	Large, dense-core
ATP	Excitatory	ADP	Mitochondrial oxidative phosphorylation; glycolysis	Hydrolysis to AMP and adenosine	Small, clear
Neuropeptides	Excitatory and inhibitory	Amino acids (protein synthesis)	Synthesis and transport	Proteases	Large, dense-core
Endocannabinoids	Inhibits inhibition	Membrane lipids	Enzymatic modification of lipids	Hydrolysis by FAAH	None
Nitric oxide	Excitatory and inhibitory	Arginine	Nitric oxide synthase	Spontaneous oxidation	None

^aThe most common postsynaptic effect is indicated; the same transmitter can elicit postsynaptic excitation or inhibition depending on the nature of the ion channels affected by transmitter binding (see Chapter 7).

Figure 6.2 Acetylcholine metabolism in cholinergic nerve terminals. The synthesis of acetylcholine from choline and acetyl CoA requires choline acetyltransferase. Acetyl CoA is derived from pyruvate generated by glycolysis, while choline is transported into the terminals via a Na^+ -dependent transporter. Acetylcholine is loaded into synaptic vesicles via a vesicular transporter. After release, acetylcholine is rapidly metabolized by acetylcholinesterase, and choline is transported back into the terminal.



transporter loads approximately 10,000 molecules of ACh into each cholinergic vesicle.

In contrast to most other small-molecule neurotransmitters, the postsynaptic actions of ACh at many cholinergic synapses (the neuromuscular junction in particular) is not terminated by reuptake but by a powerful hydrolytic enzyme, acetylcholinesterase (AChE). This enzyme is concentrated in the synaptic cleft, ensuring a rapid decrease in ACh concentration after its release from the presynaptic terminal. AChE has a very high catalytic activity (about 5000 molecules of ACh per AChE molecule per second) and hydrolyzes ACh into acetate and choline. The choline produced by ACh hydrolysis is transported back into nerve terminals and used to resynthesize ACh.

Among the many interesting drugs that interact with cholinergic enzymes are the organophosphates. This group includes some potent chemical warfare agents. One such compound is the nerve gas "Sarin," which was made notorious after a group of terrorists released this gas in Tokyo's underground rail system. Organophosphates can be lethal because they inhibit AChE, causing ACh to accumulate at cholinergic synapses. This build-up of ACh depolarizes the postsynaptic cell and renders it refractory to subsequent ACh release, causing neuromuscular paralysis and other effects. The high sensitivity of insects to these AChE inhibitors has made organophosphates popular insecticides.

Many of the postsynaptic actions of ACh are mediated by the nicotinic ACh receptor (nAChR), so named because the CNS stimulant, nicotine, also

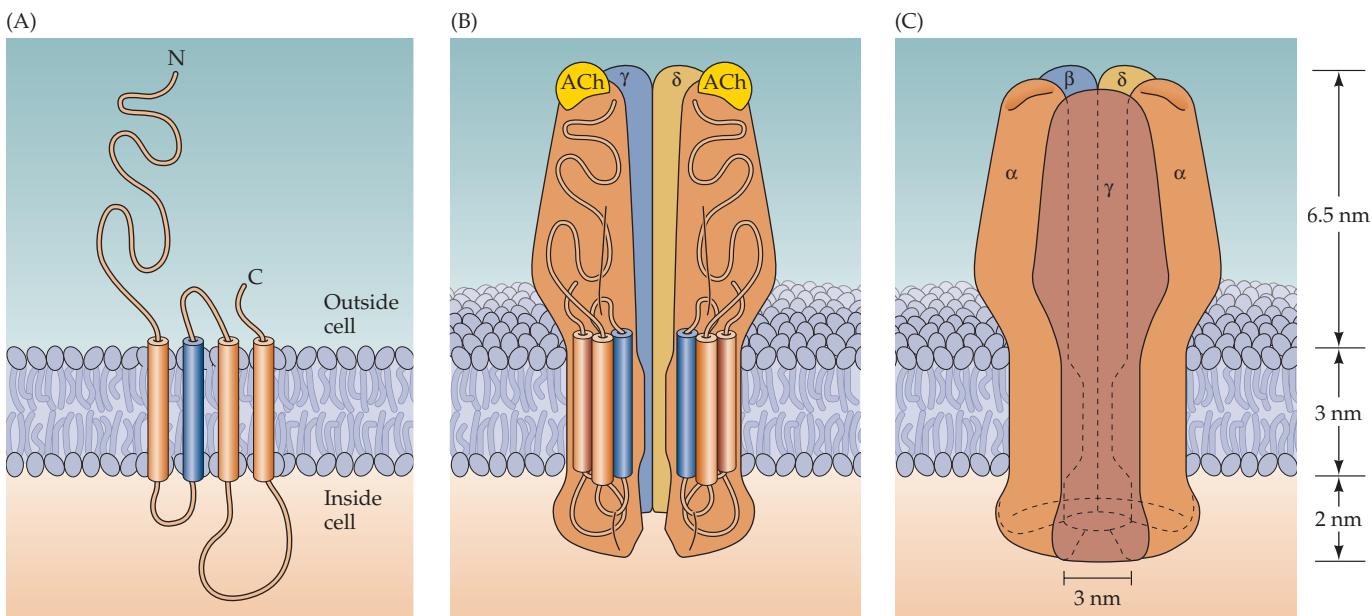
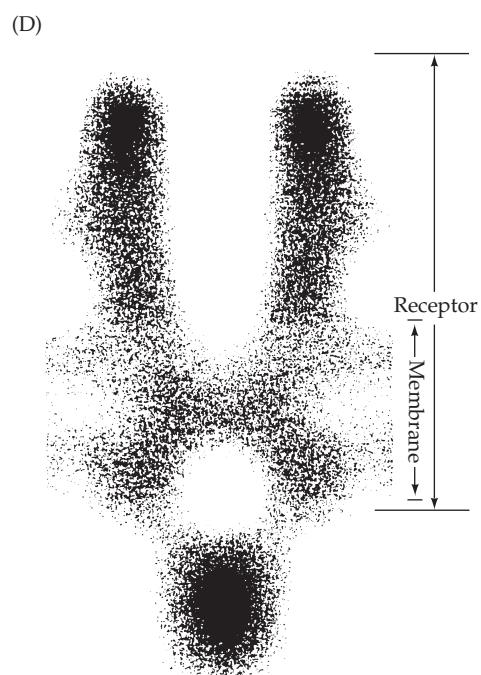


Figure 6.3 The structure of the nACh receptor/channel. (A) Each receptor sub-unit crosses the membrane four times. The membrane-spanning domain that lines the pore is shown in blue. (B) Five such subunits come together to form a complex structure containing 20 transmembrane domains that surround a central pore. (C) The openings at either end of the channel are very large—approximately 3 nm in diameter; even the narrowest region of the pore is approximately 0.6 nm in diameter. By comparison, the diameter of Na^+ or K^+ is less than 0.3 nm. (D) An electron micrograph of the nACh receptor, showing the actual position and size of the protein with respect to the membrane. (D from Toyoshima and Unwin, 1990.)

binds to these receptors. Nicotine consumption produces some degree of euphoria, relaxation, and eventually addiction (Box A), effects believed to be mediated in this case by nAChRs. Nicotinic receptors are the best-studied type of ionotropic neurotransmitter receptor. As described in Chapter 5, nAChRs are nonselective cation channels that generate excitatory postsynaptic responses. A number of biological toxins specifically bind to and block nicotinic receptors (Box B). The availability of these highly specific ligands—particularly a component of snake venom called α -bungarotoxin—has provided a valuable way to isolate and purify nAChRs. This pioneering work paved the way to cloning and sequencing the genes encoding the various subunits of the nAChR.

Based on these molecular studies, the nAChR is now known to be a large protein complex consisting of five subunits arranged around a central membrane-spanning pore (Figure 6.3). In the case of skeletal muscle AChRs, the receptor pentamer contains two α subunits, each of which binds one molecule of ACh. Because both ACh binding sites must be occupied for the channel to open, only relatively high concentrations of this neurotransmitter lead to channel activation. These subunits also bind other ligands, such as nicotine and α -bungarotoxin. At the neuromuscular junction, the two α subunits are combined with up to four other types of subunit— β , γ , δ , ϵ —in the ratio $2\alpha:\beta:\gamma:\delta$. Neuronal nAChRs typically differ from those of muscle in that they lack sensitivity to α -bungaro-



Box A

Addiction

Drug addiction is a chronic, relapsing disease with obvious medical, social, and political consequences. Addiction (also called substance dependence) is a persistent disorder of brain function in which compulsive drug use occurs despite serious negative consequences for the afflicted individual. The diagnostic manual of the American Psychiatric Association defines addiction in terms of both *physical* dependence and *psychological* dependence (in which an individual continues the drug-taking behavior despite obviously maladaptive consequences).

The range of substances that can generate this sort of dependence is wide; the primary agents of abuse at present are opioids, cocaine, amphetamines, marijuana, alcohol, and nicotine. Addiction to more “socially acceptable” agents such as alcohol and nicotine are sometimes regarded as less problematic, but in fact involve medical and behavioral consequences that are at least as great as for drugs of abuse that are considered more dangerous. Importantly, the phenomenon of addiction is

not limited to human behavior, but is demonstrable in laboratory animals. Most of these same agents are self-administered if primates, rodents, or other species are provided with the opportunity to do so.

In addition to a compulsion to take the agent of abuse, a major feature of addiction for many drugs is a constellation of negative physiological and emotional features, loosely referred to as “withdrawal syndrome,” that occur when the drug is not taken. The signs and symptoms of withdrawal are different for each agent of abuse, but in general are characterized by effects opposite those of the positive experience induced by the drug itself. Consider, as an example, cocaine, a drug that was estimated to be in regular use by 5 to 6 million Americans during the decade of the 1990s, with about 600,000 regular users either addicted or at high risk for addiction. The positive effects of the drug smoked or inhaled as a powder in the form of the alkaloidal free base is a “high” that is nearly immediate but generally lasts only a few minutes, typi-

cally leading to a desire for additional drug in as little as 10 minutes to half an hour. The “high” is described as a feeling of well-being, self-confidence, and satisfaction. Conversely, when the drug is not available, frequent users experience depression, sleepiness, fatigue, drug-craving, and a general sense of malaise.

Another aspect of addiction to cocaine or other agents is tolerance, defined as a reduction in the response to the drug upon repeated administration. Tolerance occurs as a consequence of persistent use of a number of drugs but is particularly significant in drug addiction, since it progressively increases the dose needed to experience the desired effects.

Although it is fair to say that the neurobiology of addiction is incompletely understood, for cocaine and many other agents of abuse the addictive effects involve activation of dopamine receptors in critical brain regions involved in motivation and emotional reinforcement (see Chapter 28). The most important of these areas is the midbrain dopamine system,

toxin, and comprise only two receptor subunit types (α and β), which are present in a ratio of $3\alpha:2\beta$. In all cases, however, five individual subunits assemble to form a functional, cation-selective nACh receptor.

Each subunit of the nAChR molecule contains four transmembrane domains that make up the ion channel portion of the receptor, and a long extracellular region that makes up the ACh-binding domain (Figure 6.3A). Unraveling the molecular structure of this region of the nACh receptor has provided insight into the mechanisms that allow ligand-gated ion channels to respond rapidly to neurotransmitters: The intimate association of the ACh binding sites with the pore of the channel presumably accounts for the rapid response to ACh (Figure 6.3B–D). Indeed, this general arrangement is characteristic of *all* of the ligand-gated ion channels at fast-acting synapses, as summarized in Figure 6.4. Thus, the nicotinic receptor has served as a paradigm for studies of other ligand-gated ion channels, at the same time leading to a much deeper appreciation of several neuromuscular diseases (Box C).

especially its projections from the ventral-tegmental area to the nucleus accumbens. Agents such as cocaine appear to act by raising dopamine levels in these areas, making this transmitter more available to receptors by interfering with re-uptake of synaptically released dopamine by the dopamine transporter. The reinforcement and motivation of drug-taking behaviors is thought to be related to the projections to the nucleus accumbens.

The most common opioid drug of abuse is heroin. Heroin is a derivative of the opium poppy and is not legally available for clinical purposes in the United States. The number of heroin addicts in the United States is estimated to be between 750,000 and a million individuals. The positive feelings produced by heroin, generally described as the "rush," are often compared to the feeling of sexual orgasm and begin in less than a minute after intravenous injection. There is then a feeling of general well-being (referred to as "on the nod") that lasts about an hour. The symptoms of withdrawal can be

intense; these are restlessness, irritability, nausea, muscle pain, depression, sleeplessness, and a sense of anxiety and malaise. The reinforcing aspects of the drug entail the same dopaminergic circuitry in the ventral tegmental area and nucleus accumbens as does cocaine, although additional areas are certainly involved, particularly the sites of opioid receptors described in Chapter 9.

Interestingly, addiction to heroin or any other agent is not an inevitable consequence of drug use, but depends critically on the environment. For instance, returning veterans who were heroin addicts in Vietnam typically lost their addiction upon returning to the United States. Likewise, patients given other opioids (e.g., morphine) for painful conditions rarely become addicts.

The treatment of any form of addiction is difficult and must be tailored to the circumstances of the individual. In addition to treating acute problems of withdrawal and "detoxification," patterns of behavior must be changed that may take months or years. Addiction is thus a chronic disease state that requires

continual monitoring during the lifetime of susceptible individuals.

References

- AMERICAN PSYCHIATRIC ASSOCIATION (1994) *Diagnostic and Statistical Manual of Mental Disorders*, 4th Edition (DSM IV). Washington, D.C.
- HYMAN, S. E. AND R. C. MALENKA (2001) Addiction and the brain: The neurobiology of compulsion and its persistence. *Nature Rev. Neurosci.* 2: 695–703.
- LAAKSO, A., A. R. MOHN, R. R. GAINETDINOV AND M. G. CARON (2002) Experimental genetic approaches to addiction. *Neuron* 36: 213–228.
- O'BRIEN, C. P. (2001) Goodman and Gilman's *The Pharmaceutical Basis of Therapeutics*, 10th Edition. New York: McGraw-Hill, Chapter 24, pp. 621–642..

A second class of ACh receptors is activated by muscarine, a poisonous alkaloid found in some mushrooms (see Box B), and thus they are referred to as muscarinic ACh receptors (mAChRs). mAChRs are metabotropic and mediate most of the effects of ACh in brain. Several subtypes of mAChR are known (Figure 6.5). Muscarinic ACh receptors are highly expressed in the striatum and various other forebrain regions, where they exert an inhibitory influence on dopamine-mediated motor effects. These receptors are also found in the ganglia of the peripheral nervous system. Finally, they mediate peripheral cholinergic responses of autonomic effector organs—such as heart, smooth muscle, and exocrine glands—and are responsible for the inhibition of heart rate by the vagus nerve. Numerous drugs act as mACh receptor agonists or antagonists, but most of these do not discriminate between different types of muscarinic receptors and often produce side effects. Nevertheless, mACh blockers that are therapeutically useful include atropine (used to dilate the pupil), scopolamine (effective in preventing motion sickness), and ipratropium (useful in the treatment of asthma).

Box B

Neurotoxins that Act on Postsynaptic Receptors

Poisonous plants and venomous animals are widespread in nature. The toxins they produce have been used for a variety of purposes, including hunting, healing, mind-altering, and, more recently, research. Many of these toxins have potent actions on the nervous system, often interfering with synaptic transmission by targeting neurotransmitter receptors. The poisons found in some organisms contain a single type of toxin, whereas others contain a mixture of tens or even hundreds of toxins.

Given the central role of ACh receptors in mediating muscle contraction at neuromuscular junctions in numerous species, it is not surprising that a large number of natural toxins interfere with transmission at this synapse. In fact, the classification of nicotinic and muscarinic ACh receptors is based on the sensitivity of these receptors to the toxic plant alkaloids nicotine and muscarine, which activate nicotinic and muscarinic ACh receptors, respectively. Nicotine is derived

from the dried leaves of the tobacco plant *Nicotinia tabacum*, and muscarine is from the poisonous red mushroom *Amanita muscaria*. Both toxins are stimulants that produce nausea, vomiting, mental confusion, and convulsions. Muscarine poisoning can also lead to circulatory collapse, coma, and death.

The poison α -bungarotoxin, one of many peptides that together make up the venom of the banded krait, *Bungarus multicinctus* (Figure A), blocks transmission at neuromuscular junctions and is used by the snake to paralyze its prey. This 74-amino-acid toxin blocks neuromuscular transmission by irreversibly binding to nicotinic ACh receptors, thus preventing ACh from opening postsynaptic ion channels. Paralysis ensues because skeletal muscles can no longer be activated by motor neurons. As a result of its specificity and its high affinity for nicotinic ACh receptors, α -bungarotoxin has contributed greatly to understanding the ACh receptor mole-

cule. Other snake toxins that block nicotinic ACh receptors are cobra α -neurotoxin and the sea snake peptide erabutoxin. The same strategy used by these snakes to paralyze prey was adopted by South American Indians who used curare, a mixture of plant toxins from *Chondodendron tomentosum*, as an arrowhead poison to immobilize their quarry. Curare also blocks nicotinic ACh receptors; the active agent is the alkaloid δ -tubocurarine.

Another interesting class of animal toxins that selectively block nicotinic ACh and other receptors includes the peptides produced by fish-hunting marine cone snails (Figure B). These colorful snails kill small fish by “shooting” venomous darts into them. The venom contains hundreds of peptides, known as the conotoxins, many of which target proteins that are important in synaptic transmission. There are conotoxin peptides that block Ca^{2+} channels, Na^+ channels, glutamate receptors, and ACh

(A)



(A) The banded krait *Bungarus multicinctus*.

(B)



(B) A marine cone snail (*Conus* sp.) uses venomous darts to kill a small fish.

(C)



(C) Betel nuts, *Areca catechu*, growing in Malaysia.

(A, Robert Zappalorti/Photo Researchers, Inc.; B, Zoya Maslak and Baldomera Olivera, University of Utah; C, Fletcher Baylis/Photo Researchers, Inc.)

receptors. The array of physiological responses produced by these peptides all serve to immobilize any prey unfortunate enough to encounter the cone snail. Many other organisms, including other mollusks, corals, worms and frogs, also utilize toxins containing specific blockers of ACh receptors.

Other natural toxins have mind- or behavior-altering effects and in some cases have been used for thousands of years by shamans and, more recently, physicians. Two examples are plant alkaloid toxins that block muscarinic ACh receptors: atropine from deadly nightshade (belladonna), and scopolamine from henbane. Because these plants grow wild in many parts of the world, exposure is not unusual, and poisoning by either toxin can also be fatal.

Another postsynaptic neurotoxin that, like nicotine, is used as a social drug is found in the seeds from the betel nut, *Areca catechu* (Figure C). Betel nut chewing, although unknown in the United States, is practiced by up to 25% of the population in India, Bangladesh, Ceylon, Malaysia, and the Philippines. Chewing these nuts produces a euphoria caused by arecoline, an alkaloid agonist of nicotinic ACh receptors. Like nicotine, arecoline is an addictive central nervous system stimulant.

Many other neurotoxins alter transmission at noncholinergic synapses. For example, amino acids found in certain

mushrooms, algae, and seeds are potent glutamate receptor agonists. The excitotoxic amino acids kainate, from the red alga *Digenea simplex*, and quisqualate, from the seed of *Quisqualis indica*, are used to distinguish two families of non-NMDA glutamate receptors (see text). Other neurotoxic amino acid activators of glutamate receptors include ibotenic acid and acromelic acid, both found in mushrooms, and domoate, which occurs in algae, seaweed, and mussels. Another large group of peptide neurotoxins blocks glutamate receptors. These include the α -agatoxins from the funnel web spider, NSTX-3 from the orb weaver spider, jorotoxin from the Joro spider, and β -philanthotoxin from wasp venom, as well as many cone snail toxins.

All the toxins discussed so far target excitatory synapses. The inhibitory GABA and glycine receptors, however, have not been overlooked by the exigencies of survival. Strychnine, an alkaloid extracted from the seeds of *Strychnos nux-vomica*, is the only drug known to have specific actions on transmission at glycinergic synapses. Because the toxin blocks glycine receptors, strychnine poisoning causes overactivity in the spinal cord and brainstem, leading to seizures. Strychnine has long been used commercially as a poison for rodents, although alternatives such as the anticoagulant coumadin are now more popular because they are safer for humans. Neu-

rotoxins that block GABA_A receptors include plant alkaloids such as bicuculline from Dutchman's breeches and picrotoxin from *Anamerta cocculus*. Dieldrin, a commercial insecticide, also blocks these receptors. These agents are, like strychnine, powerful central nervous system stimulants. Muscimol, a mushroom toxin that is a powerful depressant as well as a hallucinogen, activates GABA_A receptors. A synthetic analogue of GABA, baclofen, is a GABA_B agonist that reduces EPSPs in some brainstem neurons and is used clinically to reduce the frequency and severity of muscle spasms.

Chemical warfare between species has thus given rise to a staggering array of molecules that target synapses throughout the nervous system.

Although these toxins are designed to defeat normal synaptic transmission, they have also provided a set of powerful tools to understand postsynaptic mechanisms.

References

- ADAMS, M. E. AND B. M. OLIVERA (1994) Neurotoxins: Overview of an emerging research technology. *TINS* 17: 151–155.
- HUCHO, F. AND Y. OVCHINNIKOV (1990) *Toxins as Tools in Neurochemistry*. Berlin: Walter de Gruyter.
- MYERS, R. A., L. J. CRUZ, J. E. RIVIER AND B. M. OLIVERA (1993) *Conus* peptides as chemical probes for receptors and ion channels. *Chem. Rev.* 93: 1923–1926.

Glutamate

Glutamate is the most important transmitter in normal brain function. Nearly all excitatory neurons in the central nervous system are glutamatergic, and it is estimated that over half of all brain synapses release this agent. Glutamate plays an especially important role in clinical neurology because elevated concentrations of extracellular glutamate, released as a result of neural injury, are toxic to neurons (Box D).

Glutamate is a nonessential amino acid that does not cross the blood-brain barrier and therefore must be synthesized in neurons from local precursors. The most prevalent precursor for glutamate synthesis is glutamine, which is released by glial cells. Once released, glutamine is taken up into presynaptic

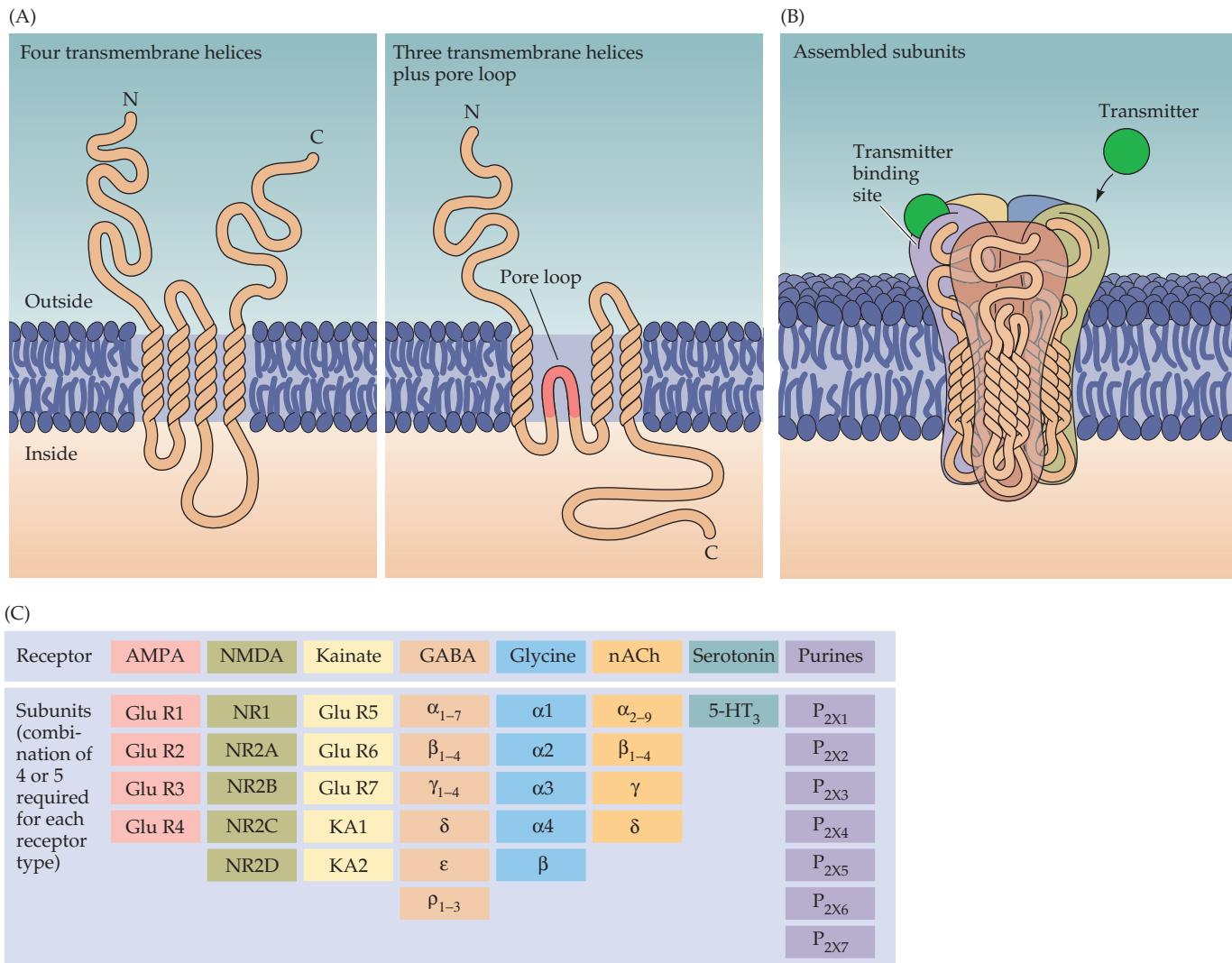


Figure 6.4 The general architecture of ligand-gated receptors. (A) One of the subunits of a complete receptor. The long N-terminal region forms the ligand-binding site, while the remainder of the protein spans the membrane either four times (left) or three times (right). (B) Assembly of either four or five subunits into a complete receptor. (C) A diversity of subunits come together to form functional ionotropic neurotransmitter receptors.

terminals and metabolized to glutamate by the mitochondrial enzyme glutaminase (Figure 6.6). Glutamate can also be synthesized by transamination of 2-oxoglutarate, an intermediate of the tricarboxylic acid cycle. Hence, some of the glucose metabolized by neurons can also be used for glutamate synthesis.

The glutamate synthesized in the presynaptic cytoplasm is packaged into synaptic vesicles by transporters, termed VGLUT. At least three different VGLUT genes have been identified. Once released, glutamate is removed from the synaptic cleft by the excitatory amino acid transporters (EAATs). There are five different types of high-affinity glutamate transporters exist, some of which are present in glial cells and others in presynaptic terminals. Glutamate taken up by glial cells is converted into glutamine by the enzyme

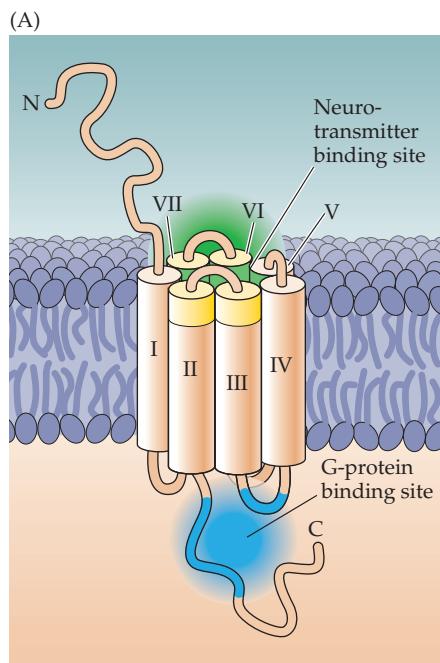


Figure 6.5 Structure and function of metabotropic receptors. (A) The transmembrane architecture of metabotropic receptors. These monomeric proteins contain seven transmembrane domains. Portions of domains II, III, VI, and VII make up the neurotransmitter-binding region. G-proteins bind to both the loop between domains V and VI and to portions of the C-terminal region. (B) Varieties of metabotropic neurotransmitter receptors.

Receptor class	Glutamate	GABA _B	Dopamine	NE, Epi	Histamine	Serotonin	Purines	Muscarinic
Receptor subtype	Class I mGlu R1 mGlu R5	GABA _B R1 GABA _B R2	D1 _A D1 _B D2	α1 α2 β1	H1 H2 H3	5-HT 1 5-HT 2 5-HT 3 5-HT 4 5-HT 5 5-HT 6 5-HT 7	A type A1 A2a A2b A3	M1 M2 M3 M4 M5
Class II	mGlu R2 mGlu R3		D3 D4	β2 β3			P type P2x P2y P2z P2t P2u	
Class III	mGlu R4 mGlu R6 mGlu R7 mGlu R8							

glutamine synthetase; glutamine is then transported out of the glial cells and into nerve terminals. In this way, synaptic terminals cooperate with glial cells to maintain an adequate supply of the neurotransmitter. This overall sequence of events is referred to as the **glutamate-glutamine cycle** (see Figure 6.6).

Several types of glutamate receptors have been identified. Three of these are ionotropic receptors called, respectively, **NMDA receptors**, **AMPA receptors**, and **kainate receptors** (Figure 6.4C). These glutamate receptors are named after the agonists that activate them: NMDA (*N*-methyl-D-aspartate), AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazole-propionate), and kainic acid. All of the ionotropic glutamate receptors are nonselective cation channels similar to the nAChR, allowing the passage of Na^+ and K^+ , and in some cases small amounts of Ca^{2+} . Hence AMPA, kainate, and NMDA receptor activation always produces excitatory postsynaptic responses. Like other ionotropic receptors, AMPA/kainate and NMDA receptors are also formed

Box C

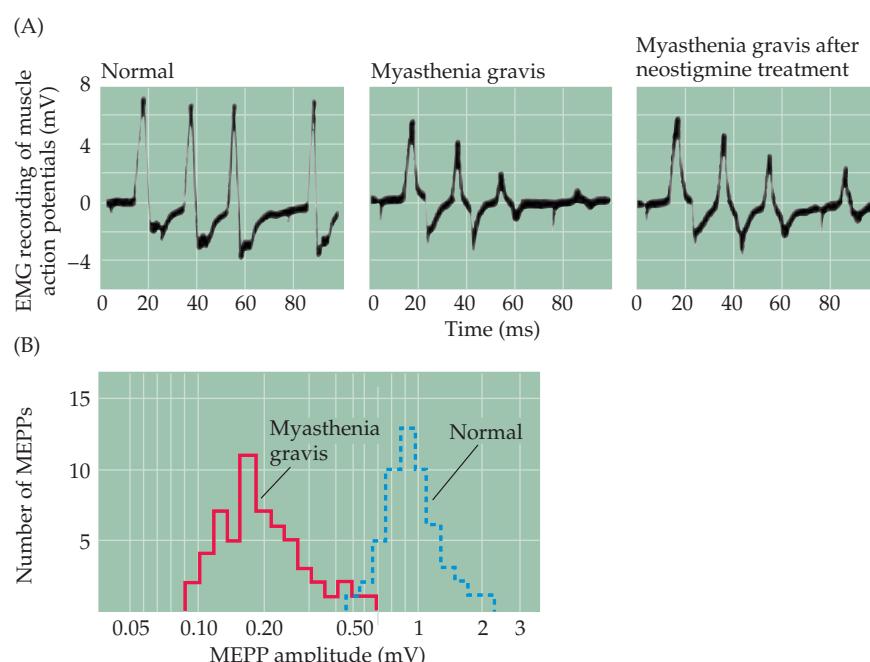
Myasthenia Gravis: An Autoimmune Disease of Neuromuscular Synapses

Myasthenia gravis is a disease that interferes with transmission between motor neurons and skeletal muscle fibers and afflicts approximately 1 of every 200,000 people. Originally described by the British physician Thomas Willis in 1685, the hallmark of the disorder is muscle weakness, particularly during sustained activity.

Although the course is variable, myasthenia commonly affects muscles controlling the eyelids (resulting in drooping of the eyelids, or ptosis) and eye movements (resulting in double vision, or diplopia). Muscles controlling facial expression, chewing, swallowing, and speaking are other common targets.

An important indication of the cause of myasthenia gravis came from the clinical observation that the muscle weakness improves following treatment with inhibitors of acetylcholinesterase, the enzyme that normally degrades acetylcholine at the neuromuscular junction. Studies of muscle obtained by biopsy from myasthenic patients showed that both end plate potentials (EPPs) and miniature end plate potentials (MEPPs) are much smaller than normal (see figure; also see Chapter 5). Because both the frequency of MEPPs and the quantal content of EPPs are normal, it seemed likely that myasthenia gravis entails a disorder of the postsynaptic muscle cells. Indeed, electron microscopy shows that the structure of neuromuscular junctions is altered, obvious changes being a widening of the synaptic cleft and an apparent reduction in the number of acetylcholine receptors in the postsynaptic membrane.

A chance observation in the early 1970s led to the discovery of the underlying cause of these changes. Jim Patrick and Jon Lindstrom, then working at the Salk Institute, were attempting to raise antibodies to nicotinic acetylcholine receptors by immunizing rabbits with



(A) Myasthenia gravis reduces the efficiency of neuromuscular transmission. Electromyographs show muscle responses elicited by stimulating motor nerves. In normal individuals, each stimulus in a train evokes the same contractile response. In contrast, transmission rapidly fatigues in myasthenic patients, although it can be partially restored by administration of acetylcholinesterase inhibitors such as neostigmine. (B) Distribution of MEPP amplitudes in muscle fibers from myasthenic patients (solid line) and controls (dashed line). The smaller size of MEPPs in myasthenics is due to a diminished number of postsynaptic receptors. (A after Harvey et al., 1941; B after Elmqvist et al., 1964.)

the receptors. Unexpectedly, the immunized rabbits developed muscle weakness that improved after treatment with acetylcholinesterase inhibitors. Subsequent work showed that the blood of myasthenic patients contains antibodies directed against the acetylcholine receptor, and that these antibodies are present at neuromuscular synapses. Removal of antibodies by plasma exchange improves the weakness. Finally, injecting the serum of myasthenic patients into mice produces myasthenic effects (because the serum carries circulating antibodies).

These findings indicate that myasthenia gravis is an autoimmune disease that targets nicotinic acetylcholine receptors. The immune response

reduces the number of functional receptors at the neuromuscular junction and can eventually destroy them altogether, diminishing the efficiency of synaptic transmission; muscle weakness thus occurs because motor neurons are less capable of exciting the postsynaptic muscle cells. This causal sequence also explains why cholinesterase inhibitors alleviate the signs and symptoms of myasthenia: The inhibitors increase the concentration of acetylcholine in the synaptic cleft, allowing more effective activation of those postsynaptic receptors not yet destroyed by the immune system.

Despite all these insights, it is still not clear what triggers the immune system to produce an autoimmune

response to acetylcholine receptors. Surgical removal of the thymus is beneficial in young patients with hyperplasia of the thymus, though precisely how the thymus contributes to myasthenia gravis is incompletely understood. Many patients are treated with combi-

nations of immunosuppression and cholinesterase inhibitors.

References

- ELMQVIST, D., W. W. HOFMANN, J. KUGELBERG AND D. M. J. QUASTEL (1964) An electrophysiological investigation of neuromuscular

transmission in myasthenia gravis. *J. Physiol. (Lond.)* 174: 417–434.

PATRICK, J. AND J. LINDSTROM (1973) Autoimmune response to acetylcholine receptor. *Science* 180: 871–872.

VINCENT, A. (2002) Unravelling the pathogenesis of myasthenia gravis. *Nature Rev. Immunol.* 2: 797–804.

from the association of several protein subunits that can combine in many ways to produce a large number of receptor isoforms (see Figure 6.4C).

NMDA receptors have especially interesting properties (Figure 6.7A). Perhaps most significant is the fact that NMDA receptor ion channels allow the entry of Ca^{2+} in addition to monovalent cations such as Na^+ and K^+ . As a result, EPSPs produced by NMDA receptors can increase the concentration of Ca^{2+} within the postsynaptic neuron; the Ca^{2+} concentration change can then act as a second messenger to activate intracellular signaling cascades (see Chapter 7). Another key property is that they bind extracellular Mg^{2+} . At hyperpolarized membrane potentials, this ion blocks the pore of the NMDA receptor channel. Depolarization, however, pushes Mg^{2+} out of the pore, allowing other cations to flow. This property provides the basis for a voltage-dependence to current flow through the receptor (dashed line in Figure 6.7B) and means that NMDA receptors pass cations (most notably Ca^{2+})

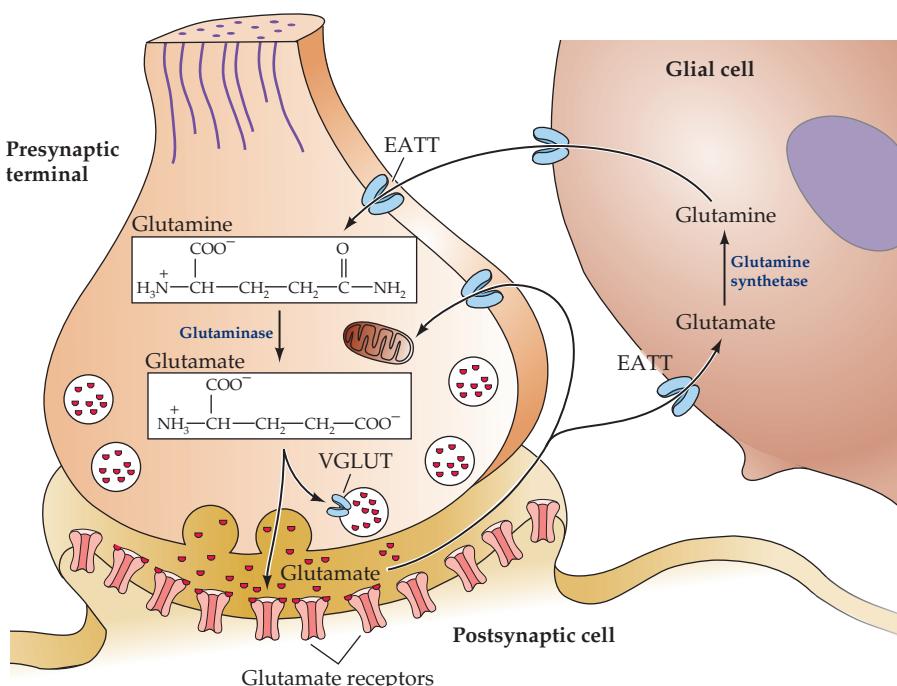


Figure 6.6 Glutamate synthesis and cycling between neurons and glia. The action of glutamate released into the synaptic cleft is terminated by uptake into neurons and surrounding glial cells via specific transporters. Within the nerve terminal, the glutamine released by glial cells and taken up by neurons is converted back to glutamate. Glutamate is transported into cells via excitatory amino acid transporters (EATTS) and loaded into synaptic vesicles via vesicular glutamate transporters (VGLUT).

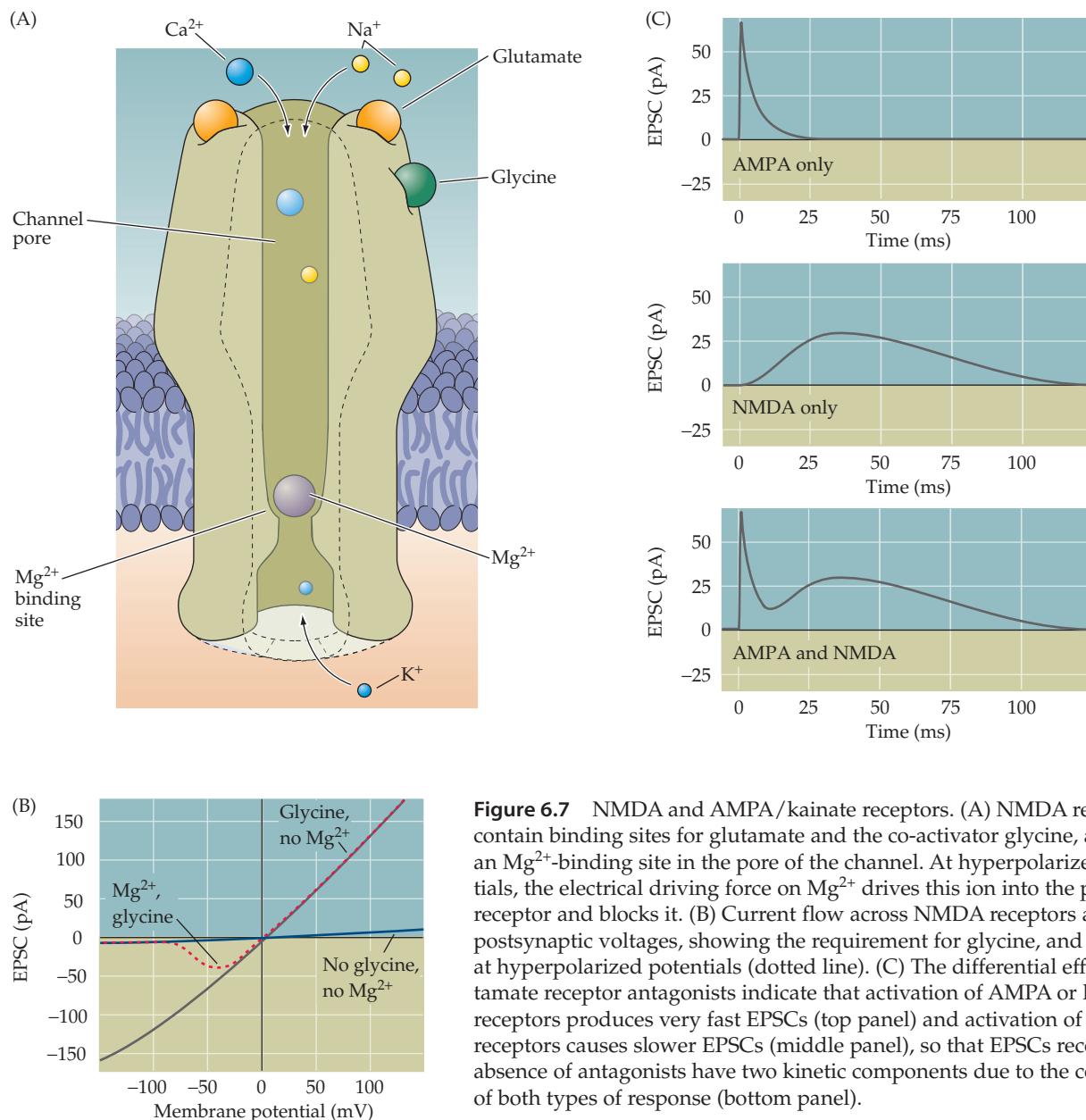


Figure 6.7 NMDA and AMPA/kainate receptors. (A) NMDA receptors contain binding sites for glutamate and the co-activator glycine, as well as an Mg^{2+} -binding site in the pore of the channel. At hyperpolarized potentials, the electrical driving force on Mg^{2+} drives this ion into the pore of the receptor and blocks it. (B) Current flow across NMDA receptors at a range of postsynaptic voltages, showing the requirement for glycine, and Mg^{2+} block at hyperpolarized potentials (dotted line). (C) The differential effects of glutamate receptor antagonists indicate that activation of AMPA or kainate receptors produces very fast EPSCs (top panel) and activation of NMDA receptors causes slower EPSCs (middle panel), so that EPSCs recorded in the absence of antagonists have two kinetic components due to the contribution of both types of response (bottom panel).

only during depolarization of the postsynaptic cell, due to either activation of a large number of excitatory inputs and/or by repetitive firing of action potentials in the presynaptic cell. These properties are widely thought to be the basis for some forms of information storage at synapses, such as memory, as described in Chapter 24. Another unusual property of NMRA receptors is that opening the channel of this receptor requires the presence of a co-agonist, the amino acid glycine (Figure 6.7A,B). There are at least five forms of NMDA receptor subunits (NMDA-R1, and NMDA-R2A through NMDA-R2D); different synapses have distinct combinations of these subunits, producing a variety of NMDA receptor-mediated postsynaptic responses.

Whereas some glutamatergic synapses have only AMPA or NMDA receptors, most possess both AMPA and NMDA receptors. An antagonist of

NMDA receptors, APV (2-amino-5-phosphono-valerate), is often used to differentiate between the two receptor types. The use of this drug has also revealed differences between the EPSPs produced by NMDA and those produced by AMPA/kainate receptors, such as the fact that the synaptic currents produced by NMDA receptors are slower and longer-lasting than those produced by AMPA/kainate receptors (see Figure 6.7C).

In addition to these ionotropic glutamate receptors, there are three types of metabotropic glutamate receptor (mGluRs) (Figure 6.5). These receptors, which modulate postsynaptic ion channels indirectly, differ in their coupling to intracellular signal transduction pathways (see Chapter 7) and in their sensitivity to pharmacological agents. Activation of many of these receptors leads to inhibition of postsynaptic Ca^{2+} and Na^+ channels. Unlike the excitatory ionotropic glutamate receptors, mGluRs cause slower postsynaptic responses that can either increase or decrease the excitability of postsynaptic cells. As a result the physiological roles of mGluRs are quite varied.

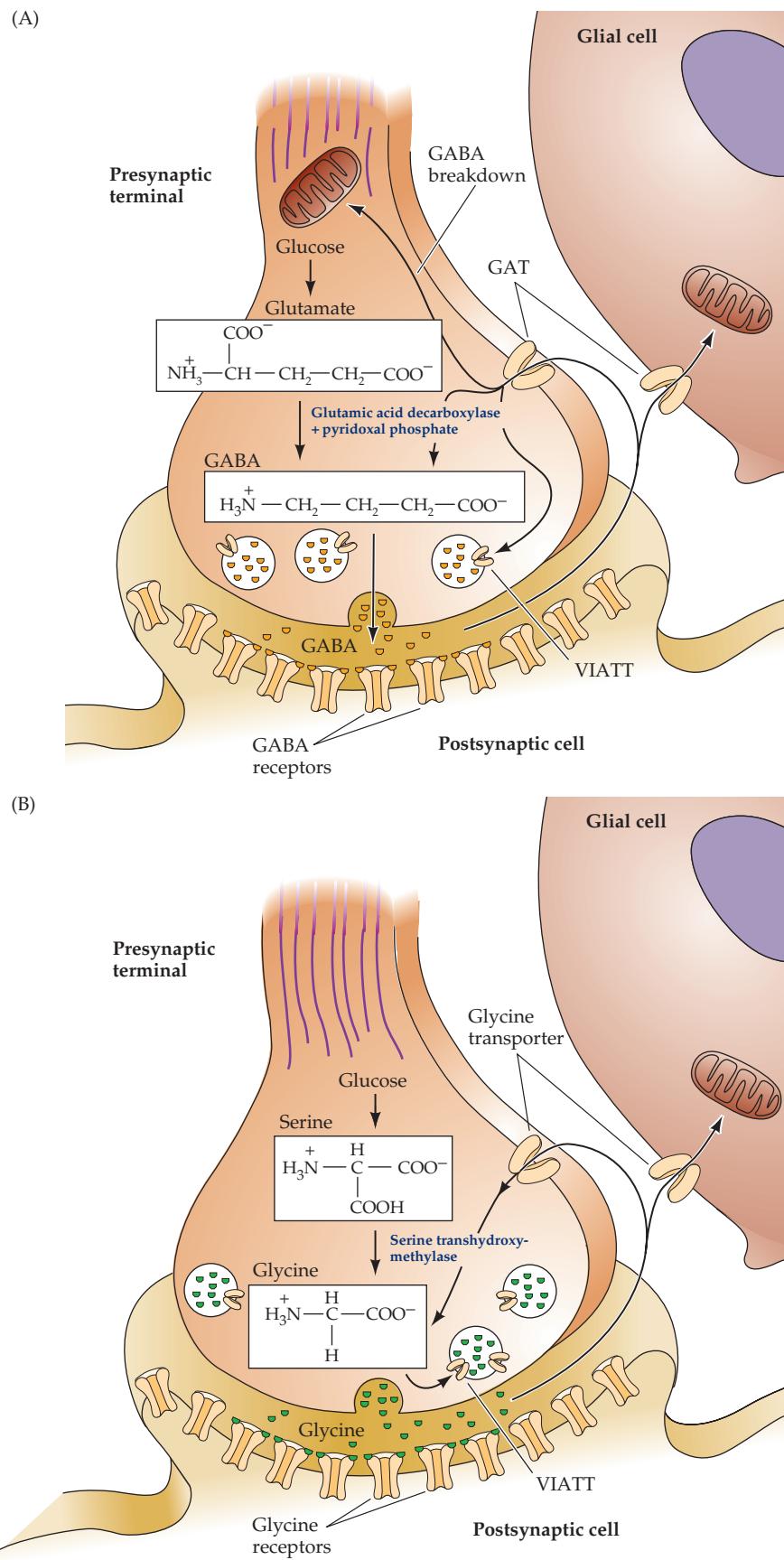
GABA and Glycine

Most inhibitory synapses in the brain and spinal cord use either γ -aminobutyric acid (GABA) or glycine as neurotransmitters. Like glutamate, GABA was identified in brain tissue during the 1950s. The details of its synthesis and degradation were worked out shortly thereafter by the work of Ernst Florey and Eugene Roberts. During this era, David Curtis and Jeffrey Watkins first showed that GABA can inhibit action potential firing in mammalian neurons. Subsequent studies by Edward Kravitz and colleagues established that GABA serves as an inhibitory transmitter at lobster neuromuscular synapses. It is now known that as many as a third of the synapses in the brain use GABA as their inhibitory neurotransmitter. GABA is most commonly found in local circuit interneurons, although cerebellar Purkinje cells provide an example of a GABAergic projection neuron (see Chapter 18).

The predominant precursor for GABA synthesis is glucose, which is metabolized to glutamate by the tricarboxylic acid cycle enzymes (pyruvate and glutamine can also act as precursors). The enzyme glutamic acid decarboxylase (GAD), which is found almost exclusively in GABAergic neurons, catalyzes the conversion of glutamate to GABA (Figure 6.8A). GAD requires a cofactor, pyridoxal phosphate, for activity. Because pyridoxal phosphate is derived from vitamin B_6 , a B_6 deficiency can lead to diminished GABA synthesis. The significance of this became clear after a disastrous series of infant deaths was linked to the omission of vitamin B_6 from infant formula. This lack of B_6 resulted in a large reduction in the GABA content of the brain, and the subsequent loss of synaptic inhibition caused seizures that in some cases were fatal. Once GABA is synthesized, it is transported into synaptic vesicles via a vesicular inhibitory amino acid transporter (VIATT).

The mechanism of GABA removal is similar to that for glutamate: Both neurons and glia contain high-affinity transporters for GABA, termed GATs (several forms of GAT have been identified). Most GABA is eventually converted to succinate, which is metabolized further in the tricarboxylic acid cycle that mediates cellular ATP synthesis. The enzymes required for this degradation, GABA transaminase and succinic semialdehyde dehydrogenase, are mitochondrial enzymes. Inhibition of GABA breakdown causes a rise in tissue GABA content and an increase in the activity of inhibitory neurons. There are also other pathways for degradation of GABA. The most noteworthy of these results in the production of γ -hydroxybutyrate, a GABA derivative that has been abused as a "date rape" drug. Oral adminis-

Figure 6.8 Synthesis, release, and reuptake of the inhibitory neurotransmitters GABA and glycine. (A) GABA is synthesized from glutamate by the enzyme glutamic acid decarboxylase, which requires pyridoxal phosphate. (B) Glycine can be synthesized by a number of metabolic pathways; in the brain, the major precursor is serine. High-affinity transporters terminate the actions of these transmitters and return GABA or glycine to the synaptic terminals for reuse, with both transmitters being loaded into synaptic vesicles via the vesicular inhibitory amino acid transporter (VIATT).



Box D

Excitotoxicity Following Acute Brain Injury

Excitotoxicity refers to the ability of glutamate and related compounds to destroy neurons by prolonged excitatory synaptic transmission. Normally, the concentration of glutamate released into the synaptic cleft rises to high levels (approximately 1 mM), but it remains at this concentration for only a few milliseconds. If abnormally high levels of glutamate accumulate in the cleft, the excessive activation of neuronal glutamate receptors can literally excite neurons to death.

The phenomenon of excitotoxicity was discovered in 1957 when D. R. Lucas and J. P. Newhouse serendipitously found that feeding sodium glutamate to infant mice destroys neurons in the retina. Roughly a decade later, John Olney at Washington University extended this discovery by showing that regions of glutamate-induced neuronal loss can occur throughout the brain. The damage was evidently restricted to the postsynaptic cells—the dendrites of the target neurons were grossly swollen—while the presynaptic terminals were spared. Olney also examined the relative potency of glutamate analogs and found that their neurotoxic actions paralleled their ability to activate postsynaptic glutamate receptors. Furthermore, glutamate receptor antagonists were effective in blocking the neurotoxic effects of glutamate. In light of this evidence, Olney postulated that glutamate destroys neurons by a mechanism similar to transmission at excitatory glutamatergic syn-

apses, and coined the term *excitotoxic* to refer to this pathological effect.

Evidence that excitotoxicity is an important cause of neuronal damage after brain injury has come primarily from studying the consequences of reduced blood flow. The most common cause of reduced blood flow to the brain (ischemia) is the occlusion of a cerebral blood vessel (i.e., a stroke; see Appendix 3). The idea that excessive synaptic activity contributes to ischemic injury emerged from the observation that concentrations of glutamate and aspartate in the extracellular space around neurons increase during ischemia. Moreover, microinjection of glutamate receptor antagonists in experimental animals protects neurons from ischemia-induced damage. Together, these findings imply that extracellular accumulation of glutamate during ischemia activates glutamate receptors excessively, and that this somehow triggers a chain of events that leads to neuronal death. The reduced supply of oxygen and glucose presumably elevates extracellular glutamate levels by slowing the energy-dependent removal of glutamate at synapses.

Excitotoxic mechanisms have now been shown to be involved in other acute forms of neuronal insult, including hypoglycemia, trauma, and repeated intense seizures (called status epilepticus). Understanding excitotoxicity therefore has important implications for treating a variety of neurological disorders. For instance, a blockade of glutamate

receptors could, in principle, protect neurons from injury due to stroke, trauma, or other causes. Unfortunately, clinical trials of glutamate receptor antagonists have not led to much improvement in the outcome of stroke. The ineffectiveness of this quite logical treatment is probably due to several factors, one of which is that substantial excitotoxic injury occurs quite soon after ischemia, prior to the typical initiation of treatment. It is also likely that excitotoxicity is only one of several mechanisms by which ischemia damages neurons, other candidates including damage secondary to inflammation. Pharmacological interventions that target all these mechanisms nonetheless hold considerable promise for minimizing brain injury after stroke and other causes.

References

- LUCAS, D. R. AND J. P. NEWHOUSE (1957) The toxic effects of sodium L-glutamate on the inner layers of the retina. *Arch. Ophthalmol.* 58: 193–201.
- OLNEY, J. W. (1969) Brain lesions, obesity and other disturbances in mice treated with monosodium glutamate. *Science* 164: 719–721.
- OLNEY, J. W. (1971) Glutamate-induced neuronal necrosis in the infant mouse hypothalamus: An electron microscopic study. *J. Neuropathol. Exp. Neurol.* 30: 75–90.
- ROTHMAN, S. M. (1983) Synaptic activity mediates death of hypoxic neurons. *Science* 220: 536–537.
- SYNTICHAKI, P. AND N. TAVERNARAKIS (2003) The biochemistry of neuronal necrosis: Rogue biology? *Nature Neurosci. Rev.* 4: 672–684.

tration of γ -hydroxybutyrate can cause euphoria, memory deficits, and unconsciousness. Presumably these effects arise from actions on GABAergic synapses in the CNS.

Inhibitory synapses employing GABA as their transmitter can exhibit three types of postsynaptic receptors, called GABA_A, GABA_B, and GABA_C. GABA_A and GABA_C receptors are ionotropic receptors, while GABA_B receptors are metabotropic. The ionotropic GABA receptors are usually

inhibitory because their associated channels are permeable to Cl^- (Figure 6.9A); the flow of the negatively charged chloride ions inhibits postsynaptic cells since the reversal potential for Cl^- is more negative than the threshold for neuronal firing (see Figure 5.19B). Like other ionotropic receptors, GABA receptors are pentamers assembled from a combination of five types of subunits (α , β , γ , δ , and ρ ; see Figure 6.4C). As a result of this subunit diversity, as well as variable stoichiometry of subunits, the function of GABA_A receptors differs widely among neuronal types. Drugs that act as agonists or modulators of postsynaptic GABA receptors, such as benzodiazepines and barbiturates, are used clinically to treat epilepsy and are effective sedatives and anesthetics. Binding sites for GABA, barbiturates, steroids, and picrotoxin are all located within the pore domain of the channel (Figure 6.9B). Another site, called the benzodiazepine binding site, lies outside the pore and modulates channel activity. Benzodiazepines, such as diazepam (Valium[®]) and chlordiazepoxide (Librium[®]), are tranquilizing (anxiety reducing) drugs that enhance GABAergic transmission by binding to the α and δ subunits of GABA_A receptors. Barbiturates, such as phenobarbital and pentobarbital, are hypnotics that bind to the α and β subunits of some GABA receptors and are used therapeutically for anesthesia and to control epilepsy. Another drug that can alter the activity of GABA-mediated inhibitory circuits is alcohol; at least some aspects of drunken behavior are caused by the alcohol-mediated alterations in ionotropic GABA receptors.

Metabotropic GABA receptors (GABA_B) are also widely distributed in brain. Like the ionotropic GABA_A receptors, GABA_B receptors are inhibitory. Rather than activating Cl^- selective channels, however, GABA_B -mediated inhibition is due to the activation of K^+ channels. A second mechanism for

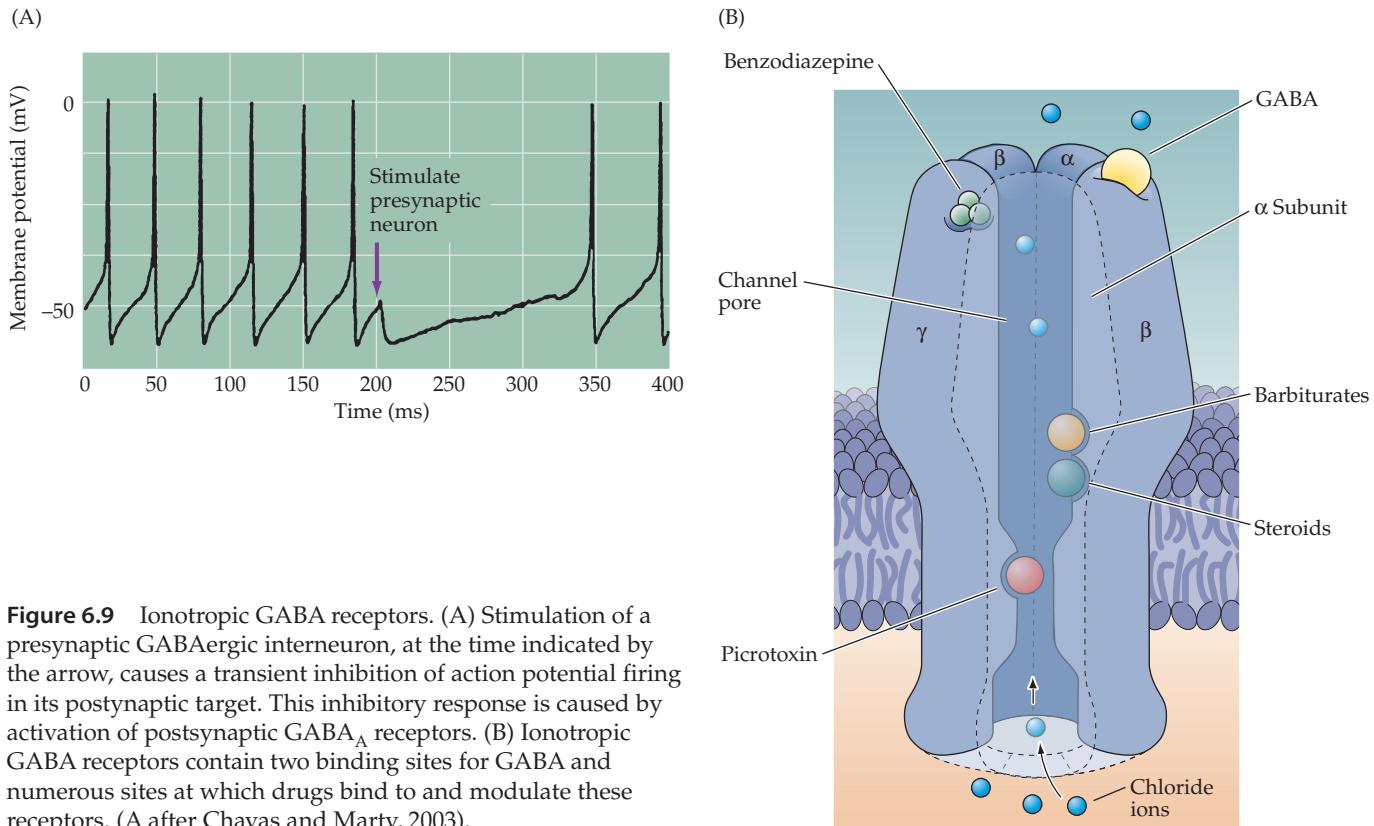


Figure 6.9 Ionotropic GABA receptors. (A) Stimulation of a presynaptic GABAergic interneuron, at the time indicated by the arrow, causes a transient inhibition of action potential firing in its postsynaptic target. This inhibitory response is caused by activation of postsynaptic GABA_A receptors. (B) Ionotropic GABA receptors contain two binding sites for GABA and numerous sites at which drugs bind to and modulate these receptors. (A after Chavas and Marty, 2003).

GABA_B-mediated inhibition is by blocking Ca²⁺ channels, which tends to hyperpolarize postsynaptic cells. Unlike most metabotropic receptors, GABA_B receptors appear to assemble as heterodimers of GABA_BR1 and R2 subunits.

The distribution of the neutral amino acid glycine in the central nervous system is more localized than that of GABA. About half of the inhibitory synapses in the spinal cord use glycine; most other inhibitory synapses use GABA. Glycine is synthesized from serine by the mitochondrial isoform of serine hydroxymethyltransferase (Figure 6.8B), and is transported into synaptic vesicles via the same vesicular inhibitory amino acid transporter that loads GABA into vesicles. Once released from the presynaptic cell, glycine is rapidly removed from the synaptic cleft by the plasma membrane glycine transporters. Mutations in the genes coding for some of these enzymes result in hyperglycinemia, a devastating neonatal disease characterized by lethargy, seizures, and mental retardation.

The receptors for glycine are also ligand-gated Cl⁻ channels, their general structure mirroring that of the GABA_A receptors. Glycine receptors are pentamers consisting of mixtures of the 4 gene products encoding glycine-binding α subunits, along with the accessory β subunit. Glycine receptors are potently blocked by strychnine, which may account for the toxic properties of this plant alkaloid (see Box B).

The Biogenic Amines

Biogenic amine transmitters regulate many brain functions and are also active in the peripheral nervous system. Because biogenic amines are implicated in such a wide range of behaviors (ranging from central homeostatic functions to cognitive phenomena such as attention), it is not surprising that defects in biogenic amines function are implicated in most psychiatric disorders. The pharmacology of amine synapses is critically important in psychotherapy, with drugs affecting the synthesis, receptor binding, or catabolism of these neurotransmitters being among the most important agents in the armamentarium of modern pharmacology (Box E). Many drugs of abuse also act on biogenic amine pathways.

There are five well-established biogenic amine neurotransmitters: the three catecholamines—dopamine, norepinephrine (noradrenaline), and epinephrine (adrenaline)—and histamine and serotonin (see Figure 6.1). All the catecholamines (so named because they share the catechol moiety) are derived from a common precursor, the amino acid tyrosine (Figure 6.10). The first step in catecholamine synthesis is catalyzed by tyrosine hydroxylase in a reaction requiring oxygen as a co-substrate and tetrahydrobiopterin as a cofactor to synthesize dihydroxyphenylalanine (DOPA). Histamine and serotonin are synthesized via other routes, as described below.

- Dopamine is present in several brain regions (Figure 6.11A), although the major dopamine-containing area of the brain is the corpus striatum, which receives major input from the substantia nigra and plays an essential role in the coordination of body movements. In Parkinson's disease, for instance, the dopaminergic neurons of the substantia nigra degenerate, leading to a characteristic motor dysfunction (see Box B in Chapter 17). Dopamine is also believed to be involved in motivation, reward, and reinforcement, and many drugs of abuse work by affecting dopaminergic synapses in the CNS (see Box A). In addition to these roles in the CNS, dopamine also plays a poorly understood role in some sympathetic ganglia.

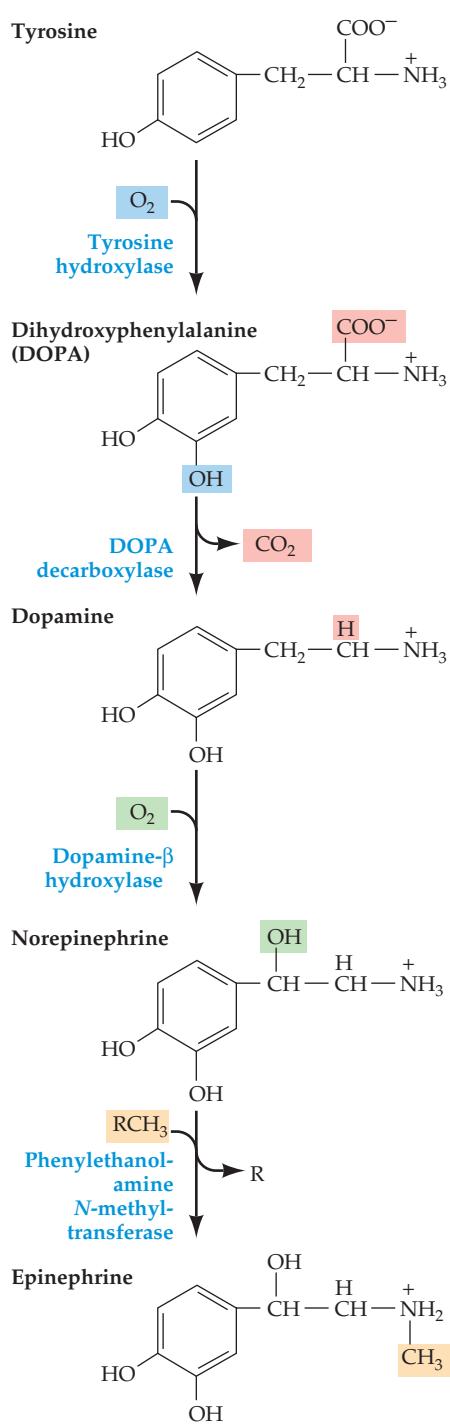


Figure 6.10 The biosynthetic pathway for the catecholamine neurotransmitters. The amino acid tyrosine is the precursor for all three catecholamines. The first step in this reaction pathway, catalyzed by tyrosine hydroxylase, is rate-limiting.

Box E

Biogenic Amine Neurotransmitters and Psychiatric Disorders

The regulation of the biogenic amine neurotransmitters is altered in a variety of psychiatric disorders. Indeed, most psychotropic drugs (defined as drugs that alter behavior, mood, or perception) selectively affect one or more steps in the synthesis, packaging, or degradation of biogenic amines. Sorting out how these drugs work has been extremely useful in beginning to understand the molecular mechanisms underlying some of these diseases.

Based on their effects on humans, psychotherapeutic drugs can be divided into several broad categories: antipsychotics, antianxiety drugs, antidepressants, and stimulants. The first antipsychotic drug used to ameliorate disorders such as schizophrenia was reserpine. Reserpine was developed in the 1950s and initially used as an antihypertensive agent; it blocks the uptake of norepinephrine into synaptic vesicles and therefore depletes the transmitter at aminergic terminals, diminishing the ability of the sympathetic division of the visceral motor system to cause vasoconstriction (see Chapter 20). A major side effect in hypertensive patients treated with reserpine—behavioral depression—suggested the possibility of using it as an antipsychotic agent in patients suffering from agitation and pathological anxiety. (Its ability to cause depression in mentally healthy individuals also suggested that aminergic transmitters are involved in mood disorders; see Box E in Chapter 28.)

Although reserpine is no longer used as an antipsychotic agent, its initial success stimulated the development of antipsychotic drugs such as chlorpromazine, haloperidol, and benperidol, which over the last several decades have radically changed the approach to treating psychotic disorders. Prior to the discovery of these drugs, psychotic patients were typically hospitalized for

long periods, sometimes indefinitely, and in the 1940s were subjected to desperate measures such as frontal lobotomy (see Box B in Chapter 25). Modern antipsychotic drugs now allow most patients to be treated on an outpatient basis after a brief hospital stay. Importantly, the clinical effectiveness of these drugs is correlated with their ability to block brain dopamine receptors, implying that activation of dopamine receptors contributes to some types of psychotic illness. A great deal of effort continues to be expended on developing more effective antipsychotic drugs with fewer side effects, and on discovering the mechanism and site of action of these medications.

The second category of psychotherapeutic drugs is the antianxiety agents. Anxiety disorders are estimated to afflict 10–35% of the population, making them the most common psychiatric problem. The two major forms of pathological anxiety—panic attacks and generalized anxiety disorder—both respond to drugs that affect aminergic transmission. The agents used to treat panic disorders include inhibitors of the enzyme monoamine oxidase (MAO inhibitors) required for the catabolism of the amine neurotransmitters, and blockers of serotonin receptors. The most effective drugs in treating generalized anxiety disorder have been benzodiazepines, such as chlordiazepoxide (Librium[®]), and diazepam (Valium[®]). In contrast to most other psychotherapeutic drugs, these agents increase the efficacy of transmission at GABA_A synapses rather than acting at aminergic synapses.

Antidepressants and stimulants also affect aminergic transmission. A large number of drugs are used clinically to treat depressive disorders. The three major classes of antidepressants—MAO inhibitors, tricyclic antidepressants, and

serotonin uptake blockers such as fluoxetine (Prozac[®]) and trazodone—all influence various aspects of aminergic transmission. MAO inhibitors such as phenelzine block the breakdown of amines, whereas the tricyclic antidepressants such as desipramine block the reuptake of norepinephrine and other amines. The extraordinarily popular antidepressant fluoxetine (Prozac[®]) selectively blocks the reuptake of serotonin without affecting the reuptake of catecholamines. Stimulants such as amphetamine are also used to treat some depressive disorders. Amphetamine stimulates the release of norepinephrine from nerve terminals; the transient “high” resulting from taking amphetamine may reflect the emotional opposite of the depression that sometimes follows reserpine-induced norepinephrine depletion.

Despite the relatively small number of aminergic neurons in the brain, this litany of pharmacological actions emphasizes that these neurons are critically important in the maintenance of mental health.

References

- FRANKLE, W. G., J. LERMA AND M. LARUELLE (2003) The synaptic hypothesis of schizophrenia. *Neuron* 39: 205–216.
- FREEDMAN, R. (2003) Schizophrenia. *N. Engl. J. Med.* 349: 1738–1749.
- LEWIS, D. A. AND P. LEVITT (2002) Schizophrenia as a disorder of neurodevelopment. *Annu. Rev. Neurosci.* 25: 409–432.
- NESTLER, E. J., M. BARROT, R. J. DiLEONE, A. J. EISCH, S. J. GOLD AND L. M. MONTEGGIA (2002) Neurobiology of depression. *Neuron* 34: 13–25.

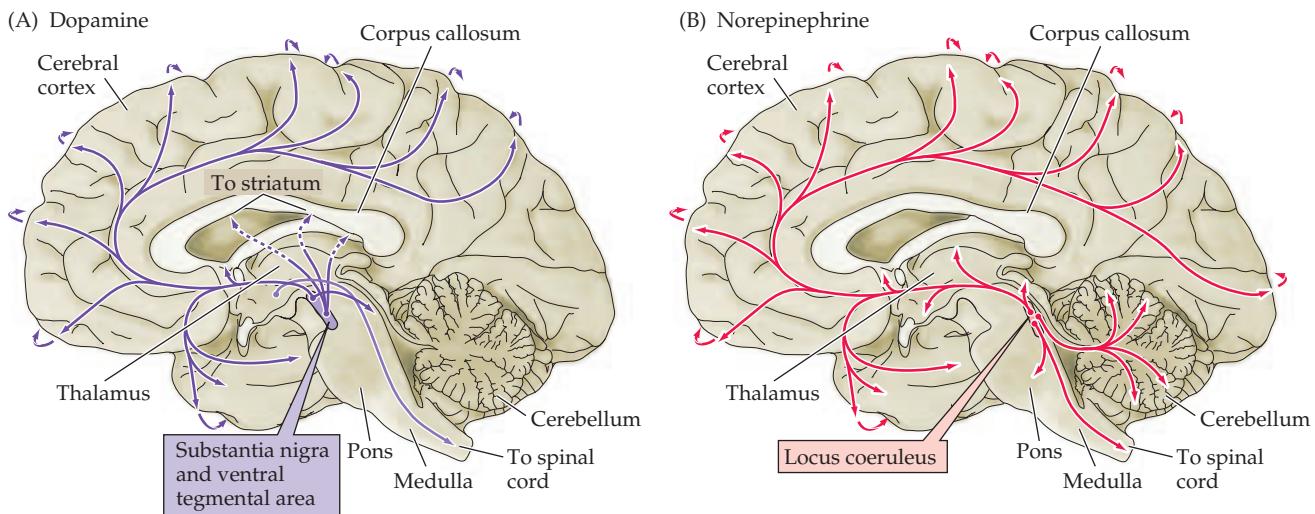


Figure 6.11 The distribution in the human brain of neurons and their projections (arrows) containing catecholamine neurotransmitters. Curved arrows along the perimeter of the cortex indicate the innervation of lateral cortical regions not shown in this midsagittal plane of section.



Dopamine is produced by the action of DOPA decarboxylase on DOPA (see Figure 6.10). Following its synthesis in the cytoplasm of presynaptic terminals, dopamine is loaded into synaptic vesicles via a vesicular monoamine transporter (VMAT). Dopamine action in the synaptic cleft is terminated by reuptake of dopamine into nerve terminals or surrounding glial cells by a Na^+ -dependent dopamine transporter, termed DAT. Cocaine apparently produces its psychotropic effects by binding to and inhibiting DAT, yielding a net increase in dopamine release from specific brain areas. Amphetamine, another addictive drug, also inhibits DAT as well as the transporter for norepinephrine (see below). The two major enzymes involved in the catabolism of dopamine are monoamine oxidase (MAO) and catechol O-methyltransferase (COMT). Both neurons and glia contain mitochondrial MAO and cytoplasmic COMT. Inhibitors of these enzymes, such as phenelzine and tranylcypromine, are used clinically as antidepressants (see Box E).

Once released, dopamine acts exclusively by activating G-protein-coupled receptors. These are mainly dopamine-specific receptors, although β -adrenergic receptors also serve as important targets of norepinephrine and epinephrine (see below). Most dopamine receptor subtypes (see Figure 6.5B)

act by either activating or inhibiting adenylyl cyclase (see Chapter 7). Activation of these receptors generally contribute to complex behaviors; for example, administration of dopamine receptor agonists elicits hyperactivity and repetitive, stereotyped behavior in laboratory animals. Activation of another type of dopamine receptor in the medulla inhibits vomiting. Thus, antagonists of these receptors are used as emetics to induce vomiting after poisoning or a drug overdose. Dopamine receptor antagonists can also elicit catalepsy, a state in which it is difficult to initiate voluntary motor movement, suggesting a basis for this aspect of some psychoses.

- *Norepinephrine* (also called noradrenaline) is used as a neurotransmitter in the locus coeruleus, a brainstem nucleus that projects diffusely to a variety of forebrain targets (Figure 6.11B) and influences sleep and wakefulness, attention, and feeding behavior. Perhaps the most prominent noradrenergic neurons are sympathetic ganglion cells, which employ norepinephrine as the major peripheral transmitter in this division of the visceral motor system (see Chapter 20).

Norepinephrine synthesis requires dopamine β -hydroxylase, which catalyzes the production of norepinephrine from dopamine (see Figure 6.10). Norepinephrine is then loaded into synaptic vesicles via the same VMAT involved in vesicular dopamine transport. Norepinephrine is cleared from the synaptic cleft by the norepinephrine transporter (NET), which also is capable of taking up dopamine. As mentioned, NET serves as a molecular target of amphetamine, which acts as a stimulant by producing a net increase in the release of norepinephrine and dopamine. A mutation in the NET gene is a cause of orthostatic intolerance, a disorder that produces lightheadedness while standing up. Like dopamine, norepinephrine is degraded by MAO and COMT.

Norepinephrine, as well as epinephrine, acts on α - and β -adrenergic receptors (Figure 6.5B). Both types of receptor are G-protein-coupled; in fact, the β -adrenergic receptor was the first identified metabotropic neurotransmitter receptor. Two subclasses of α -adrenergic receptors are now known. Activation of α_1 receptors usually results in a slow depolarization linked to the inhibition of K^+ channels, while activation of α_2 receptors produces a slow hyperpolarization due to the activation of a different type of K^+ channel. There are three subtypes of β -adrenergic receptor, two of which are expressed in many types of neurons. Agonists and antagonists of adrenergic receptors, such as the β blocker propanolol (Inderol[®]), are used clinically for a variety of conditions ranging from cardiac arrhythmias to migraine headaches. However, most of the actions of these drugs are on smooth muscle receptors, particularly in the cardiovascular and respiratory systems (see Chapter 20).

- *Epinephrine* (also called adrenaline) is found in the brain at lower levels than the other catecholamines and also is present in fewer brain neurons than other catecholamines. Epinephrine-containing neurons in the central nervous system are primarily in the lateral tegmental system and in the medulla and project to the hypothalamus and thalamus (Figure 6.11C). The function of these epinephrine-secreting neurons is not known.

The enzyme that synthesizes epinephrine, phenylethanolamine-*N*-methyltransferase (see Figure 6.10), is present only in epinephrine-secreting neurons. Otherwise, the metabolism of epinephrine is very similar to that of norepinephrine. Epinephrine is loaded into vesicles via the VMAT. No plasma membrane transporter specific for epinephrine has been identified, though the NET is capable of transporting epinephrine. As already noted, epinephrine acts on both α - and β -adrenergic receptors.

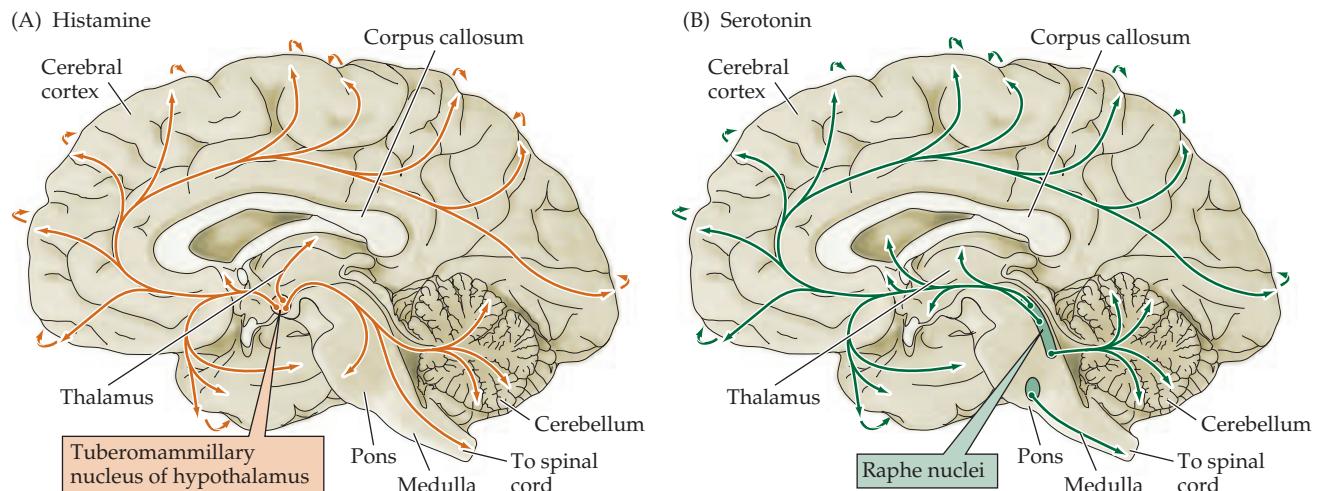


Figure 6.12 The distribution in the human brain of neurons and their projections (arrows) containing histamine (A) or serotonin (B). Curved arrows along the perimeter of the cortex indicate the innervation of lateral cortical regions not shown in this midsagittal plane of section.



- *Histamine* is found in neurons in the hypothalamus that send sparse but widespread projections to almost all regions of the brain and spinal cord (Figure 6.12A). The central histamine projections mediate arousal and attention, similar to central ACh and norepinephrine projections. Histamine also controls the reactivity of the vestibular system. Allergic reactions or tissue damage cause release of histamine from mast cells in the bloodstream. The close proximity of mast cells to blood vessels, together with the potent actions of histamine on blood vessels, also raises the possibility that histamine may influence brain blood flow.

Histamine is produced from the amino acid histidine by a histidine decarboxylase (Figure 6.13A) and is transported into vesicles via the same VMAT as the catecholamines. No plasma membrane histamine transporter has been identified yet. Histamine is degraded by the combined actions of histamine methyltransferase and MAO.

There are three known types of histamine receptors, all of which are G-protein-coupled receptors (Figure 6.5B). Because of the importance of histamine receptors in the mediation of allergic responses, many histamine receptor antagonists have been developed as antihistamine agents. Antihistamines that cross the blood-brain barrier, such as diphenhydramine (Benadryl®), act as sedatives by interfering with the roles of histamine in CNS arousal. Antagonists of the H₁ receptor also are used to prevent motion sickness, perhaps because of the role of histamine in controlling vestibular function. H₂ receptors control the secretion of gastric acid in the digestive system, allowing H₂ receptor antagonists to be used in the treatment of a variety of upper gastrointestinal disorders (e.g., peptic ulcers).

- *Serotonin*, or 5-hydroxytryptamine (5-HT), was initially thought to increase vascular tone by virtue of its presence in serum (hence the name serotonin). Serotonin is found primarily in groups of neurons in the raphe region of the pons and upper brainstem, which have widespread projections to the forebrain (see Figure 6.12B) and regulate sleep and wakefulness (see Chapter 27). 5-HT occupies a place of prominence in neuropharmacology because a large number of antipsychotic drugs that are valuable in the treatment of depression and anxiety act on serotonergic pathways (see Box E).

5-HT is synthesized from the amino acid tryptophan, which is an essential dietary requirement. Tryptophan is taken up into neurons by a plasma mem-

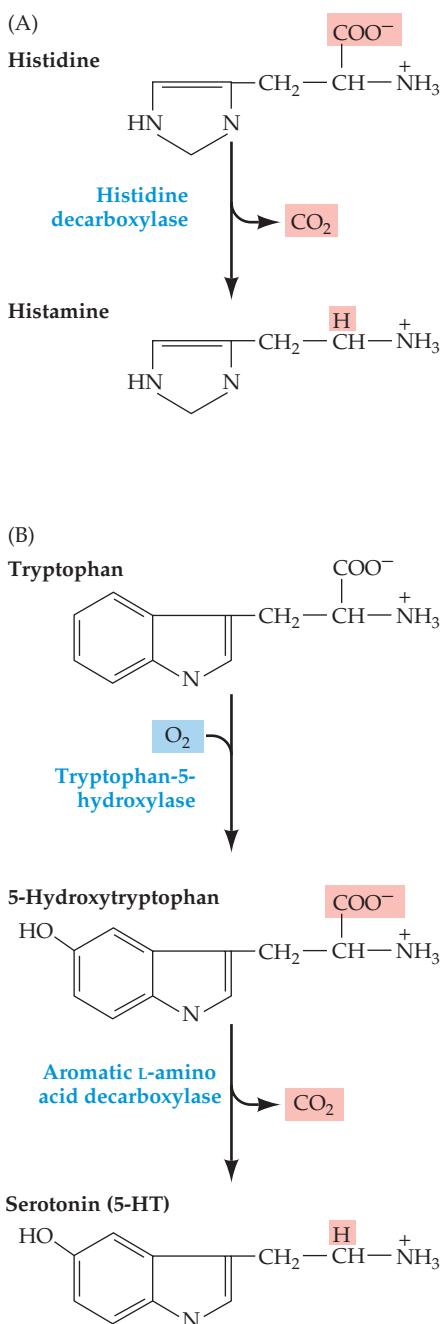


Figure 6.13 Synthesis of histamine and serotonin. (A) Histamine is synthesized from the amino acid histidine. (B) Serotonin is derived from the amino acid tryptophan by a two-step process that requires the enzymes tryptophan-5-hydroxylase and a decarboxylase.

brane transporter and hydroxylated in a reaction catalyzed by the enzyme tryptophan-5-hydroxylase (Figure 6.13B), the rate-limiting step for 5-HT synthesis. Loading of 5-HT into synaptic vesicles is done by the VMAT that is also responsible for loading of other monoamines into synaptic vesicles. The synaptic effects of serotonin are terminated by transport back into nerve terminals via a specific serotonin transporter (SERT). Many antidepressant drugs are selective serotonin reuptake inhibitors (SSRIs) that inhibit transport of 5-HT by SERT. Perhaps the best known example of an SSRI is Prozac (see Box E). The primary catabolic pathway for 5-HT is mediated by MAO.

A large number of 5-HT receptors have been identified. Most 5-HT receptors are metabotropic (see Figure 6.5B). These have been implicated in behaviors, including the emotions, circadian rhythms, motor behaviors, and state of mental arousal. Impairments in the function of these receptors have been implicated in numerous psychiatric disorders, such as depression, anxiety disorders, and schizophrenia (see Chapter 28), and drugs acting on serotonin receptors are effective treatments for a number of these conditions. Activation of 5-HT receptors also mediates satiety and decreased food consumption, which is why serotonergic drugs are sometimes useful in treating eating disorders.

Only one group of serotonin receptors, called the 5-HT₃ receptors, are ligand-gated ion channels (see Figure 6.4C). These are non-selective cation channels and therefore mediate excitatory postsynaptic responses. Their general structure, with functional channels formed by assembly of multiple subunits, is similar to the other ionotropic receptors described in the chapter. Two types of 5-HT₃ subunit are known, and form functional channels by assembling as a heteromultimer. 5-HT receptors are targets for a wide variety of therapeutic drugs including ondansetron (Zofran®) and granisetron (Kytril®), which are used to prevent postoperative nausea and chemotherapy-induced emesis.

ATP and Other Purines

Interestingly, all synaptic vesicles contain ATP, which is co-released with one or more “classical” neurotransmitters. This observation raises the possibility that ATP acts as a co-transmitter. It has been known since the 1920s that the extracellular application of ATP (or its breakdown products AMP and adenosine) can elicit electrical responses in neurons. The idea that some purines (so named because all these compounds contain a purine ring; see Figure 6.1) are also neurotransmitters has now received considerable experimental support. ATP acts as an excitatory neurotransmitter in motor neurons of the spinal cord, as well as sensory and autonomic ganglia. Postsynaptic actions of ATP have also been demonstrated in the central nervous system, specifically for dorsal horn neurons and in a subset of hippocampal neurons. Adenosine, however, cannot be considered a classical neurotransmitter because it is not stored in synaptic vesicles or released in a Ca²⁺-dependent manner. Rather, it is generated from ATP by the action of extracellular enzymes. A number of enzymes, such as apyrase and ecto-5' nucleotidase, as well as nucleoside transporters are involved in the rapid catabolism and removal of purines from extracellular locations. Despite the relative novelty of this evidence, it suggests that excitatory transmission via purinergic synapses is widespread in the mammalian brain.

In accord with this evidence, receptors for both ATP and adenosine are widely distributed in the nervous system, as well as many other tissues.

Three classes of these purinergic receptors are now known. One of these classes consists of ligand-gated ion channels (see Figure 6.4C); the other two are G-protein-coupled metabotropic receptors (see Figure 6.5B). Like many ionotropic transmitter receptors, the ligand-gated channels are nonselective cation channels that mediate excitatory postsynaptic responses. The genes encoding these channels, however, are unique in that they appear to have only two transmembrane domains. Ionotropic purinergic receptors are widely distributed in central and peripheral neurons. In sensory nerves they evidently play a role in mechanosensation and pain; their function in most other cells, however, is not known.

The two types of metabotropic receptors activated by purines differ in their sensitivity to agonists: One type is preferentially stimulated by adenosine, whereas the other is preferentially activated by ATP. Both receptor types are found throughout the brain, as well as in peripheral tissues such as the heart, adipose tissue, and the kidney. Xanthines such as caffeine and theophylline block adenosine receptors, and this activity is thought to be responsible for the stimulant effects of these agents.

Peptide Neurotransmitters

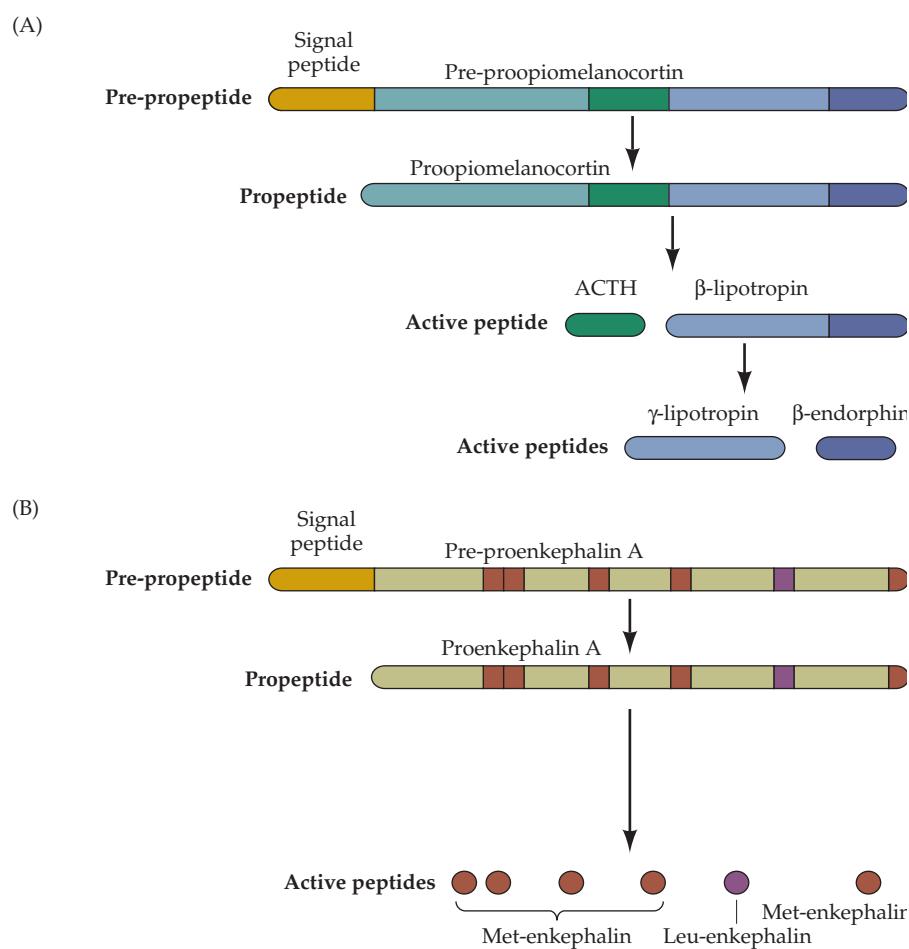
Many peptides known to be hormones also act as neurotransmitters. Some peptide transmitters have been implicated in modulating emotions (see Chapter 28). Others, such as substance P and the opioid peptides, are involved in the perception of pain (see Chapter 9). Still other peptides, such as melanocyte-stimulating hormone, adrenocorticotropin, and β -endorphin, regulate complex responses to stress.

The mechanisms responsible for the synthesis and packaging of peptide transmitters are fundamentally different from those used for the small-molecule neurotransmitters and are much like the synthesis of proteins that are secreted from non-neuronal cells (pancreatic enzymes, for instance). Peptide-secreting neurons generally synthesize polypeptides in their cell bodies that are much larger than the final, "mature" peptide. Processing these polypeptides in their cell bodies, which are called **pre-propeptides** (or pre-proproteins), takes place by a sequence of reactions in several intracellular organelles. Pre-propeptides are synthesized in the rough endoplasmic reticulum, where the signal sequence of amino acids—that is, the sequence indicating that the peptide is to be secreted—is removed. The remaining polypeptide, called a **propeptide** (or proprotein), then traverses the Golgi apparatus and is packaged into vesicles in the *trans*-Golgi network. The final stages of peptide neurotransmitter processing occur after packaging into vesicles and involve proteolytic cleavage, modification of the ends of the peptide, glycosylation, phosphorylation, and disulfide bond formation.

Propeptide precursors are typically larger than their active peptide products and can give rise to more than one species of neuropeptide (Figure 6.14). This means that multiple neuroactive peptides can be released from a single vesicle. In addition, neuropeptides often are co-released with small-molecule neurotransmitters. Thus, peptidergic synapses often elicit complex postsynaptic responses. Peptides are catabolized into inactive amino acid fragments by enzymes called peptidases, usually located on the extracellular surface of the plasma membrane.

The biological activity of the peptide neurotransmitters depends on their amino acid sequence (Figure 6.15). Based on their amino acid sequences, neuropeptide transmitters have been loosely grouped into five categories:

Figure 6.14 Proteolytic processing of the pre-propeptides pre-proopiomelanocortin (A) and pre-proenkephalin A (B). For each pre-propeptide, the signal sequence is indicated in orange at the left; the locations of active peptide products are indicated by different colors. The maturation of the pre-propeptides involves cleaving the signal sequence and other proteolytic processing. Such processing can result in a number of different neuroactive peptides such as ACTH, γ -lipotropin, and β -endorphin (A), or multiple copies of the same peptide, such as met-enkephalin (B).



the brain/gut peptides, opioid peptides, pituitary peptides, hypothalamic releasing hormones, and a catch-all category containing other peptides that are not easily classified.

Substance P is an example of the first of these categories (Figure 6.15A). The study of neuropeptides actually began more than 60 years ago with the accidental discovery of substance P, a powerful hypotensive agent. (The peculiar name derives from the fact that this molecule was an unidentified component of *powder* extracts from brain and intestine.) This 11-amino-acid peptide (see Figure 6.15) is present in high concentrations in the human hippocampus, neocortex, and also in the gastrointestinal tract; hence its classification as a brain/gut peptide. It is also released from C fibers, the small-diameter afferents in peripheral nerves that convey information about pain and temperature (as well as postganglionic autonomic signals). Substance P is a sensory neurotransmitter in the spinal cord, where its release can be inhibited by opioid peptides released from spinal cord interneurons, resulting in the suppression of pain (see Chapter 9). The diversity of neuropeptides is highlighted by the finding that the gene coding for substance P encodes a number of other neuroactive peptides including neuropeptide A, neuropeptide K, and neuropeptide γ .

An especially important category of peptide neurotransmitters is the family of opioids (Figure 6.15B). These peptides are so named because they bind

(A) Brain-gut peptides



(B) Opioid peptides



Amino acid properties

- Hydrophobic
- Polar, uncharged
- Acidic
- Basic

(C) Pituitary peptides



(D) Hypothalamic-releasing peptides



(E) Miscellaneous peptides



to the same postsynaptic receptors activated by opium. The opium poppy has been cultivated for at least 5000 years, and its derivatives have been used as an analgesic since at least the Renaissance. The active ingredients in opium are a variety of plant alkaloids, predominantly morphine. Morphine, named for Morpheus, the Greek god of dreams, is still one of the most effective analgesics in use today, despite its addictive potential (see Box A). Synthetic opiates such as meperidine and methadone are also used as analgesics, and fentanyl, a drug with 80 times the analgesic potency of morphine, is widely used in clinical anesthesiology.

The opioid peptides were discovered in the 1970s during a search for endorphins, *endogenous* compounds that mimicked the actions of *morphine*. It was hoped that such compounds would be analgesics, and that understanding them would shed light on drug addiction. The endogenous ligands of the opioid receptors have now been identified as a family of more than 20 opioid peptides that fall into three classes: the endorphins, the enkephalins, and the dynorphins (Table 6.2). Each of these classes are liberated from an inactive pre-propeptide (pre-proopiomelanocortin, pre-proenkephalin A, and pre-prodynorphin), derived from distinct genes (see Figure 6.14). Opioid precursor processing is carried out by tissue-specific processing enzymes that are packaged into vesicles, along with the precursor peptide, in the Golgi apparatus.

Figure 6.15 Neuropeptides vary in length, but usually contain between 3 and 36 amino acids. The sequence of amino acids determines the biological activity of each peptide.

TABLE 6.2
Endogenous Opioid Peptides

Name	Amino acid sequence ^a
Endorphins	
α -Endorphin	<i>Tyr-Gly-Gly-Phe-Met-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr</i>
α -Neoendorphin	<i>Tyr-Gly-Gly-Phe-Leu-Arg-Lys-Tyr-Pro-Lys</i>
β -Endorphin	<i>Tyr-Gly-Gly-Phe-Met-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr-Leu-Phe-Lys-Asn-Ala-Ile-Val-Lys-Asn-Ala-His-Lys-Gly-Gln</i>
γ -Endorphin	<i>Tyr-Gly-Gly-Phe-Met-Thr-Ser-Glu-Lys-Ser-Gln-Thr-Pro-Leu-Val-Thr-Leu</i>
Enkephalins	
Leu-enkephalin	<i>Tyr-Gly-Gly-Phe-Leu</i>
Met-enkephalin	<i>Tyr-Gly-Gly-Phe-Met</i>
Dynorphins	
Dynorphin A	<i>Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Ile-Arg-Pro-Lys-Leu-Lys-Trp-Asp-Asn-Gln</i>
Dynorphin B	<i>Tyr-Gly-Gly-Phe-Leu-Arg-Arg-Gln-Phe-Lys-Val-Val-Thr</i>

^aNote the initial homology, indicated by italics.

Opioid peptides are widely distributed throughout the brain and are often co-localized with other small-molecule neurotransmitters, such as GABA and 5-HT. In general, these peptides tend to be depressants. When injected intracerebrally in experimental animals, they act as analgesics; on the basis of this and other evidence, opioids are likely to be involved in the mechanisms underlying acupuncture-induced analgesia. Opioids are also involved in complex behaviors such as sexual attraction and aggressive/submissive behaviors. They have also been implicated in psychiatric disorders such as schizophrenia and autism, although the evidence for this is debated. Unfortunately, the repeated administration of opioids leads to tolerance and addiction.

Virtually all neuropeptides initiate their effects by activating G-protein-coupled receptors. The study of these metabotropic peptide receptors in the brain has been difficult because few specific agonists and antagonists are known. Peptides activate their receptors at low (nM to μ M) concentrations compared to the concentrations required to activate receptors for small-molecule neurotransmitters. These properties allow the postsynaptic targets of peptides to be quite far removed from presynaptic terminals and to modulate the electrical properties of neurons that are simply in the vicinity of the site of peptide release. Neuropeptide receptor activation is especially important in regulating the postganglionic output from sympathetic ganglia and the activity of the gut (see Chapter 20). Peptide receptors, particularly the neuropeptide Y receptor, are also implicated in the initiation and maintenance of feeding behavior leading to satiety or obesity.

Other behaviors ascribed to peptide receptor activation include anxiety and panic attacks, and antagonists of cholecystokinin receptors are clinically useful in the treatment of these afflictions. Other useful drugs have been developed by targeting the opiate receptors. Three well-defined opioid receptor subtypes (μ , δ , and κ) play a role in reward mechanisms as well as addiction. The μ -opiate receptor has been specifically identified as the primary site for drug reward mediated by opiate drugs.

Unconventional Neurotransmitters

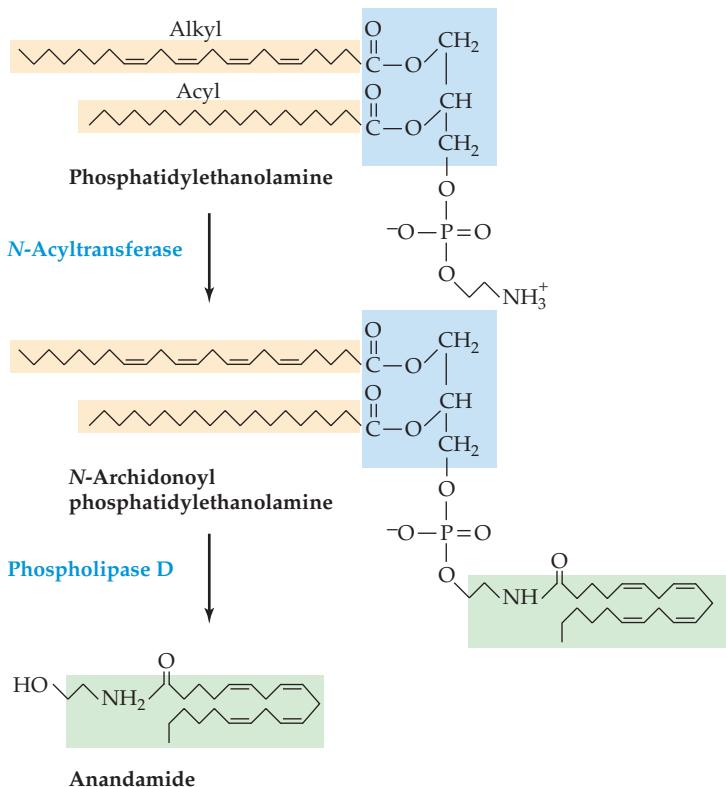
In addition to the conventional neurotransmitters already described, some unusual molecules are also used for signaling between neurons and their targets. These chemical signals can be considered as neurotransmitters because of their roles in interneuronal signaling and because their release from neurons is regulated by Ca^{2+} . However, they are unconventional, in comparison to other neurotransmitters, because they are not stored in synaptic vesicles and are not released from presynaptic terminals via exocytotic mechanisms. In fact, these unconventional neurotransmitters need not be released from presynaptic terminals at all and are often associated with "retrograde" signaling from postsynaptic cells back to presynaptic terminals.

- *Endocannabinoids* are a family of related endogenous signals that interact with cannabinoid receptors. These receptors are the molecular targets of Δ^9 -tetrahydrocannabinol, the psychoactive component of the marijuana plant, *Cannabis* (Box F). While some members of this emerging group of chemical signals remain to be determined, anandamide and 2-arachidonylglycerol (2-AG) have been established as endocannabinoids. These signals are unsaturated fatty acid with polar head groups and are produced by enzymatic degradation of membrane lipids (Figure 6.16A,B). Production of endocannabinoids is stimulated by a second messenger signal within postsynaptic neurons, typically a rise in postsynaptic Ca^{2+} concentration. Although the mechanism of endocannabinoid release is not entirely clear, it is likely that these hydrophobic signals diffuse through the postsynaptic membrane to reach cannabinoid receptors on other nearby cells. Endocannabinoid action is terminated by carrier-mediated transport of these signals back into the postsynaptic neuron. There they are hydrolyzed by the enzyme fatty acid hydrolase (FAAH).

At least two types of cannabinoid receptor have been identified, with most actions of endocannabinoids in the CNS mediated by the type termed CB1. CB1 is a G-protein-coupled receptor that is related to the metabotropic receptors for ACh, glutamate, and the other conventional neurotransmitters. Several compounds that are structurally related to endocannabinoids and that bind to the CB1 receptor have been synthesized (see Figure 6.16C). These compounds act as agonists or antagonists of the CB1 receptor and serve as both tools for elucidating the physiological functions of endocannabinoids and as targets for developing therapeutically useful drugs.

Endocannabinoids participate in several forms of synaptic regulation. The best-documented action of these agents is to inhibit communication between postsynaptic target cells and their presynaptic inputs. In both the hippocampus and the cerebellum, among other regions, endocannabinoids serve as retrograde signals to regulate GABA release at certain inhibitory terminals. At such synapses, depolarization of the postsynaptic neuron causes a transient reduction in inhibitory postsynaptic responses (Figure 6.17). Depolarization reduces synaptic transmission by elevating the concentration of Ca^{2+} within the postsynaptic neuron. This rise in Ca^{2+} triggers synthesis and release of endocannabinoids from the postsynaptic cells. The endocannabinoids then make their way to the presynaptic terminals and bind to CB1 receptors on these terminals. Activation of the CB1 receptors inhibits the amount of GABA released in response to presynaptic action potentials, thereby reducing inhibitory transmission. These mechanisms responsible for the reduction in GABA release are not entirely clear, but probably involve effects on voltage-gated Ca^{2+} channels and/or K^+ channels in the presynaptic neurons.

(A)



(C)

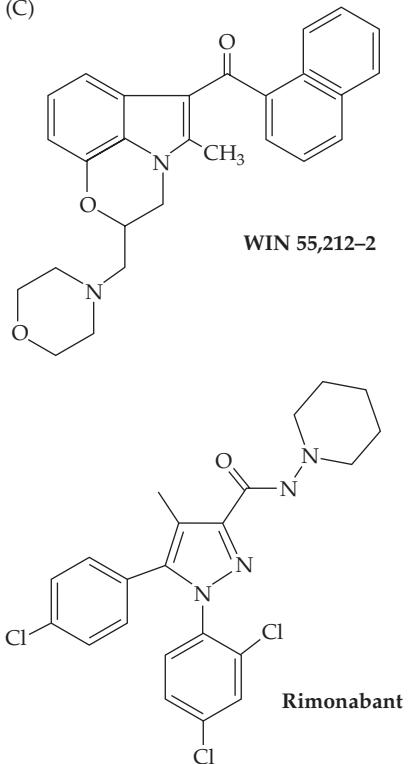
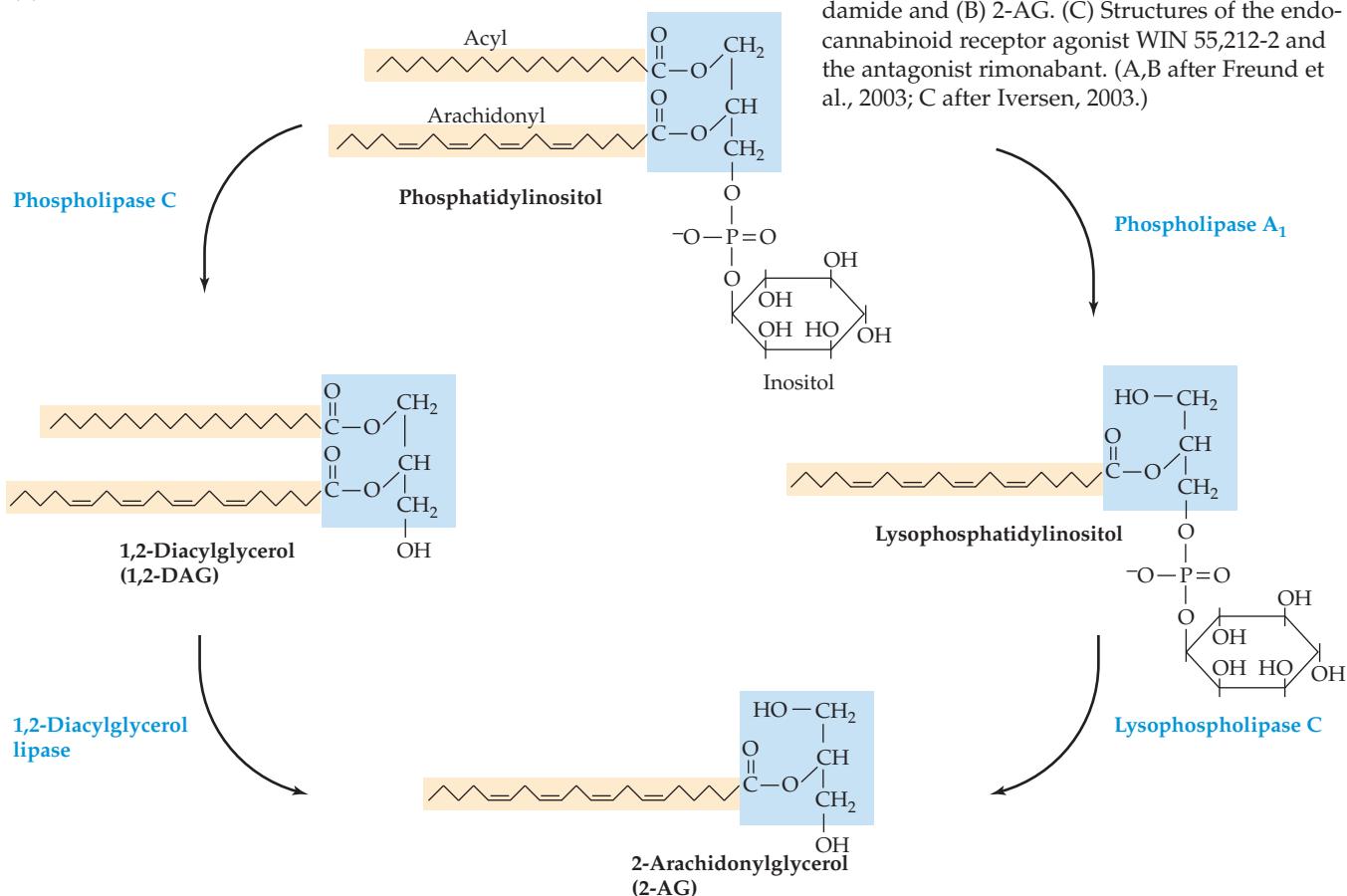


Figure 6.16 Endocannabinoid signals involved in synaptic transmission. Possible mechanism of production of the endocannabinoids (A) anandamide and (B) 2-AG. (C) Structures of the endocannabinoid receptor agonist WIN 55,212-2 and the antagonist rimonabant. (A,B after Freund et al., 2003; C after Iversen, 2003.)

(B)



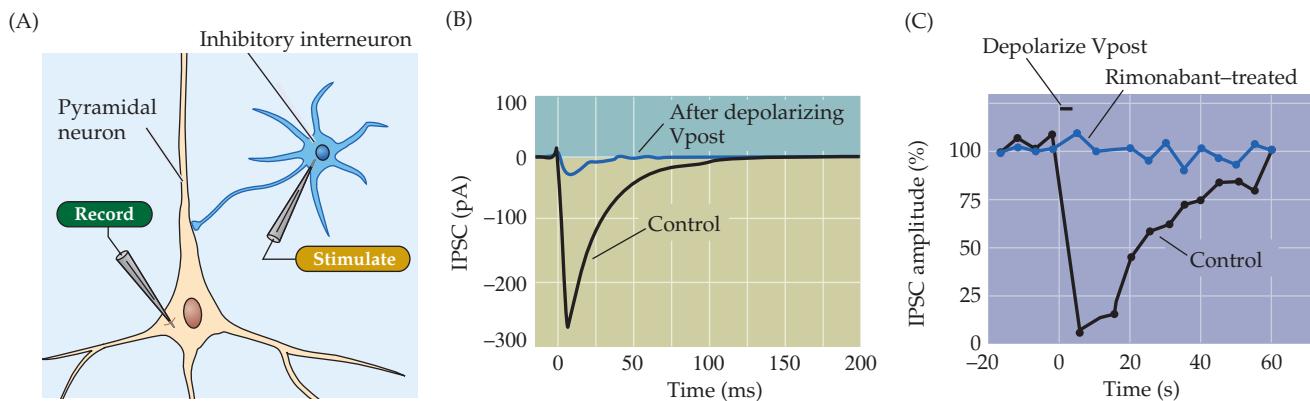
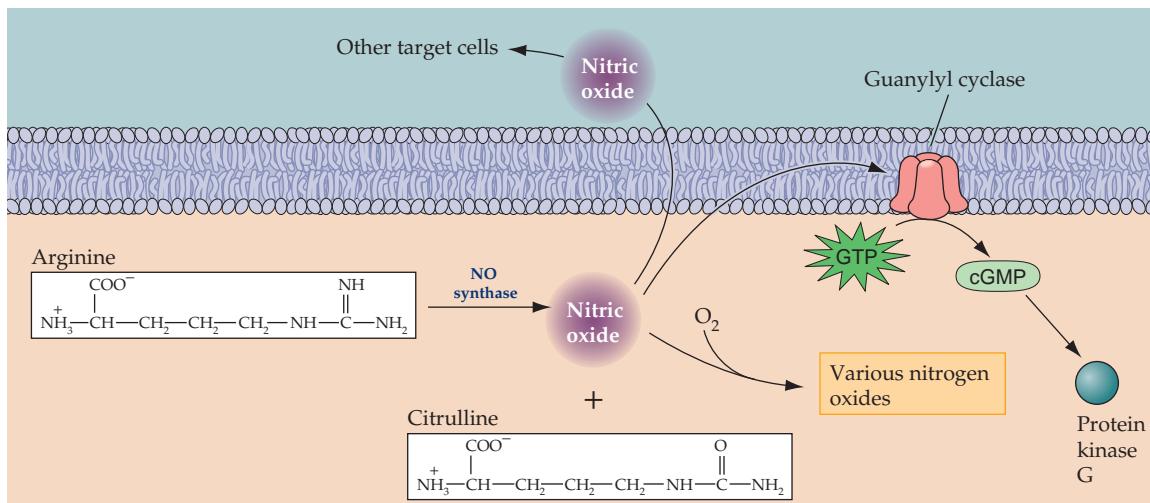


Figure 6.17 Endocannabinoid-mediated retrograde control of GABA release. (A) Experimental arrangement. Stimulation of a presynaptic interneuron causes release of GABA onto a postsynaptic pyramidal neuron. (B) IPSCs elicited by the inhibitory synapse (control) are reduced in amplitude following a brief depolarization of the postsynaptic neuron. This reduction in the IPSC is due to less GABA being released from the presynaptic interneuron. (C) The reduction in IPSC amplitude produced by postsynaptic depolarization lasts a few seconds and is mediated by endocannabinoids, because it is prevented by the endocannabinoid receptor antagonist rimonabant. (B,C after Ohno-Shosaku et al., 2001.)

• **Nitric oxide (NO)** is an unusual but especially interesting chemical signal. NO is a gas that is produced by the action of nitric oxide synthase, an enzyme that converts the amino acid arginine into a metabolite (citrulline) and simultaneously generates NO (Figure 6.18). NO is produced by an enzyme, nitric oxide synthase. Neuronal NO synthase is regulated by Ca^{2+} binding to the Ca^{2+} sensor protein calmodulin (see Chapter 7). Once produced, NO can permeate the plasma membrane, meaning that NO generated inside one cell can travel through the extracellular medium and act within nearby cells. Thus, this gaseous signal has a range of influence that extends well beyond the cell of origin, diffusing a few tens of micrometers from its site of production before it is degraded. This property makes NO a

Figure 6.18 Synthesis, release, and termination of NO.



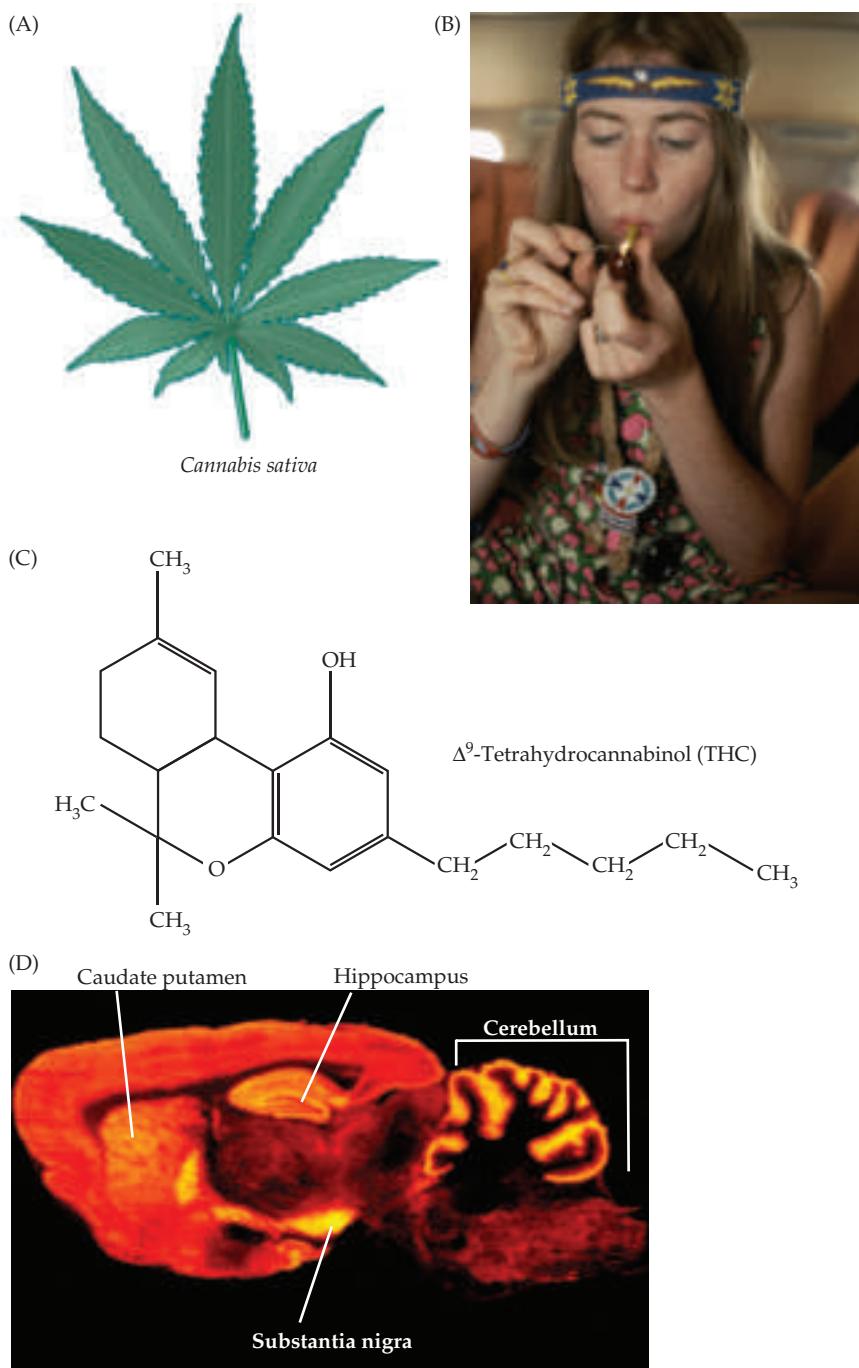
Box F

Marijuana and the Brain

Medicinal use of the marijuana plant, *Cannabis sativa* (Figure A), dates back thousands of years. Ancient civilizations—including both Greek and Roman societies in Europe, as well as Indian and Chinese cultures in Asia—appreciated that this plant was capable of producing relaxation, euphoria, and a number of other psychopharmacological actions. In more recent times, medicinal use of marijuana has largely subsided (although it remains useful in relieving the symptoms of terminal cancer patients); the recreational use of marijuana (Figure B) has become so popular that some societies have decriminalized its use.

Understanding the brain mechanisms underlying the actions of marijuana was advanced by the discovery that a cannabinoid, Δ^9 -tetrahydrocannabinol (THC; Figure C), is the active component of marijuana. This finding led to the development of synthetic derivatives, such as WIN 55,212-2 and rimonabant (see Figure 6.16), that have served as valuable tools for probing the brain actions of THC. Of particular interest is that receptors for these cannabinoids exist in the brain and exhibit marked regional variations in distribution, being especially enriched in the brain areas—such as substantia nigra and caudate putamen—that have been implicated in drug abuse (Figure D). The presence of these brain receptors for cannabinoids led in turn to a search for endogenous cannabinoid compounds in the brain, culminating in the discovery of endocannabinoids such as 2-AG and anandamide (see Figure 6.16). This path of discovery closely parallels the identification of endogenous opioid peptides, which resulted from the search for endogenous morphine-like compounds in the brain (see text and Table 6.2).

Given that THC interacts with brain endocannabinoid receptors, particularly



(A) Leaf of *Cannabis sativa*, the marijuana plant. (B) Smoking ground-up *Cannabis* leaves is a popular method of achieving the euphoric effects of marijuana. (C) Structure of THC (Δ^9 -tetrahydrocannabinol), the active ingredient of marijuana. (D) Distribution of brain CB1 receptors, visualized by examining the binding of CP-55,940, a CB1 receptor ligand. (B photo © Henry Diltz/Corbis; C after Iversen, 2003; D courtesy of M. Herkenham, NIMH.)

the CB1 receptor, it is likely that such actions are responsible for the behavioral consequences of marijuana use. Indeed, many of the well-documented effects of marijuana are consistent with the distribution and actions of brain CB1 receptors. For example, marijuana effects on perception could be due to CB1 receptors in the neocortex, effects on psychomotor control due to endocannabinoid receptors in the basal ganglia and cerebellum, effects on short-term memory due to cannabinoid

receptors in the hippocampus, and the well-known effects of marijuana on stimulating appetite due to hypothalamic actions. While formal links between these behavioral consequences of marijuana and the underlying brain mechanisms are still being forged, studies of the actions of this drug have shed substantial light on basic synaptic mechanisms, which promise to further elucidate the mode of action of one of the world's most popular drugs.

References

- ADAMS, A. R. (1941) Marihuana. *Harvey Lect.* 37: 168–168.
- FREUND, T. F., I. KATONA AND D. PIOMELLI (2003) Role of endogenous cannabinoids in synaptic signaling. *Physiol. Rev.* 83: 1017–1066.
- GERDEMAN, G. L., J. G. PARTRIDGE, C. R. LUPICA AND D. M. LOVINGER (2003) It could be habit forming: Drugs of abuse and striatal synaptic plasticity. *Trends Neurosci.* 26: 184–192.
- IVERSEN, L. (2003) Cannabis and the brain. *Brain* 126: 1252–1270.
- MECHOULAM, R. (1970) Marihuana chemistry. *Science* 168: 1159–1166.

potentially useful agent for coordinating the activities of multiple cells in a very localized region and may mediate certain forms of synaptic plasticity that spread within small networks of neurons.

All of the known actions of NO are mediated within its cellular targets; for this reason, NO often is considered a second messenger rather than a neurotransmitter. Some of these actions of NO are due to the activation of the enzyme guanylyl cyclase, which then produces the second messenger cGMP within target cells (see Chapter 7). Other actions of NO are the result of covalent modification of target proteins via nitrosylation, the addition of a nitryl group to selected amino acids within the proteins. NO decays spontaneously by reacting with oxygen to produce inactive nitrogen oxides. As a result, NO signals last for only a short time, on the order of seconds or less. NO signaling evidently regulates a variety of synapses that also employ conventional neurotransmitters; so far, presynaptic terminals that release glutamate are the best-studied target of NO in the central nervous system. NO may also be involved in some neurological diseases. For example, it has been proposed that an imbalance between nitric oxide and superoxide generation underlies some neurodegenerative diseases.

Summary

The complex synaptic computations occurring at neural circuits throughout the brain arise from the actions of a large number of neurotransmitters, which act on an even larger number of postsynaptic neurotransmitter receptors. Glutamate is the major excitatory neurotransmitter in the brain, whereas GABA and glycine are the major inhibitory neurotransmitters. The actions of these small-molecule neurotransmitters are typically faster than those of the neuropeptides. Thus, most small-molecule transmitters mediate synaptic transmission when a rapid response is essential, whereas the neuropeptide transmitters, as well as the biogenic amines and some small-molecule neurotransmitters, tend to modulate ongoing activity in the brain or in peripheral target tissues in a more gradual and ongoing way. Two broadly different families of neurotransmitter receptors have evolved to carry out the postsynaptic signaling actions of neurotransmitters. Ionotropic or ligand-

gated ion channels combine the neurotransmitter receptor and ion channel in one molecular entity, and therefore give rise to rapid postsynaptic electrical responses. Metabotropic receptors regulate the activity of postsynaptic ion channels indirectly, usually via G-proteins, and induce slower and longer-lasting electrical responses. Metabotropic receptors are especially important in regulating behavior, and drugs targeting these receptors have been clinically valuable in treating a wide range of behavioral disorders. The postsynaptic response at a given synapse is determined by the combination of receptor subtypes, G-protein subtypes, and ion channels that are expressed in the postsynaptic cell. Because each of these features can vary both within and among neurons, a tremendous diversity of transmitter-mediated effects is possible. Drugs that influence transmitter actions have enormous importance in the treatment of neurological and psychiatric disorders, as well as in a broad spectrum of other medical problems.

Additional Reading

Reviews

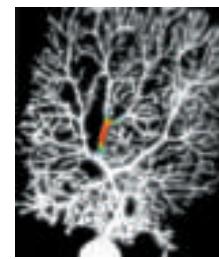
- BARNES, N. M. AND T. SHARP (1999) A review of central 5-HT receptors and their function. *Neuropharmac. 38:* 1083–1152.
- BOURIN, M., G. B. BAKER AND J. BRADWEIJN (1998) Neurobiology of panic disorder. *J. Psychosomatic Res. 44:* 163–180.
- BURNSTOCK, G. (1999) Current status of purinergic signalling in the nervous system. *Prog. Brain Res. 120:* 3–10.
- CARLSSON, A. (1987) Perspectives on the discovery of central monoaminergic neurotransmission. *Annu. Rev. Neurosci. 10:* 19–40.
- CHANGEUX, J.-P. (1993) Chemical signaling in the brain. *Sci. Amer. 269* (May): 58–62.
- CIVELLI, O. (1998) Functional genomics: The search for novel neurotransmitters and neuropeptides. *FEBS Letters 430:* 55–58.
- FREDHOLM, B. B. (1995) Adenosine, adenosine receptors and the actions of caffeine. *Pharmacol. Toxicol. 76:* 93–101.
- FREUND, T. F., I. KATONA AND D. PIOMELLI (2003) Role of endogenous cannabinoids in synaptic signaling. *Physiol Rev. 83:* 1017–1066.
- HÖKFELT, T. D. AND 10 OTHERS (1987) Coexistence of peptides with classical neurotransmitters. *Experientia Suppl. 56:* 154–179.
- HYLAND, K. (1999) Neurochemistry and defects of biogenic amine neurotransmitter metabolism. *J. Inher. Metab. Dis. 22:* 353–363.
- INESTROSA, N. C. AND A. PERELMAN (1989) Distribution and anchoring of molecular forms of acetylcholinesterase. *Trends Pharmacol. Sci. 10:* 325–329.
- IVERSEN, L. (2003) Cannabis and the brain. *Brain 126:* 1252–1270.
- KOOB, G. F., P. P. SANNA AND F. E. BLOOM (1998) Neuroscience of addiction. *Neuron 21:* 467–476.
- KUPFERMANN, I. (1991) Functional studies of cotransmission. *Physiol. Rev. 71:* 683–732.
- LAUBE, B., G. MAKSY, R. SCHEMM AND H. BETZ (2002) Modulation of glycine receptor function: A novel approach for therapeutic intervention at inhibitory synapses? *Trends Pharmacol. Sci. 23:* 519–527.
- LOVINGER, D. M. (1999) 5-HT₃ receptors and the neural actions of alcohols: An increasingly exciting topic. *Neurochem. Internat. 35:* 125–30.
- MACKENZIE, A. B., A. SURPRENANT AND R. A. NORTH (1999) Functional and molecular diversity of purinergic ion channel receptors. *Ann. NY Acad. Sci. 868:* 716–729.
- MASSON, J., C. SAGN, M. HAMON AND S. E. MESTIKAWY (1999) Neurotransmitter transporters in the central nervous system. *Pharmacol. Rev. 51:* 439–464.
- MELDRUM, B. AND J. GARTHWAITE (1990) Glutamate neurotoxicity may underlie slowly progressive degenerative diseases such as Huntington's Disease and Alzheimer's Disease. *Trends Pharmacol. Sci. 11:* 379–387.
- NAKANISHI, S. (1992) Molecular diversity of glutamate receptors and implication for brain function. *Science 258:* 597–603.
- PERRY, E., M. WALKER, J. GRACE AND R. PERRY (1999) Acetylcholine in mind: A neurotransmitter correlate of consciousness? *Trends Neurosci. 22:* 273–280.
- PIERCE, K. L., R. T. PREMONT AND R. J. LEFKOWITZ (2002) Seven-transmembrane receptors. *Nature Rev. Mol. Cell Biol. 3:* 639–650.
- SCHWARTZ, J. C., J. M. ARRANG, M. GARBARG, H. POLLARD AND M. RUAT (1991) Histaminergic transmission in the mammalian brain. *Physiol. Rev. 71:* 1–51.
- SCHWARTZ, M. W., S. C. WOODS, D. PORTE JR., R. J. SEELEY AND D. G. BASKIN (2000) Central nervous system control of food intake. *Nature 404:* 661–671.
- STAMLER, J. S., E. J. TOONE, S. A. LIPTON AND N. SUCHER (1997) (S)NO Signals: Translocation, regulation, and a consensus motif. *Neuron 18:* 691–696.
- TUCEK, S., J. RICNY AND V. DOLEZAL (1990) Advances in the biology of cholinergic neurons. *Adv. Neurol. 51:* 109–115.
- WEBB, T. E. AND E. A. BARNARD (1999) Molecular biology of P2Y receptors expressed in the nervous system. *Prog. Brain Res. 120:* 23–31.
- WILSON, R. I. AND R. A. NICOLL (2002) Endocannabinoid signaling in the brain. *Science 296:* 678–682.

Important Original Papers

- BRENowitz, S. D. AND W. G. REGEHR (2003) Calcium dependence of retrograde inhibition by endocannabinoids at synapses onto Purkinji cells. *J. Neurosci. 23:* 6373–6384.
- CHAVAS, J. AND A. MARTY (2003) Coexistence of excitatory and inhibitory GABA synapses in the cerebellar interneuron network. *J. Neurosci. 23:* 2019–2031.
- CHEN, Z. P., A. LEVY AND S. L. LIGHTMAN (1995) Nucleotides as extracellular signalling molecules. *J. Neuroendocrinol. 7:* 83–96.
- CURTIS, D. R., J. W. PHILLIS AND J. C. WATKINS (1959) Chemical excitation of spinal neurons. *Nature 183:* 611–612.
- DALE, H. H., W. FELDBERG AND M. VOGL (1936) Release of acetylcholine at voluntary motor nerve endings. *J. Physiol. 86:* 353–380.
- DAVIES, P. A. AND 6 OTHERS (1999) The 5-HT_{3B} subunit is a major determinant of serotonin-receptor function. *Nature 397:* 359–363.
- GOMEZA, J. AND 6 OTHERS (2003) Inactivation of the glycine transporter 1 gene discloses vital role of glial glycine uptake in glycinergic inhibition. *Neuron 40:* 785–796.
- GU, J. G. AND A. B. MACDERMOTT (1997) Activation of ATP P2X receptors elicits glutamate release from sensory neuron synapses. *Nature 389:* 749–753.
- HÖKFELT, T., O. JOHANSSON, A. LJUNGDAHL, J.

- M. LUNDBERG AND M. SCHULTZBERG (1980) Peptidergic neurons. *Nature* 284: 515–521.
- HOLLMANN, M., C. MARON AND S. HEINEMANN (1994) N-glycosylation site tagging suggests a three transmembrane domain topology for the glutamate receptor GluR1. *Neuron* 13: 1331–1343.
- HUGHES, J., T. W. SMITH, H. W. KOSTERLITZ, L. A. FOTHERGILL, B. A. MORGAN AND H. R. MORRIS (1975) Identification of two related pentapeptides from the brain with potent opiate agonist activity. *Nature* 258: 577–580.
- KAUPMANN, K. AND 10 OTHERS (1997) Expression cloning of GABA β receptors uncovers similarity to metabotropic glutamate receptors. *Nature* 386: 239–246.
- KREITZER, A. C. AND W. G. REGEHR (2001) Retrograde inhibition of presynaptic calcium influx by endogenous cannabinoids at excitatory synapses onto Purkinje cells. *Neuron* 29: 717–727.
- LEDEBT, C. AND 9 OTHERS (1997) Aggressiveness, hypoalgesia and high blood pressure in mice lacking the adenosine A2a receptor. *Nature* 388: 674–678.
- NAVEILHAN, P. AND 10 OTHERS (1999) Normal feeding behavior, body weight and leptin response require the neuropeptide Y Y2 receptor. *Nature Med.* 5: 1188–1193.
- OHNO-SHOSAKU, T., T. MAEJIMA, AND M. KANO (2001) Endogenous cannabinoids mediate retrograde signals from depolarized postsynaptic neurons to presynaptic terminals. *Neuron* 29: 729–738.
- ROSENKRANTZ, C., Y. STERN-BACH AND C. F. STEVENS (1998) The tetrameric structure of a glutamate receptor channel. *Science* 280: 1596–1599.
- THOMAS, S. A. AND R. D. PALMITER (1995) Targeted disruption of the tyrosine hydroxylase gene reveals that catecholamines are required for mouse fetal development. *Nature* 374: 640–643.
- UNWIN, N. (1995) Acetylcholine receptor channels imaged in the open state. *Nature* 373: 37–43.
- WANG, Y.M. AND 8 OTHERS (1997) Knockout of the vesicular monoamine transporter 2 gene results in neonatal death and supersensitivity to cocaine and amphetamine. *Neuron* 19: 1285–1296.
- Books**
- BRADFORD, H. F. (1986) *Chemical Neurobiology*. New York: W. H. Freeman.
- COOPER, J. R., F. E. BLOOM AND R. H. ROTH (2003) *The Biochemical Basis of Neuropharmacology*. New York: Oxford University Press.
- FELDMAN, R. S., J. S. MEYER AND L. F. QUENZER (1997) *Principles of Neuropharmacology*, 2nd Edition. Sunderland, MA: Sinauer Associates.
- HALL, Z. (1992) *An Introduction to Molecular Neurobiology*. Sunderland, MA: Sinauer Associates.
- HILLE, B. (2002) *Ion Channels of Excitable Membranes*, 3rd Edition. Sunderland, MA: Sinauer Associates.
- MYCEK, M. J., R. A. HARVEY AND P. C. CHAMPE (2000) *Pharmacology*, 2nd Edition. Philadelphia, New York: Lippincott/Williams and Wilkins Publishers.
- NICHOLLS, D. G. (1994) *Proteins, Transmitters, and Synapses*. Boston: Blackwell Scientific.
- SIEGEL, G.J., B. W. AGRANOFF, R. W. ALBERS, S. K. FISHER AND M. D. UHLER (1999) *Basic Neurochemistry*. Philadelphia: Lippincott-Raven.

Chapter 7



Molecular Signaling within Neurons

Overview

As is apparent in the preceding chapters, electrical and chemical signaling mechanisms allow one nerve cell to receive and transmit information to another. This chapter focuses on the related events within neurons and other cells that are triggered by the interaction of a chemical signal with its receptor. This intracellular processing typically begins when extracellular chemical signals, such as neurotransmitters, hormones, and trophic factors, bind to specific receptors located either on the surface or within the cytoplasm or nucleus of the target cells. Such binding activates the receptors and in so doing stimulates cascades of intracellular reactions involving GTP-binding proteins, second messenger molecules, protein kinases, ion channels, and many other effector proteins whose modulation temporarily changes the physiological state of the target cell. These same intracellular signal transduction pathways can also cause longer-lasting changes by altering the transcription of genes, thus affecting the protein composition of the target cells on a more permanent basis. The large number of components involved in intracellular signaling pathways allows precise temporal and spatial control over the function of individual neurons, thereby allowing the coordination of electrical and chemical activity in the related populations of neurons that comprise neural circuits and systems.

Strategies of Molecular Signaling

Chemical communication coordinates the behavior of individual nerve and glial cells in physiological processes that range from neural differentiation to learning and memory. Indeed, molecular signaling ultimately mediates and modulates all brain functions. To carry out such communication, a series of extraordinarily diverse and complex chemical signaling pathways has evolved. The preceding chapters have described in some detail the electrical signaling mechanisms that allow neurons to generate action potentials for conduction of information. These chapters also described synaptic transmission, a special form of chemical signaling that transfers information from one neuron to another. Chemical signaling is not, however, limited to synapses (Figure 7.1A). Other well-characterized forms of chemical communication include **paracrine** signaling, which acts over a longer range than synaptic transmission and involves the secretion of chemical signals onto a group of nearby target cells, and **endocrine** signaling, which refers to the secretion of hormones into the bloodstream where they can affect targets throughout the body.

Chemical signaling of any sort requires three components: a molecular *signal* that transmits information from one cell to another, a *receptor* molecule

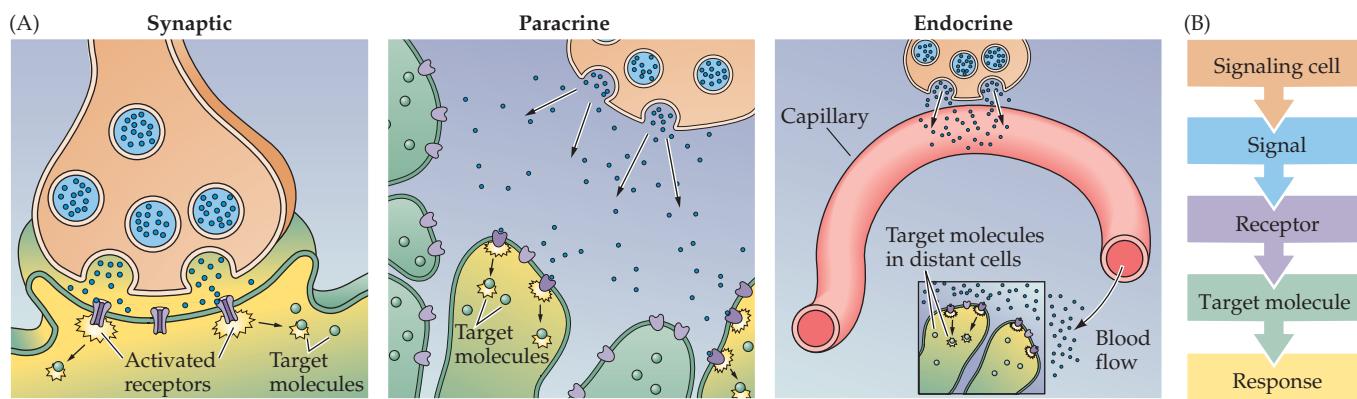
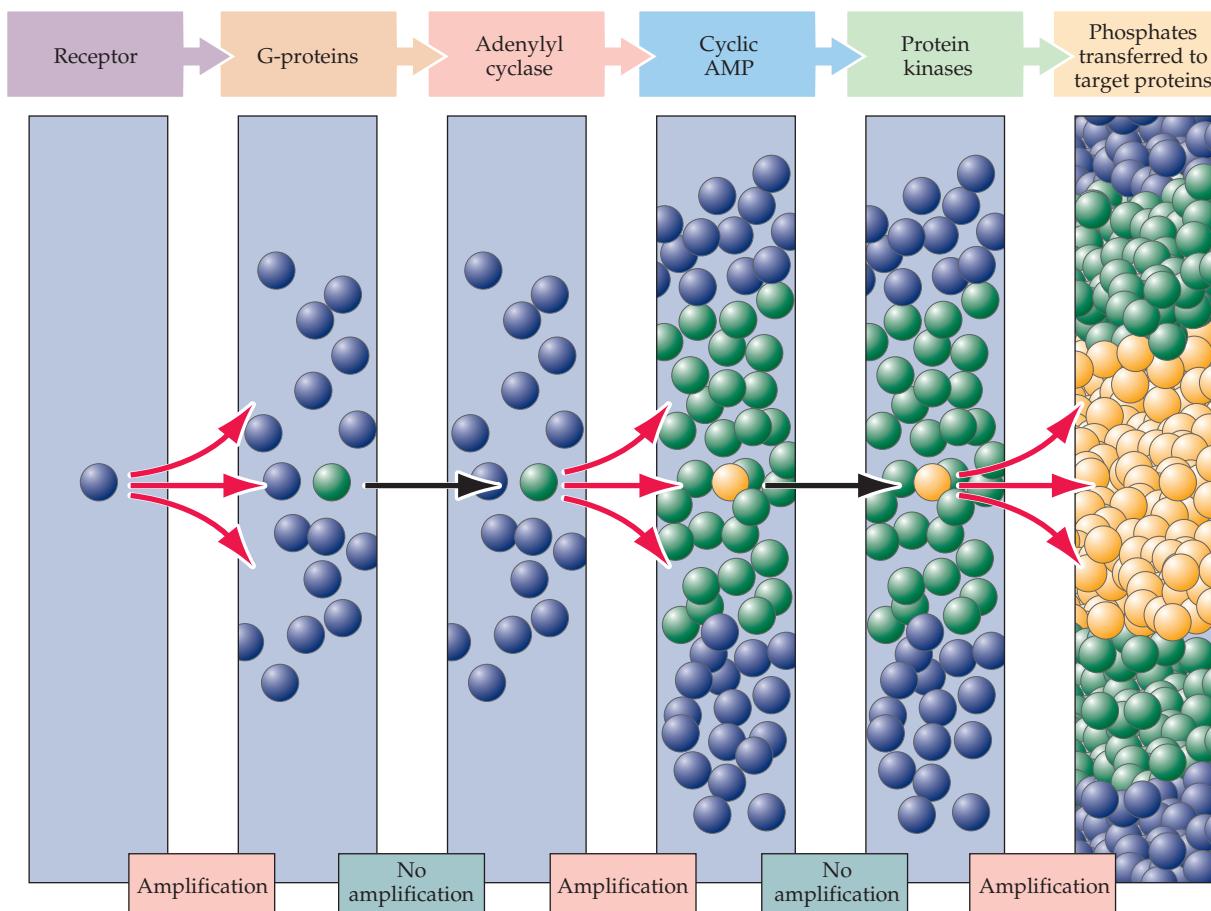


Figure 7.1 Chemical signaling mechanisms. (A) Forms of chemical communication include synaptic transmission, paracrine signaling, and endocrine signaling. (B) The essential components of chemical signaling are: cells that initiate the process by releasing signaling molecules; specific receptors on target cells; second messenger target molecules; and subsequent cellular responses.

that transduces the information provided by the signal, and a *target molecule* that mediates the cellular response (Figure 7.1B). The part of this process that take place within the confines of the target cell is called **intracellular signal transduction**. A good example of transduction in the context of *intercellular* communication is the sequence of events triggered by chemical synaptic transmission (see Chapter 5): Neurotransmitters serve as the signal, neurotransmitter receptors serve as the transducing receptor, and the target molecule is an ion channel that is altered to cause the electrical response of the postsynaptic cell. In many cases, however, synaptic transmission activates additional *intracellular* pathways that have a variety of functional consequences. For example, the binding of the neurotransmitter norepinephrine to its receptor activates GTP-binding proteins, which produces second messengers within the postsynaptic target, activates enzyme cascades, and eventually changes the chemical properties of numerous target molecules within the affected cell.

A general advantage of chemical signaling in both intercellular and intracellular contexts is **signal amplification**. Amplification occurs because individual signaling reactions can generate a much larger number of molecular products than the number of molecules that initiate the reaction. In the case of norepinephrine signaling, for example, a single norepinephrine molecule binding to its receptor can generate many thousands of second messenger molecules (such as cyclic AMP), yielding an amplification of tens of thousands of phosphates transferred to target proteins (Figure 7.2). Similar amplification occurs in all signal transduction pathways. Because the transduction processes often are mediated by a sequential set of enzymatic reactions, each with its own amplification factor, a small number of signal molecules ultimately can activate a very large number of target molecules. Such amplification guarantees that a physiological response is evoked in the face of other, potentially countervailing, influences.

Another rationale for these complex signal transduction schemes is to permit precise control of cell behavior over a wide range of times. Some molecular interactions allow information to be transferred rapidly, while others are slower and longer lasting. For example, the signaling cascades associated with synaptic transmission at neuromuscular junctions allow a person to respond to rapidly changing cues, such as the trajectory of a pitched ball, while the slower responses triggered by adrenal medullary hormones (epinephrine and norepinephrine) secreted during a challenging game produce slower (and longer lasting) effects on muscle metabolism (see Chapter 20) and emotional state (see Chapter 29). To encode information that varies so



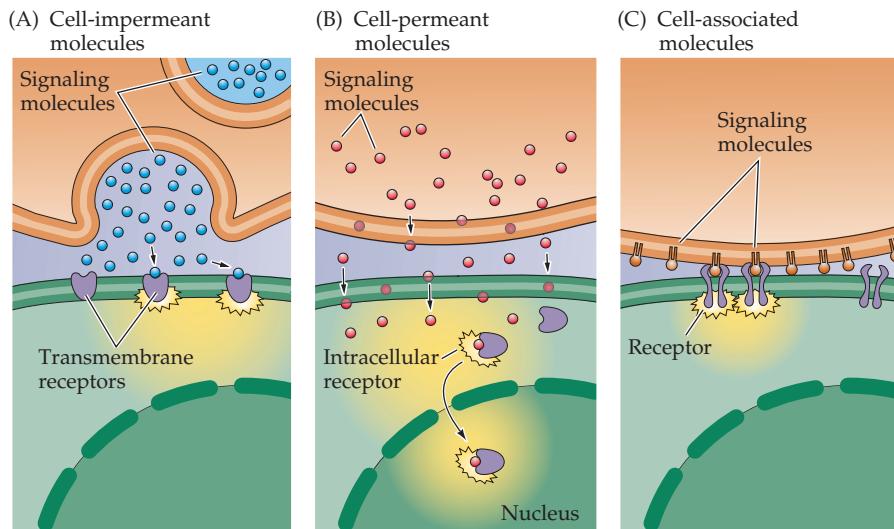
widely over time, the concentration of the relevant signaling molecules must be carefully controlled. On one hand, the concentration of every signaling molecule within the signaling cascade must return to subthreshold values before the arrival of another stimulus. On the other hand, keeping the intermediates in a signaling pathway activated is critical for a sustained response. Having multiple levels of molecular interactions facilitates the intricate timing of these events.

The Activation of Signaling Pathways

The molecular components of these signal transduction pathways are always activated by a chemical signaling molecule. Such signaling molecules can be grouped into three classes: **cell-impermeant**, **cell-permeant**, and **cell-associated signaling molecules** (Figure 7.3). The first two classes are secreted molecules and thus can act on target cells removed from the site of signal synthesis or release. Cell-impermeant signaling molecules typically bind to receptors associated with cell membranes. Hundreds of secreted molecules have now been identified, including the neurotransmitters discussed in Chapter 6, as well as proteins such as neurotrophic factors (see Chapter 22), and peptide hormones such as glucagon, insulin, and various reproductive hormones. These signaling molecules are typically short-lived, either because they are rapidly metabolized or because they are internalized by endocytosis once bound to their receptors.

Figure 7.2 Amplification in signal transduction pathways. The activation of a single receptor by a signaling molecule, such as the neurotransmitter norepinephrine, can lead to the activation of numerous G-proteins inside cells. These activated proteins can bind to other signaling molecules, such as the enzyme adenylyl cyclase. Each activated enzyme molecule generates a large number of cAMP molecules. cAMP binds to and activates another family of enzymes, protein kinases. These enzymes can then phosphorylate many target proteins. While not every step in this signaling pathway involves amplification, overall the cascade results in a tremendous increase in the potency of the initial signal.

Figure 7.3 Three classes of cell signaling molecules. (A) Cell-impermeant molecules, such as neurotransmitters, cannot readily traverse the plasma membrane of the target cell and must bind to the extracellular portion of transmembrane receptor proteins. (B) Cell-permeant molecules are able to cross the plasma membrane and bind to receptors in the cytoplasm or nucleus of target cells. (C) Cell-associated molecules are presented on the extracellular surface of the plasma membrane. These signals activate receptors on target cells only if they are directly adjacent to the signaling cell.



Cell-permeant signaling molecules can cross the plasma membrane to act directly on receptors that are inside the cell. Examples include numerous steroid (glucocorticoids, estradiol, and testosterone) and thyroid (thyroxin) hormones, and retinoids. These signaling molecules are relatively insoluble in aqueous solutions and are often transported in blood and other extracellular fluids by binding to specific carrier proteins. In this form, they may persist in the bloodstream for hours or even days.

The third group of chemical signaling molecules, cell-associated signaling molecules, are arrayed on the extracellular surface of the plasma membrane. As a result, these molecules act only on other cells that are physically in contact with the cell that carries such signals. Examples include proteins such as the integrins and neural cell adhesion molecules (NCAMs) that influence axonal growth (see Chapter 22). Membrane-bound signaling molecules are more difficult to study, but are clearly important in neuronal development and other circumstances where physical contact between cells provides information about cellular identities.

Receptor Types

Regardless of the nature of the initiating signal, cellular responses are determined by the presence of receptors that specifically bind the signaling molecules. Binding of signal molecules causes a conformational change in the receptor, which then triggers the subsequent signaling cascade within the affected cell. Given that chemical signals can act either at the plasma membrane or within the cytoplasm (or nucleus) of the target cell, it is not surprising that receptors are actually found on both sides of the plasma membrane. The receptors for impermeant signal molecules are membrane-spanning proteins. The extracellular domain of such receptors includes the binding site for the signal, while the intracellular domain activates intracellular signaling cascades after the signal binds. A large number of these receptors have been identified and are grouped into families defined by the mechanism used to transduce signal binding into a cellular response (Figure 7.4).

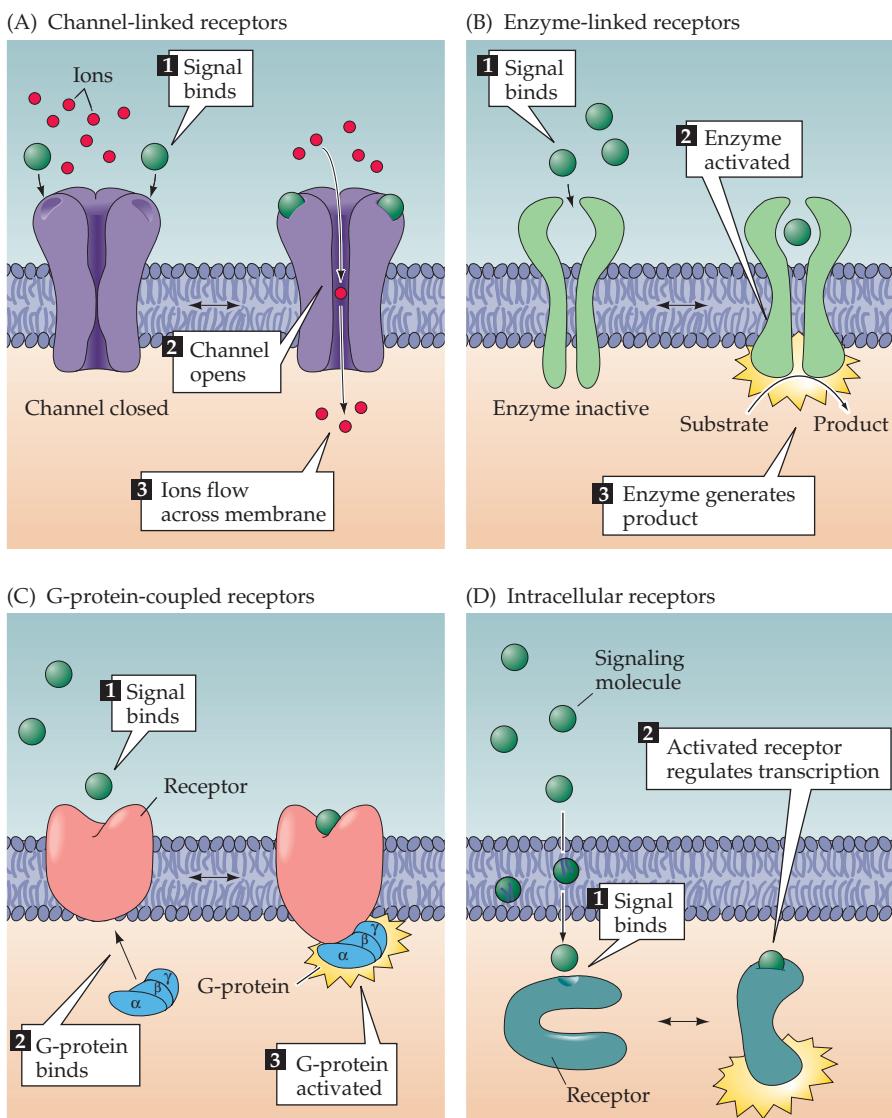


Figure 7.4 Categories of cellular receptors. Membrane-impermeant signaling molecules can bind to and activate either channel-linked receptors (A), enzyme-linked receptors (B), or G-protein-coupled receptors (C). Membrane permeant signaling molecules activate intracellular receptors (D).

Channel-linked receptors (also called ligand-gated ion channels) have the receptor and transducing functions as part of the same protein molecule. Interaction of the chemical signal with the binding site of the receptor causes the opening or closing of an ion channel pore in another part of the same molecule. The resulting ion flux changes the membrane potential of the target cell and, in some cases, can also lead to entry of Ca^{2+} ions that serve as a second messenger signal within the cell. Good examples of such receptors are the ionotropic neurotransmitter receptors described in Chapters 5 and 6.

Enzyme-linked receptors also have an extracellular binding site for chemical signals. The intracellular domain of such receptors is an enzyme whose catalytic activity is regulated by the binding of an extracellular signal. The great majority of these receptors are **protein kinases**, often tyrosine kinases, that phosphorylate intracellular target proteins, thereby changing the physiological function of the target cells. Noteworthy members of this

group of receptors are the Trk family of neurotrophin receptors (see Chapter 22) and other receptors for growth factors.

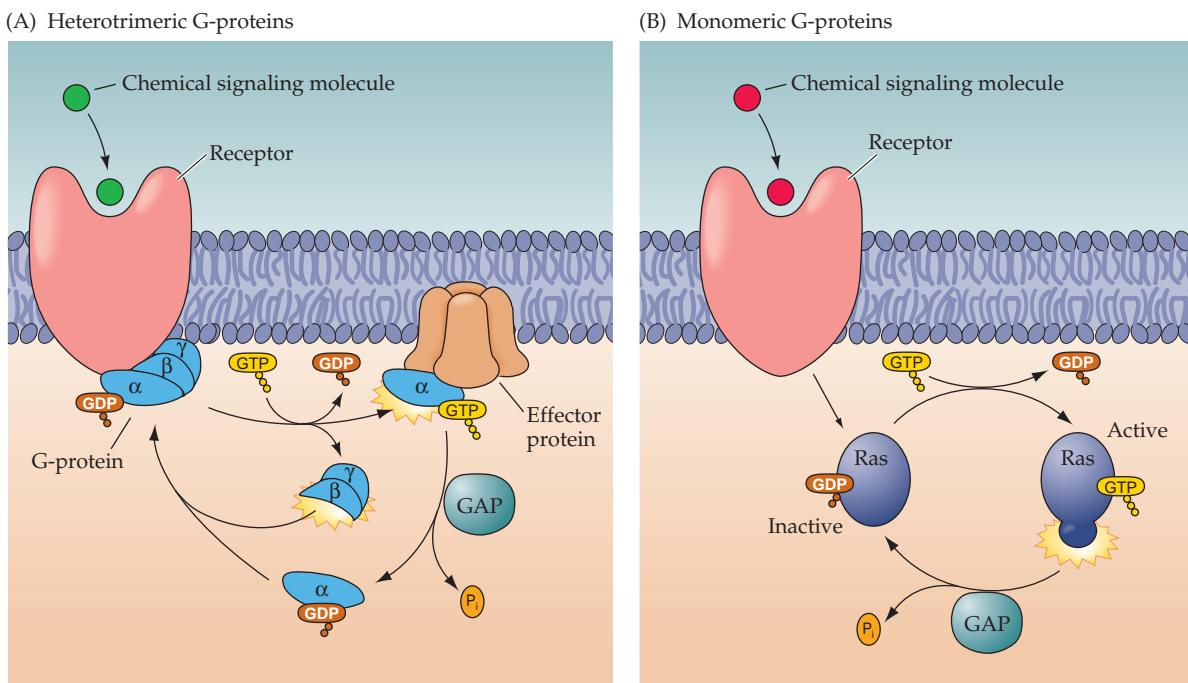
G-protein-coupled receptors regulate intracellular reactions by an indirect mechanism involving an intermediate transducing molecule, called the **GTP-binding proteins** (or **G-proteins**). Because these receptors all share the structural feature of crossing the plasma membrane seven times, they are also referred to as 7-transmembrane receptors (or metabotropic receptors; see Chapter 5). Hundreds of different G-protein-linked receptors have been identified. Well-known examples include the β -adrenergic receptor, the muscarinic type of acetylcholine receptor, metabotropic glutamate receptors, receptors for odorants in the olfactory system, and many types of receptors for peptide hormones. Rhodopsin, a light-sensitive, 7-transmembrane protein in retinal photoreceptors, is another form of G-protein-linked receptor (see Chapter 10).

Intracellular receptors are activated by cell-permeant or lipophilic signaling molecules (Figure 7.4D). Many of these receptors lead to the activation of signaling cascades that produce new mRNA and protein within the target cell. Often such receptors comprise a receptor protein bound to an inhibitory protein complex. When the signaling molecule binds to the receptor, the inhibitory complex dissociates to expose a DNA-binding domain on the receptor. This activated form of the receptor can then move into the nucleus and directly interact with nuclear DNA, resulting in altered transcription. Some intracellular receptors are located primarily in the cytoplasm, while others are in the nucleus. In either case, once these receptors are activated they can affect gene expression by altering DNA transcription.

G-Proteins and Their Molecular Targets

Both G-protein-linked receptors and enzyme-linked receptors can activate biochemical reaction cascades that ultimately modify the function of target proteins. For both these receptor types, the coupling between receptor activation and their subsequent effects are the GTP-binding proteins. There are two general classes of GTP-binding protein (Figure 7.5). **Heterotrimeric G-proteins** are composed of three distinct subunits (α , β , and γ). There are many different α , β , and γ subunits, allowing a bewildering number of G-protein permutations. Regardless of the specific composition of the heterotrimeric G-protein, its α subunit binds to guanine nucleotides, either GTP or GDP. Binding of GDP then allows the α subunit to bind to the β and γ subunits to form an inactive trimer. Binding of an extracellular signal to a G-protein-coupled receptor in turn allows the G-protein to bind to the receptor and causes GDP to be replaced with GTP (Figure 7.5A). When GTP is bound to the G-protein, the α subunit dissociates from the $\beta\gamma$ complex and activates the G-protein. Following activation, both the GTP-bound α subunit and the free $\beta\gamma$ complex can bind to downstream effector molecules that mediate a variety of responses in the target cell.

The second class of GTP-binding proteins are **monomeric G-proteins** (also called **small G-proteins**). These monomeric GTPases also relay signals from activated cell surface receptors to intracellular targets such as the cytoskeleton and the vesicle trafficking apparatus of the cell. The first small G-protein was discovered in a virus that causes *rat* sarcoma tumors and was therefore called **ras**. Ras is a molecule that helps regulate cell differentiation and proliferation by relaying signals from receptor kinases to the nucleus; the viral form of ras is defective, which accounts for the ability of the virus to cause the uncontrolled cell proliferation that leads to tumors. Since then, a



large number of small GTPases have been identified and can be sorted into five different subfamilies with different functions. For instance, some are involved in vesicle trafficking in the presynaptic terminal or elsewhere in the neuron, while others play a central role in protein and RNA trafficking in and out of the nucleus.

Termination of signaling by both heterotrimeric and monomeric G-proteins is determined by hydrolysis of GTP to GDP. The rate of GTP hydrolysis is an important property of a particular G-protein that can be regulated by other proteins, termed GTPase-activating proteins (GAPs). By replacing GTP with GDP, GAPs return G-proteins to their inactive form. GAPs were first recognized as regulators of small G-proteins, but recently similar proteins have been found to regulate the α subunits of heterotrimeric G-proteins. Hence, monomeric and trimeric G-proteins function as molecular timers that are active in their GTP-bound state, and become inactive when they have hydrolyzed the bound GTP to GDP (Figure 7.5B).

Activated G-proteins alter the function of many downstream effectors. Most of these effectors are enzymes that produce intracellular second messengers. Effector enzymes include adenylyl cyclase, guanylyl cyclase, phospholipase C, and others (Figure 7.6). The second messengers produced by these enzymes trigger the complex biochemical signaling cascades discussed in the next section. Because each of these cascades is activated by specific G-protein subunits, the pathways activated by a particular receptor are determined by the specific identity of the G-protein subunits associated with it.

As well as activating effector molecules, G-proteins can also directly bind to and activate ion channels. For example, some neurons, as well as heart muscle cells, have G-protein-coupled receptors that bind acetylcholine. Because these receptors are also activated by the agonist muscarine, they are usually called muscarinic receptors (see Chapters 6 and 20). Activation of muscarinic receptors can open K^+ channels, thereby inhibiting the rate at which the neuron fires action potentials, or slowing the heartbeat of muscle

Figure 7.5 Types of GTP-binding protein. (A) Heterotrimeric G-proteins are composed of three distinct subunits (α , β , and γ). Receptor activation causes the binding of the G-protein and the α subunit to exchange GDP for GTP, leading to a dissociation of the α and $\beta\gamma$ subunits. The biological actions of these G-proteins are terminated by hydrolysis of GTP, which is enhanced by GTPase-activating (GAP) proteins. (B) Monomeric G-proteins use similar mechanisms to relay signals from activated cell surface receptors to intracellular targets. Binding of GTP stimulates the biological actions of these G-proteins, and their activity is terminated by hydrolysis of GTP, which is also regulated by GAP proteins.

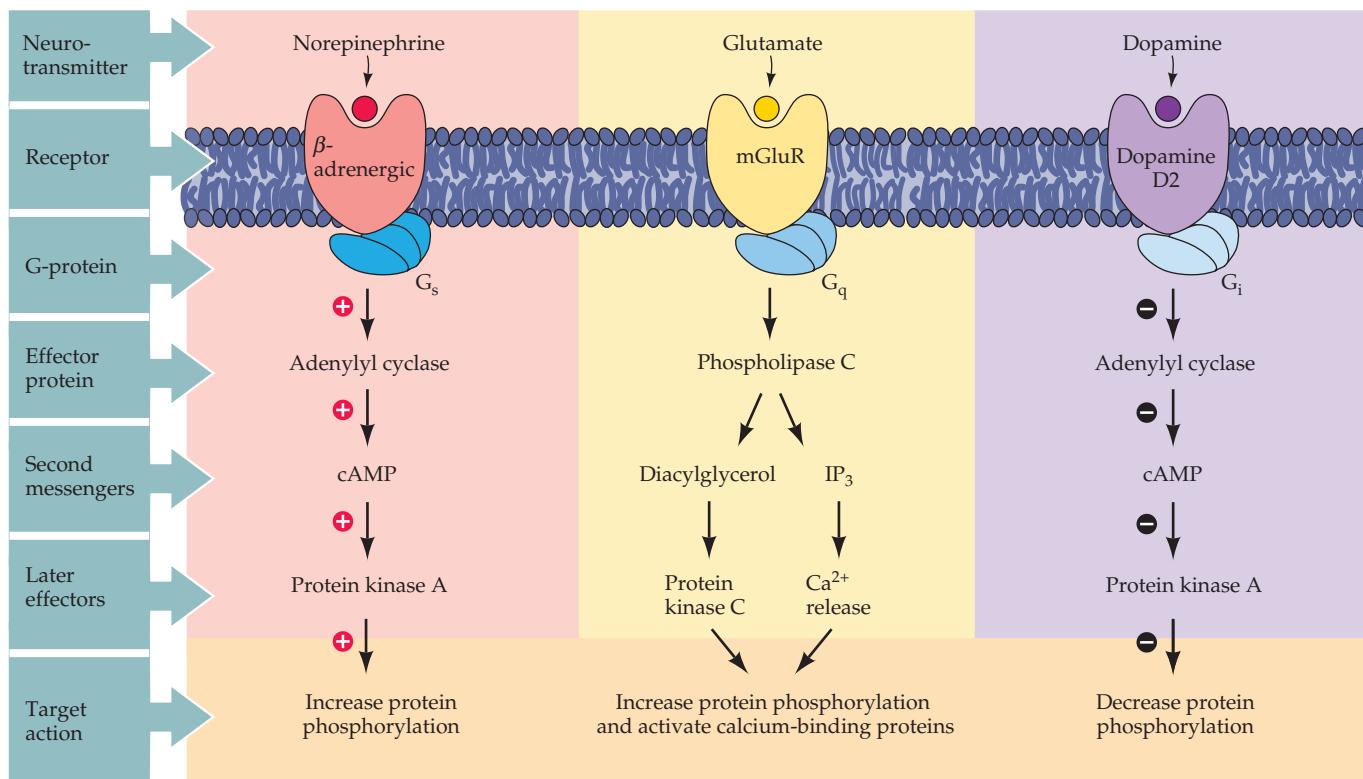


Figure 7.6 Effector pathways associated with G-protein-coupled receptors. In all three examples shown here, binding of a neurotransmitter to such a receptor leads to activation of a G-protein and subsequent recruitment of second messenger pathways. G_s, G_q, and G_i refer to three different types of heterotrimeric G-protein.

cells. These inhibitory responses are believed to be the result of $\beta\gamma$ subunits of G-proteins binding to the K⁺ channels. The activation of α subunits can also lead to the rapid closing of voltage-gated Ca²⁺ and Na⁺ channels. Because these channels carry inward currents involved in generating action potentials, closing them makes it more difficult for target cells to fire (see Chapters 3 and 4).

In summary, the binding of chemical signals to their receptors activates cascades of signal transduction events in the cytosol of target cells. Within such cascades, G-proteins serve a pivotal function as the molecular transducing elements that couple membrane receptors to their molecular effectors within the cell. The diversity of G-proteins and their downstream targets leads to many types of physiological responses. By directly regulating the gating of ion channels, G-proteins can influence the membrane potential of target cells.

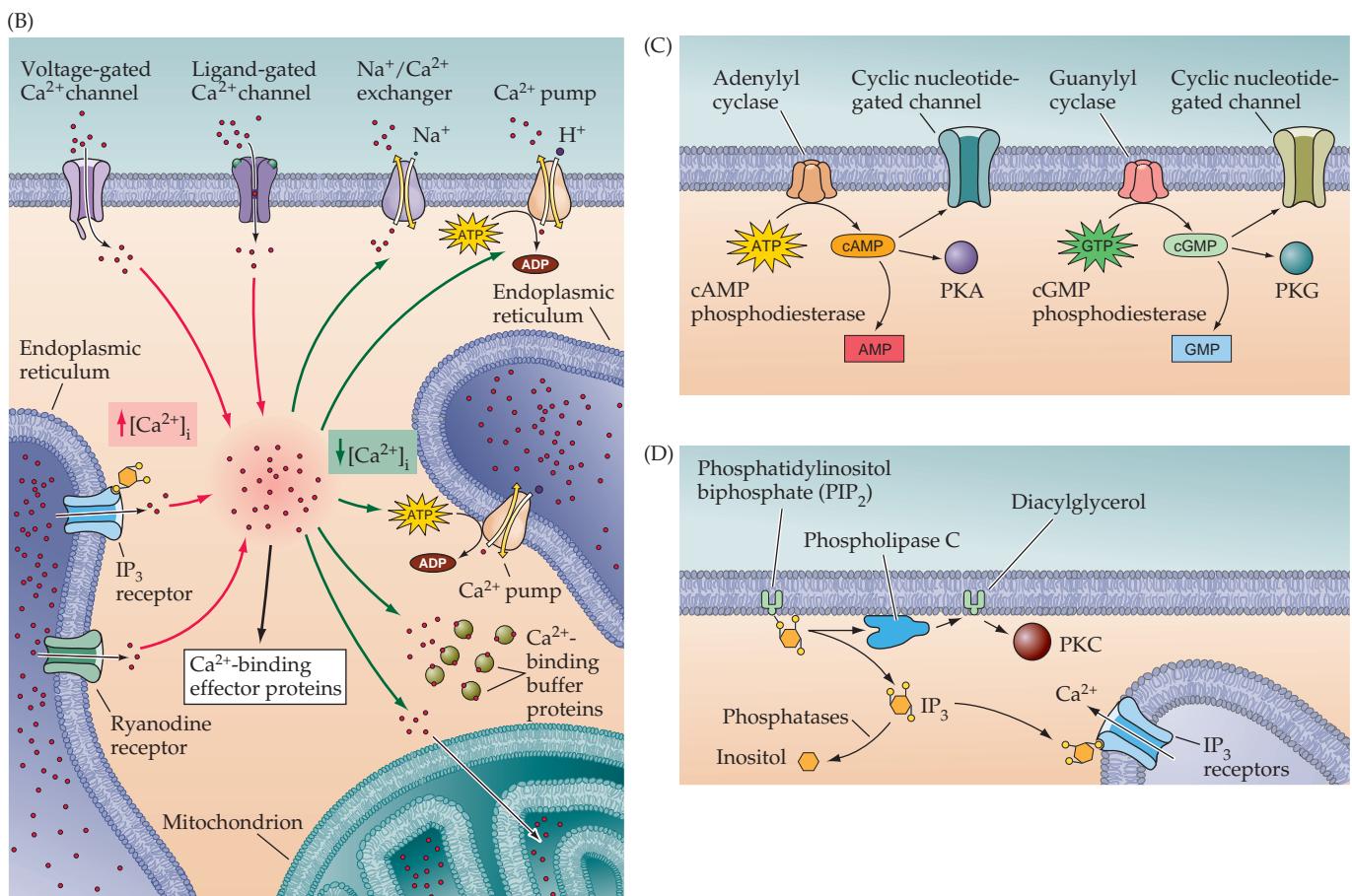
Second Messengers

Neurons use many different second messengers as intracellular signals. These messengers differ in the mechanism by which they are produced and removed, as well as their downstream targets and effects (Figure 7.7A). This section summarizes the attributes of some of the principal second messengers.

- **Calcium.** The calcium ion (Ca²⁺) is perhaps the most common intracellular messenger in neurons. Indeed, few neuronal functions are immune to the influence—direct or indirect—of Ca²⁺. In all cases, information is transmitted by a transient rise in the cytoplasmic calcium concentration, which

(A) Second messenger	Sources	Intracellular targets	Removal mechanisms
Ca ²⁺	Plasma membrane: Voltage-gated Ca ²⁺ channels Various ligand-gated channels Endoplasmic reticulum: IP ₃ receptors Ryanodine receptors	Calmodulin Protein kinases Protein phosphatases Ion channels Synaptotagmin Many other Ca ²⁺ -binding proteins	Plasma membrane: Na ⁺ /Ca ²⁺ exchanger Ca ²⁺ pump Endoplasmic reticulum: Ca ²⁺ pump Mitochondria
Cyclic AMP	Adenylyl cyclase acts on ATP	Protein kinase A Cyclic nucleotide-gated channels	cAMP phosphodiesterase
Cyclic GMP	Guanylyl cyclase acts on GTP	Protein kinase G Cyclic nucleotide-gated channels	cGMP phosphodiesterase
IP ₃	Phospholipase C acts on PIP ₂	IP ₃ receptors on endoplasmic reticulum	Phosphatases
Diacylglycerol	Phospholipase C acts on PIP ₂	Protein kinase C	Various enzymes

Figure 7.7 Neuronal second messengers. (A) Mechanisms responsible for producing and removing second messengers, as well as the downstream targets of these messengers. (B) Proteins involved in delivering calcium to the cytoplasm and in removing calcium from the cytoplasm. (C) Mechanisms of production and degradation of cyclic nucleotides. (D) Pathways involved in production and removal of diacylglycerol (DAG) and IP₃.



allows Ca^{2+} to bind to a large number of Ca^{2+} -binding proteins that serve as molecular targets. One of the most thoroughly studied targets of Ca^{2+} is **calmodulin**, a Ca^{2+} -binding protein abundant in the cytosol of all cells. Binding of Ca^{2+} to calmodulin activates this protein, which then initiates its effects by binding to still other downstream targets, such as protein kinases.

Ordinarily the concentration of Ca^{2+} ions in the cytosol is extremely low, typically 50–100 nanomolar (10^{-9} M). The concentration of Ca^{2+} ions outside neurons—in the bloodstream or cerebrospinal fluid, for instance—is several orders of magnitude higher, typically several millimolar (10^{-3} M). This steep Ca^{2+} gradient is maintained by a number of mechanisms (Figure 7.7B). Most important in this maintenance are two proteins that translocate Ca^{2+} from the cytosol to the extracellular medium: an ATPase called the **calcium pump**, and an $\text{Na}^+/\text{Ca}^{2+}$ **exchanger**, which is a protein that replaces intracellular Ca^{2+} with extracellular sodium ions (see Chapter 4). In addition to these plasma membrane mechanisms, Ca^{2+} is also pumped into the endoplasmic reticulum and mitochondria. These organelles can thus serve as storage depots of Ca^{2+} ions that are later released to participate in signaling events. Finally, nerve cells contain other Ca^{2+} -binding proteins—such as **calbindin**—that serve as Ca^{2+} buffers. Such buffers reversibly bind Ca^{2+} and thus blunt the magnitude and kinetics of Ca^{2+} signals within neurons.

The Ca^{2+} ions that act as intracellular signals enter cytosol by means of one or more types of Ca^{2+} -permeable ion channels (see Chapter 4). These can be voltage-gated Ca^{2+} channels or ligand-gated channels in the plasma membrane, both of which allow Ca^{2+} to flow down the Ca^{2+} gradient and into the cell from the extracellular medium. In addition, other channels allow Ca^{2+} to be released from the interior of the endoplasmic reticulum into the cytosol. These intracellular Ca^{2+} -releasing channels are gated, so they can be opened or closed in response to various intracellular signals. One such channel is the **inositol trisphosphate (IP₃) receptor**. As the name implies, these channels are regulated by IP₃, a second messenger described in more detail below. A second type of intracellular Ca^{2+} -releasing channel is the **ryanodine receptor**, named after a drug that binds to and partially opens these receptors. Among the biological signals that activate ryanodine receptors are cytoplasmic Ca^{2+} and, at least in muscle cells, depolarization of the plasma membrane.

These various mechanisms for elevating and removing Ca^{2+} ions allow precise control of both the timing and location of Ca^{2+} signaling within neurons, which in turn permit Ca^{2+} to control many different signaling events. For example, voltage-gated Ca^{2+} channels allow Ca^{2+} concentrations to rise very rapidly and locally within presynaptic terminals to trigger neurotransmitter release, as already described in Chapter 5. Slower and more widespread rises in Ca^{2+} concentration regulate a wide variety of other responses, including gene expression in the cell nucleus.

- **Cyclic nucleotides.** Another important group of second messengers are the cyclic nucleotides, specifically cyclic adenosine monophosphate (cAMP) and cyclic guanosine monophosphate (cGMP) (Figure 7.7C). Cyclic AMP is a derivative of the common cellular energy storage molecule, ATP. Cyclic AMP is produced when G-proteins activate adenylyl cyclase in the plasma membrane. This enzyme converts ATP into cAMP by removing two phosphate groups from the ATP. Cyclic GMP is similarly produced from GTP by the action of guanylyl cyclase. Once the intracellular concentration of cAMP or cGMP is elevated, these nucleotides can bind to two different classes of targets. The most common targets of cyclic nucleotide action are protein kinases, either the cAMP-dependent protein kinase (PKA) or the cGMP-dependent

protein kinase (PKG). These enzymes mediate many physiological responses by phosphorylating target proteins, as described in the following section. In addition, cAMP and cGMP can bind to certain ligand-gated ion channels, thereby influencing neuronal signaling. These cyclic nucleotide-gated channels are particularly important in phototransduction and other sensory transduction processes, such as olfaction. Cyclic nucleotide signals are degraded by phosphodiesterases, enzymes that cleave phosphodiester bonds and convert cAMP into AMP or cGMP into GMP.

- **Diacylglycerol and IP₃.** Remarkably, membrane lipids can also be converted into intracellular second messengers (Figure 7.7D). The two most important messengers of this type are produced from phosphatidylinositol bisphosphate (PIP₂). This lipid component is cleaved by phospholipase C, an enzyme activated by certain G-proteins and by calcium ions. Phospholipase C splits the PIP₂ into two smaller molecules that each act as second messengers. One of these messengers is diacylglycerol (DAG), a molecule that remains within the membrane and activates protein kinase C, which phosphorylates substrate proteins in both the plasma membrane and elsewhere. The other messenger is inositol trisphosphate (IP₃), a molecule that leaves the cell membrane and diffuses within the cytosol. IP₃ binds to IP₃ receptors, channels that release calcium from the endoplasmic reticulum. Thus, the action of IP₃ is to produce yet another second messenger (perhaps a third messenger, in this case!) that triggers a whole spectrum of reactions in the cytosol. The actions of DAG and IP₃ are terminated by enzymes that convert these two molecules into inert forms that can be recycled to produce new molecules of PIP₂.

Second Messenger Targets: Protein Kinases and Phosphatases

As already mentioned, second messengers typically regulate neuronal functions by modulating the phosphorylation state of intracellular proteins (Figure 7.8). Phosphorylation (the addition of phosphate groups) rapidly and reversibly changes protein function. Proteins are phosphorylated by a wide variety of **protein kinases**; phosphate groups are removed by other enzymes called **protein phosphatases**. The degree of phosphorylation of a target protein thus reflects a balance between the competing actions of protein kinases and phosphatases, thus integrating a host of cellular signaling pathways. The substrates of protein kinases and phosphatases include enzymes, neurotransmitter receptors, ion channels, and structural proteins.

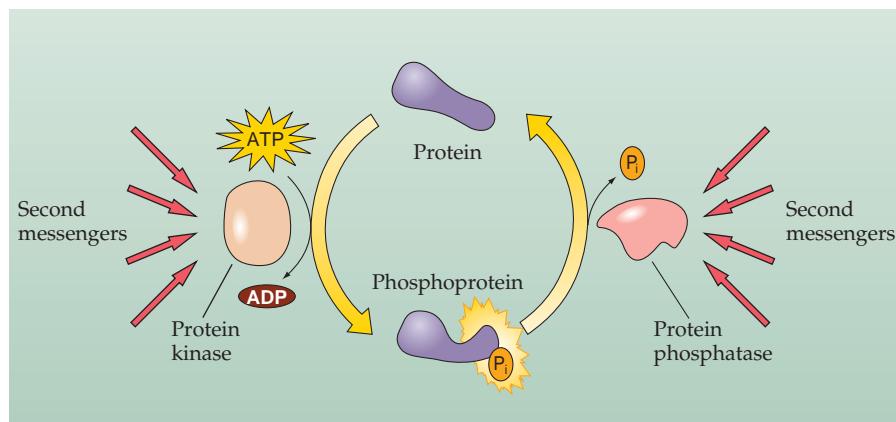


Figure 7.8 Regulation of cellular proteins by phosphorylation. Protein kinases transfer phosphate groups (P_i) from ATP to serine, threonine, or tyrosine residues on substrate proteins. This phosphorylation reversibly alters the structure and function of cellular proteins. Removal of the phosphate groups is catalyzed by protein phosphatases. Both kinases and phosphatases are regulated by a variety of intracellular second messengers.

Protein kinases and phosphatases typically act either on the serine and threonine residues (Ser/Thr kinases or phosphatases) or the tyrosine residues (Tyr kinases or phosphatases) of their substrates. Some of these enzymes act specifically on only one or a handful of protein targets, while others are multifunctional and have a broad range of substrate proteins. The activity of protein kinases and phosphatases can be regulated either by second messengers, such as cAMP or Ca^{2+} , or by extracellular chemical signals, such as growth factors (see Chapter 22). Typically, second messengers activate Ser/Thr kinases, whereas extracellular signals activate Tyr kinases. Although thousands of protein kinases are expressed in the brain, a relatively small number function as regulators of neuronal signaling.

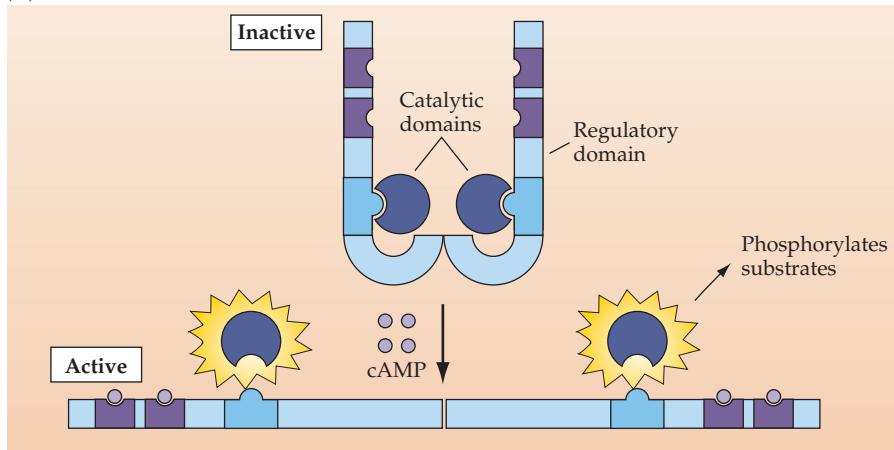
- *cAMP-dependent protein kinase (PKA)*. The primary effector of cAMP is the cAMP-dependent protein kinase (PKA). PKA is a tetrameric complex of two catalytic subunits and two inhibitory (regulatory) subunits. cAMP activates PKA by binding to the regulatory subunits and causing them to release active catalytic subunits. Such displacement of inhibitory domains is a general mechanism for activation of several protein kinases by second messengers (Figure 7.9A). The catalytic subunit of PKA phosphorylates serine and threonine residues of many different target proteins. Although this subunit is similar to the catalytic domains of other protein kinases, distinct amino acids allow the PKA to bind to specific target proteins, thus allowing only those targets to be phosphorylated in response to intracellular cAMP signals.

- *Ca^{2+} /calmodulin-dependent protein kinase type II (CaMKII)*. Ca^{2+} ions binding to calmodulin can regulate protein phosphorylation/dephosphorylation. In neurons, the most abundant Ca^{2+} /calmodulin-dependent protein kinase is CaMKII, a multifunctional Ser/Thr protein kinase. CaMKII is composed of approximately 14 subunits, which in the brain are the α and β types. Each subunit contains a catalytic domain and a regulatory domain, as well as other domains that allow the enzyme to oligomerize and target to the proper region within the cell. Ca^{2+} /calmodulin activates CaMKII by displacing the inhibitory domain from the catalytic site (Figure 7.9B). CaMKII phosphorylates a large number of substrates, including ion channels and other proteins involved in intracellular signal transduction.

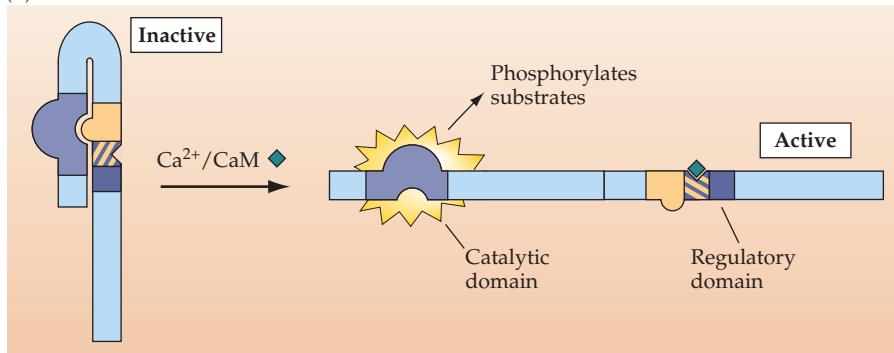
- *Protein kinase C (PKC)*. Another important group of Ser/Thr protein kinases is protein kinase C (PKC). PKCs are diverse monomeric kinases activated by the second messengers DAG and Ca^{2+} . DAG causes PKC to move from the cytosol to the plasma membrane, where it also binds Ca^{2+} and phosphatidylserine, a membrane phospholipid (Figure 7.9C). These events relieve autoinhibition and cause PKC to phosphorylate various protein substrates. PKC also diffuses to sites other than the plasma membrane—such as the cytoskeleton, perinuclear sites, and the nucleus—where it phosphorylates still other substrate proteins. Prolonged activation of PKC can be accomplished with phorbol esters, tumor-promoting compounds that activate PKC by mimicking DAG.

- *Protein tyrosine kinases*. Two classes of protein kinases transfer phosphate groups to tyrosine residues on substrate proteins. Receptor tyrosine kinases are transmembrane proteins with an extracellular domain that binds to protein ligands (growth factors, neurotrophic factors, or cytokines) and an intracellular catalytic domain that phosphorylates the relevant substrate proteins. Non-receptor tyrosine kinases are cytoplasmic or membrane-associated enzymes that are indirectly activated by extracellular signals. Tyrosine phosphorylation is less common than Ser/Thr phosphorylation, and it often serves to recruit signaling molecules to the phosphorylated protein. Tyrosine

(A) PKA



(B) CaMKII



(C) PKC

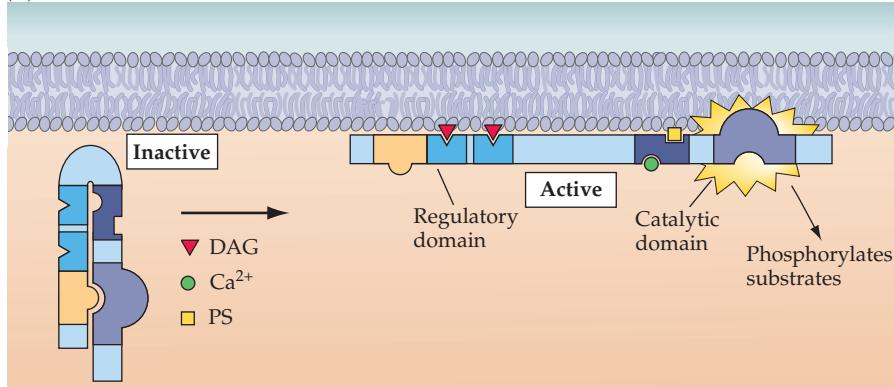


Figure 7.9 Mechanism of activation of protein kinases. Protein kinases contain several specialized domains with specific functions. Each of the kinases has homologous catalytic domains responsible for transferring phosphate groups to substrate proteins. These catalytic domains are kept inactive by the presence of an autoinhibitory domain that occupies the catalytic site. Binding of second messengers, such as cAMP, DAG, and Ca^{2+} , to the appropriate regulatory domain of the kinase removes the autoinhibitory domain and allows the catalytic domain to be activated. For some kinases, such as PKC and CaMKII, the autoinhibitory and catalytic domains are part of the same molecule. For other kinases, such as PKA, the autoinhibitory domain is a separate subunit.

kinases are particularly important for cell growth and differentiation (see Chapters 21 and 22).

- **Mitogen-activated protein kinase (MAPK).** In addition to protein kinases that are directly activated by second messengers, some of these molecules can be activated by other signals, such as phosphorylation by another protein kinase. Important examples of such protein kinases are the mitogen-activated protein kinases (MAPKs), also called extracellular signal-regulated kinases (ERKs). MAPKs were first identified as participants in the control of cell growth and are now known to have many other signaling functions.

MAPKs are normally inactive in neurons but become activated when they are phosphorylated by other kinases. In fact, MAPKs are part of a kinase cascade in which one protein kinase phosphorylates and activates the next protein kinase in the cascade. The extracellular signals that trigger these kinase cascades are often extracellular growth factors that bind to receptor tyrosine kinases that, in turn, activate monomeric G-proteins such as ras. Once activated, MAPKs can phosphorylate transcription factors, proteins that regulate gene expression. Among the wide variety of other MAPK substrates are various enzymes, including other protein kinases, and cytoskeletal proteins.

The best-characterized protein phosphatases are the Ser/Thr phosphatases PP1, PP2A, and PP2B (also called calcineurin). In general, protein phosphatases display less substrate specificity than protein kinases. Their limited specificity may arise from the fact that the catalytic subunits of the three major protein phosphatases are highly homologous, though each still associates with specific targeting or regulatory subunits. PP1 dephosphorylates a wide array of substrate proteins and is probably the most prevalent Ser/Thr protein phosphatase in mammalian cells. PP1 activity is regulated by several inhibitory proteins expressed in neurons. PP2A is a multisubunit enzyme with a broad range of substrates that overlap with PP1. PP2B, or calcineurin, is present at high levels in neurons. A distinctive feature of this phosphatase is its activation by Ca^{2+} /calmodulin. PP2B is composed of a catalytic and a regulatory subunit. Ca^{2+} /calmodulin activates PP2B primarily by binding to the catalytic subunit and displacing the inhibitory regulatory domain. PP2B generally does not have the same molecular targets as CaMKII, even though both enzymes are activated by Ca^{2+} /calmodulin.

In summary, activation of membrane receptors can elicit complex cascades of enzyme activation, resulting in second messenger production and protein phosphorylation or dephosphorylation. These cytoplasmic signals produce a variety of rapid physiological responses by transiently regulating enzyme activity, ion channels, cytoskeletal proteins, and many other cellular processes. In addition, such signals can propagate to the nucleus to cause long-lasting changes in gene expression.

Nuclear Signaling

Second messengers elicit prolonged changes in neuronal function by promoting the synthesis of new RNA and protein. The resulting accumulation of new proteins requires at least 30–60 minutes, a time frame that is orders of magnitude slower than the responses mediated by ion fluxes or phosphorylation. Likewise, the reversal of such events requires hours to days. In some cases, genetic “switches” can be thrown to permanently alter a neuron, as in neuronal differentiation (see Chapter 21).

The amount of protein present in cells is determined primarily by the rate of transcription of DNA into RNA (Figure 7.10). The first step in RNA synthesis is the decondensation of the structure of chromatin to provide binding sites for the RNA polymerase complex and for **transcriptional activator proteins**, also called **transcription factors**. Transcriptional activator proteins attach to binding sites that are present on the DNA molecule near the start of the target gene sequence; they also bind to other proteins that promote unwrapping of DNA. The net result of these actions is to allow RNA polymerase, an enzyme complex, to assemble on the **promoter** region of the DNA and begin transcription. In addition to clearing the promoter for RNA polymerase, activator proteins can stimulate transcription by interacting

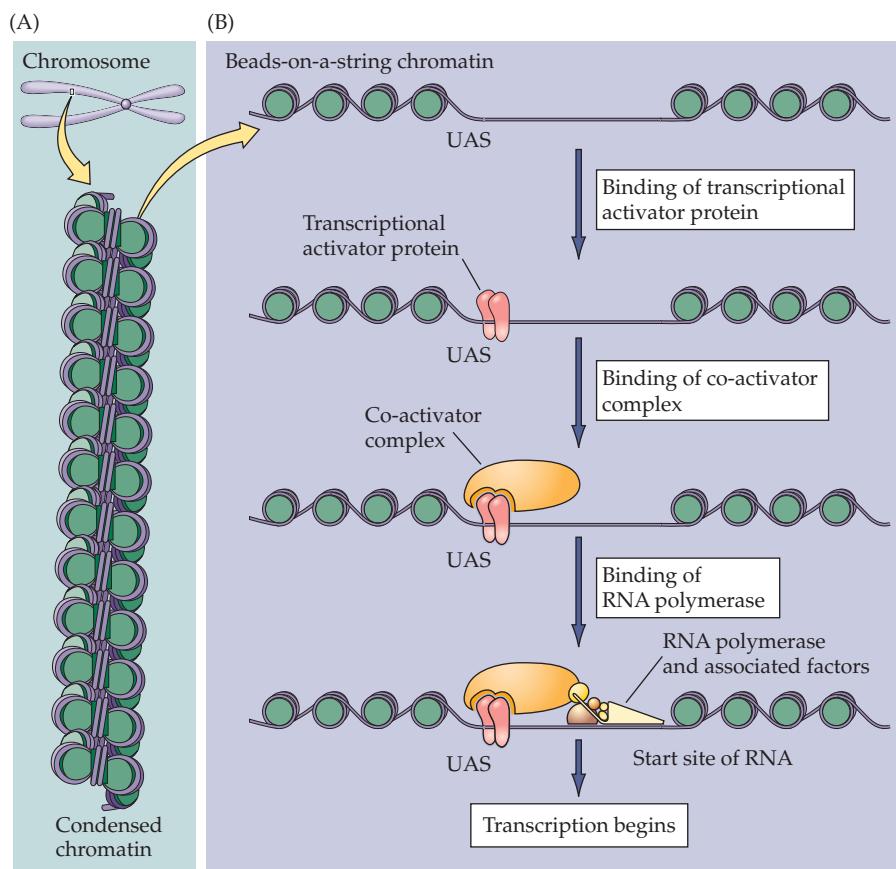


Figure 7.10 Steps involved in transcription of DNA into RNA. Condensed chromatin (A) is decondensed into a beads-on-a-DNA-string array (B) in which an upstream activator site (UAS) is free of proteins and is bound by a sequence-specific transcriptional activator protein (transcription factor). The transcriptional activator protein then binds co-activator complexes that enable the RNA polymerase with its associated factors to bind at the start site of transcription and initiate RNA synthesis.

with the RNA polymerase complex or by interacting with other activator proteins that influence the polymerase.

Intracellular signal transduction cascades regulate gene expression by converting transcriptional activator proteins from an inactive state to an active state in which they are able to bind to DNA. This conversion comes about in several ways. The key activator proteins and the mechanisms that allow them to regulate gene expression in response to signaling events are briefly summarized in the following sections.

- **CREB.** The cAMP response element binding protein, usually abbreviated **CREB**, is a ubiquitous transcriptional activator (Figure 7.11). CREB is normally bound to its binding site on DNA (called the cAMP response element, or CRE), either as a homodimer or bound to another, closely related transcription factor. In unstimulated cells, CREB is not phosphorylated and has little or no transcriptional activity. However, phosphorylation of CREB greatly potentiates transcription. Several signaling pathways are capable of causing CREB to be phosphorylated. Both PKA and the ras pathway, for example, can phosphorylate CREB. CREB can also be phosphorylated in response to increased intracellular calcium, in which case the CRE site is also called the CaRE (calcium response element) site. The calcium-dependent phosphorylation of CREB is primarily caused by Ca^{2+} /calmodulin kinase IV (a relative of CaMKII) and by MAP kinase, which leads to prolonged CREB phosphorylation. CREB phosphorylation must be maintained long enough for transcription to ensue, even though neuronal electrical activity only tran-

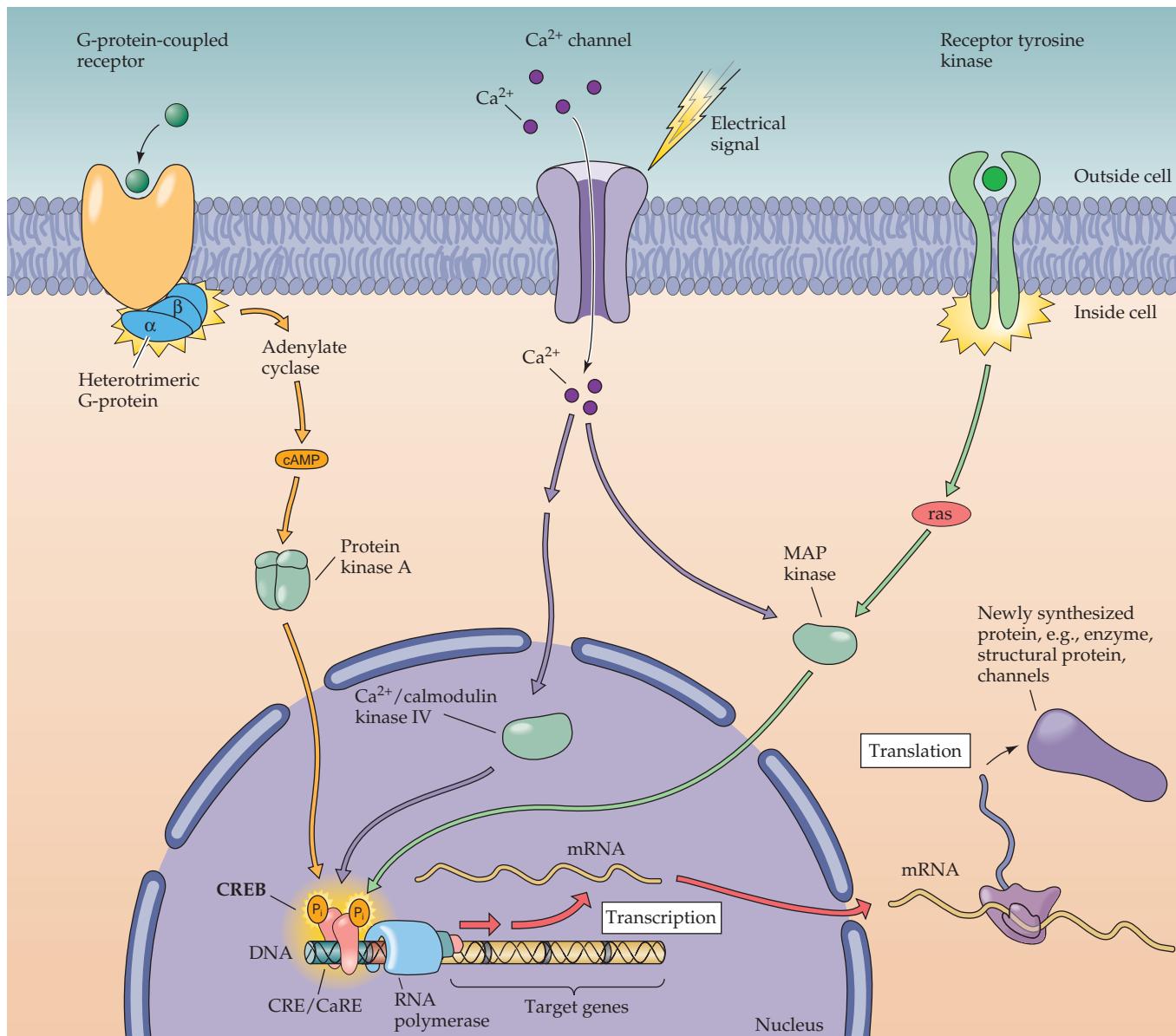


Figure 7.11 Transcriptional regulation by CREB. Multiple signaling pathways converge by activating kinases that phosphorylate CREB. These include PKA, Ca²⁺/calmodulin kinase IV, and MAP kinase. Phosphorylation of CREB allows it to bind co-activators (not shown in the figure), which then stimulate RNA polymerase to begin synthesis of RNA. RNA is then processed and exported to the cytoplasm, where it serves as mRNA for translation into protein.

siently raises intracellular calcium concentration. Such signaling cascades can potentiate CREB-mediated transcription by inhibiting a protein phosphatase that dephosphorylates CREB. CREB is thus an example of the convergence of multiple signaling pathways onto a single transcriptional activator.

Many genes whose transcription is regulated by CREB have been identified. CREB-sensitive genes include the immediate early gene, *c-fos* (see below), the neurotrophin BDNF (see Chapter 22), the enzyme tyrosine hydroxylase (which is important for synthesis of catecholamine neurotransmitters; see Chapter 6), and many neuropeptides (including somatostatin, enkephalin, and corticotropin releasing hormone). CREB also is thought to mediate long-lasting changes in brain function. For example, CREB has been implicated in spatial learning, behavioral sensitization, long-term memory of odorant-conditioned behavior, and long-term synaptic plasticity (see Chapters 23 and 24).

- *Nuclear receptors.* Nuclear receptors for membrane-permeant ligands also are transcriptional activators. The receptor for glucocorticoid hormones illustrates one mode of action of such receptors. In the absence of glucocorticoid hormones, the receptors are located in the cytoplasm. Binding of glucocorticoids causes the receptor to unfold and move to the nucleus, where it binds a specific recognition site on the DNA. This DNA binding activates the relevant RNA polymerase complex to initiate transcription and subsequent gene expression. Thus, a critical regulatory event for steroid receptors is their translocation to the nucleus to allow DNA binding.

The receptors for thyroid hormone (TH) and other non-steroid nuclear receptors illustrate a second mode of regulation. In the absence of TH, the receptor is bound to DNA and serves as a potent repressor of transcription. Upon binding TH, the receptor undergoes a conformational change that ultimately opens the promoter for polymerase binding. Hence, TH binding switches the receptor from being a repressor to being an activator of transcription.

- *c-fos.* A different strategy of gene regulation is apparent in the function of the transcriptional activator protein, *c-fos*. In resting cells, *c-fos* is present at a very low concentration. However, stimulation of the target cell causes *c-fos* to be synthesized, and the amount of this protein rises dramatically over 30–60 minutes. Therefore, *c-fos* is considered to be an **immediate early gene** because its synthesis is directly triggered by the stimulus. Once synthesized, *c-fos* protein can act as a transcriptional activator to induce synthesis of second-order genes. These are termed **delayed response genes** because their activity is delayed by the fact that an immediate early gene—*c-fos* in this case—needs to be activated first.

Multiple signals converge on *c-fos*, activating different transcription factors that bind to at least three distinct sites in the promoter region of the gene. The regulatory region of the *c-fos* gene contains a binding site that mediates transcriptional induction by cytokines and ciliary neurotropic factor. Another site is targeted by growth factors such as neurotrophins through ras and protein kinase C, and a CRE/CaRE that can bind to CREB and thereby respond to cAMP or calcium entry resulting from electrical activity. In addition to synergistic interactions among these *c-fos* sites, transcriptional signals can be integrated by converging on the same activator, such as CREB.

Nuclear signaling events typically result in the generation of a large and relatively stable complex composed of a functional transcriptional activator protein, additional proteins that bind to the activator protein, and the RNA polymerase and associated proteins bound at the start site of transcription. Most of the relevant signaling events act to “seed” this complex by generating an active transcriptional activator protein by phosphorylation, by inducing a conformational change in the activator upon ligand binding, by fostering nuclear localization, by removing an inhibitor, or simply by making more activator protein.

Examples of Neuronal Signal Transduction

Understanding the general properties of signal transduction processes at the plasma membrane, in the cytosol, and within the nucleus make it possible to consider how these processes work in concert to mediate specific functions in the brain. Three important signal transduction pathways illustrate some of the roles of intracellular signal transduction processes in the nervous system.

- **NGF/TrkA.** The first of these is signaling by the **nerve growth factor (NGF)**. This protein is a member of the neurotrophin growth factor family and is required for the differentiation, survival, and synaptic connectivity of sympathetic and sensory neurons (see Chapter 22). NGF works by binding to a high-affinity tyrosine kinase receptor, TrkA, found on the plasma membrane of these target cells (Figure 7.12). NGF binding causes TrkA receptors to dimerize, and the intrinsic tyrosine kinase activity of each receptor then phosphorylates its partner receptor. Phosphorylated TrkA receptors trigger the ras cascade, resulting in the activation of multiple protein kinases. Some of these kinases translocate to the nucleus to activate transcriptional activators, such as CREB. This ras-based component of the NGF pathway is primarily responsible for inducing and maintaining differentiation of NGF-sensitive neurons. Phosphorylation of TrkA also causes this receptor to stimulate the activity of phospholipase C, which increases production of IP₃ and DAG. IP₃ induces release of Ca²⁺ from the endoplasmic reticulum, and diacylglycerol activates PKC. These two second messengers appear to target many of the same downstream effectors as ras. Finally, activation of TrkA receptors also causes activation of other protein kinases (such as Akt kinase) that inhibit cell death. This pathway, therefore, primarily mediates the NGF-dependent survival of sympathetic and sensory neurons described in Chapter 22.

- **Long-term depression (LTD).** The interplay between several intracellular signals can be observed at the excitatory synapses that innervate Purkinje

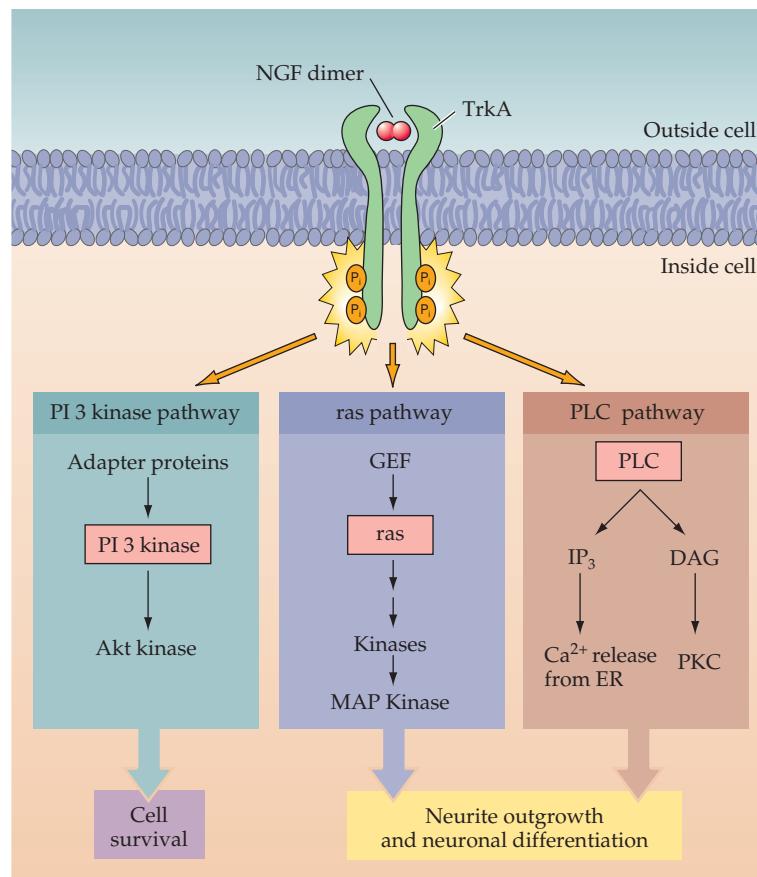


Figure 7.12 Mechanism of action of NGF. NGF binds to a high-affinity tyrosine kinase receptor, TrkA, on the plasma membrane to induce phosphorylation of TrkA at two different tyrosine residues. These phosphorylated tyrosines serve to tether various adapter proteins or phospholipase C (PLC), which, in turn, activate three major signaling pathways: the PI 3 kinase pathway leading to activation of Akt kinase, the ras pathway leading to MAP kinases, and the PLC pathway leading to release of intracellular Ca²⁺ and activation of PKC. The ras and PLC pathways primarily stimulate processes responsible for neuronal differentiation, while the PI 3 kinase pathway is primarily involved in cell survival.

cells in the cerebellum. These synapses are central to information flow through the cerebellar cortex, which in turn helps coordinate motor movements (see Chapter 18). One of the synapses is between the parallel fibers (PFs) and their Purkinje cell targets. LTD is a form of synaptic plasticity that causes the PF synapses to become less effective (see Chapter 24). When PFs are active, they release the neurotransmitter glutamate onto the dendrites of Purkinje cells. This activates AMPA-type receptors, which are ligand-gated ion channels (see Chapter 6), and causes a small EPSP that briefly depolarizes the Purkinje cell. In addition to this electrical signal, PF synaptic transmission also generates two second messengers within the Purkinje cell (Figure 7.13). The glutamate released by PFs activates metabotropic glutamate receptors, which stimulates phospholipase C to produce IP_3 and DAG. When the PF synapses alone are active, these intracellular signals are insufficient to open IP_3 receptors or to stimulate PKC.

LTD is induced when PF synapses are activated at the same time as the glutamatergic climbing fiber synapses that also innervate Purkinje cells. The climbing fiber synapses produce large EPSPs that strongly depolarize the membrane potential of the Purkinje cell. This depolarization allows Ca^{2+} to

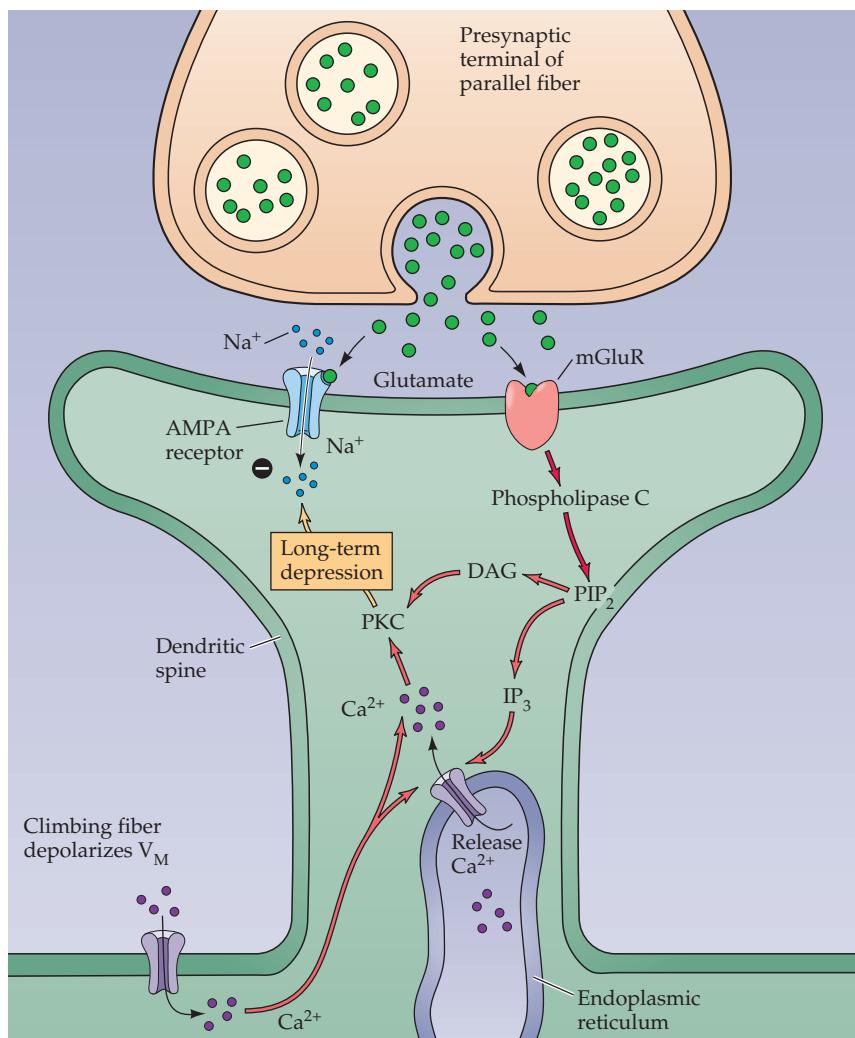


Figure 7.13 Signaling at cerebellar parallel fiber synapses. Glutamate released by parallel fibers activates both AMPA-type and metabotropic receptors. The latter produces IP_3 and DAG within the Purkinje cell. When paired with a rise in Ca^{2+} associated with activity of climbing fiber synapses, the IP_3 causes Ca^{2+} to be released from the endoplasmic reticulum, while Ca^{2+} and DAG together activate protein kinase C. These signals together change the properties of AMPA receptors to produce LTD.

enter the Purkinje cell via voltage-gated Ca^{2+} channels. When both synapses are simultaneously activated, the rise in intracellular Ca^{2+} concentration caused by the climbing fiber synapse enhances the sensitivity of IP_3 receptors to the IP_3 produced by PF synapses and allows the IP_3 receptors within the Purkinje cell to open. This releases Ca^{2+} from the endoplasmic reticulum and further elevates Ca^{2+} concentration locally near the PF synapses. This larger rise in Ca^{2+} , in conjunction with the DAG produced by the PF synapses, activates PKC. PKC in turn phosphorylates a number of substrate proteins. Ultimately, these signaling processes change AMPA-type receptors at the PF synapse, so that these receptors produce smaller electrical signals in response to the glutamate released from the PFs. This weakening of the PF synapse is the final cause of LTD.

In short, transmission at Purkinje cell synapses produces brief electrical signals and chemical signals that last much longer. The temporal interplay between these signals allows LTD to occur only when both PF and climbing fiber synapses are active. The actions of IP_3 , DAG and Ca^{2+} also are restricted to small parts of the Purkinje cell dendrite, which is a more limited spatial range than the EPSPs, which spread throughout the entire dendrite and cell body of the Purkinje cell. Thus, in contrast to the electrical signals, the second messenger signals can impart precise information about the location of active synapses and allow LTD to occur only in the vicinity of active PFs.

- *Phosphorylation of tyrosine hydroxylase.* A third example of intracellular signaling in the nervous system is the regulation of the enzyme tyrosine hydroxylase. Tyrosine hydroxylase governs the synthesis of the catecholamine neurotransmitters: dopamine, norepinephrine, and epinephrine (see Chapter 6). A number of signals, including electrical activity, other neurotransmitters, and NGF, increase the rate of catecholamine synthesis by increasing the catalytic activity of tyrosine hydroxylase (Figure 7.14). The rapid increase of tyrosine hydroxylase activity is largely due to phosphorylation of this enzyme.

Tyrosine hydroxylase is a substrate for several protein kinases, including PKA, CaMKII, MAP kinase, and PKC. Phosphorylation causes conformational changes that increase the catalytic activity of tyrosine hydroxylase. Stimuli that elevate cAMP, Ca^{2+} , or DAG can all increase tyrosine hydroxylase activity and thus increase the rate of catecholamine biosynthesis. This regulation by several different signals allows for close control of tyrosine hydroxylase activity, and illustrates how several different pathways can converge to influence a key enzyme involved in synaptic transmission.

Summary

A diversity of signal transduction pathways exist within all neurons. Activation of these pathways typically is initiated by chemical signals such as neurotransmitters and hormones. These molecules bind to receptors that include ligand-gated ion channels, G-protein-coupled receptors and tyrosine kinase receptors. Many of these receptors activate either heterotrimeric or monomeric G-proteins that regulate intracellular enzyme cascades and/or ion channels. A common outcome of the activation of these receptors is the production of second messengers, such as cAMP, Ca^{2+} , and IP_3 , that bind to effector enzymes. Particularly important effectors are protein kinases and phosphatases that regulate the phosphorylation state of their substrates, and thus their function. These substrates can be metabolic enzymes or other signal transduction molecules, such as ion channels, protein kinases, or transcription factors that regulate gene expression. Examples of transcription

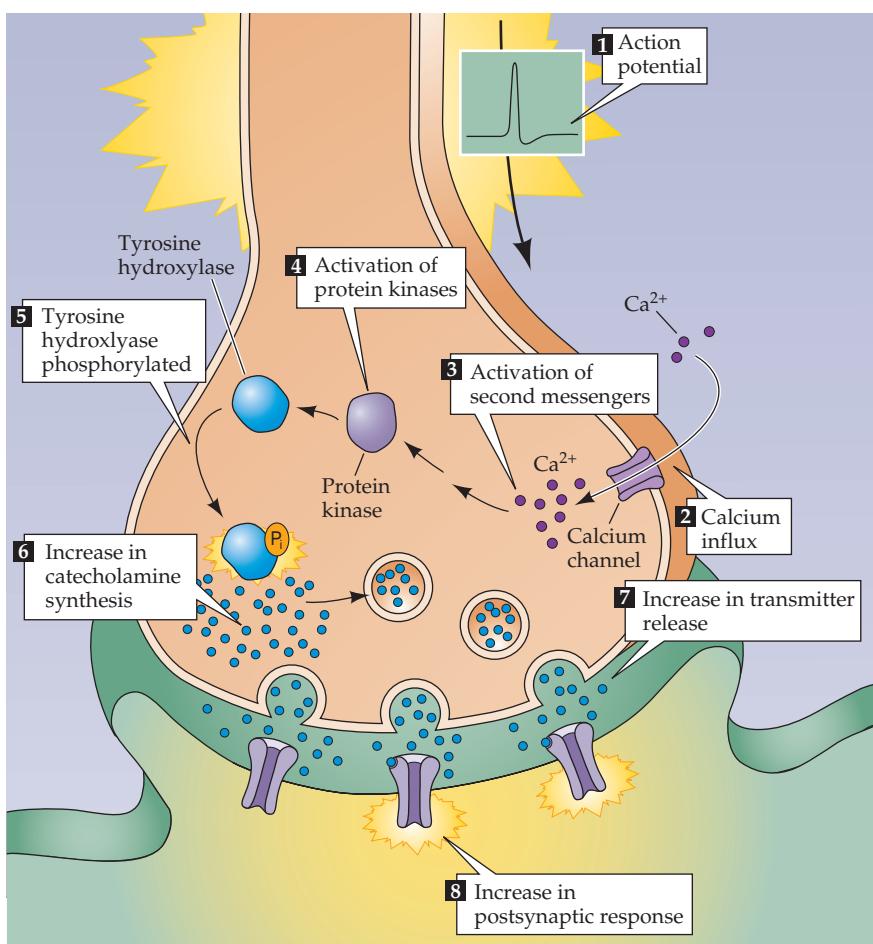


Figure 7.14 Regulation of tyrosine hydroxylase by protein phosphorylation. This enzyme governs the synthesis of the catecholamine neurotransmitters and is stimulated by a number of intracellular signals. In the example shown here, neuronal electrical activity (1) causes influx of Ca^{2+} (2). The resultant rise in intracellular Ca^{2+} concentration (3) activates protein kinases (4), which phosphorylates tyrosine hydroxylase (5) to stimulate catecholamine synthesis (6). This, in turn, increases release of catecholamines (7) and enhances the postsynaptic response produced by the synapse (8).

factors include CREB, steroid hormone receptors, and c-fos. This plethora of molecular components allows intracellular signal transduction pathways to generate responses over a wide range of times and distances, greatly augmenting and refining the information-processing ability of neuronal circuits and ultimately systems.

Additional Reading

Reviews

AUGUSTINE, G. J., F. SANTAMARIA AND K. TANAKA (2003) Local calcium signaling in neurons. *Neuron* 40: 331–346.

DEISSEROTH, K., P. G. MERMELSTEIN, H. XIA AND R. W. TSien (2003) Signaling from synapse to nucleus: The logic behind the mechanisms. *Curr. Opin. Neurobiol.* 13: 354–365.

EXTON, J. H. (1998) Small GTPases. *J. Biol. Chem.* 273: 19923.

FISCHER, E. H. (1999) Cell signaling by protein tyrosine phosphorylation. *Adv. Enzyme Regul.* Review 39: 359–369.

FRIEDMAN, W. J. AND L. A. GREENE (1999) Neurotrophin signaling via Trks and p75. *Exp. Cell Res.* 253: 131–142.

GILMAN, A. G. (1984) G proteins and dual control of adenylate cyclase. *Cell* 36: 577–579.

GRAVES J. D. AND E. G. KREBS (1999) Protein phosphorylation and signal transduction. *Pharmacol. Ther.* 82: 111–121.

KENNEDY, M. B. (2000) Signal-processing machines at the postsynaptic density. *Science* 290: 750–754.

KUMER, S. AND K. VRANA (1996) Intricate regulation of tyrosine hydroxylase activity and gene expression. *J. Neurochem.* 67: 443–462.

LEVITAN, I. B. (1999) Modulation of ion channels by protein phosphorylation. How the brain works. *Adv. Second Mess. Phosphoprotein Res.* 33: 3–22.

NEER, E. J. (1995) Heterotrimeric G proteins: Organizers of transmembrane signals. *Cell* 80: 249–257.

RODBELL, M. (1995) Nobel Lecture. Signal transduction: Evolution of an idea. *Bioscience Reports* 15: 117–133.

SHENG, M. AND M. J. KIM (2002) Postsynaptic signaling and plasticity mechanisms. *Science* 298: 776–780.

WEST, A. E. AND 8 OTHERS (2001) Calcium regulation of neuronal gene expression. *Proc. Natl. Acad. Sci. USA* 98: 11024–11031.

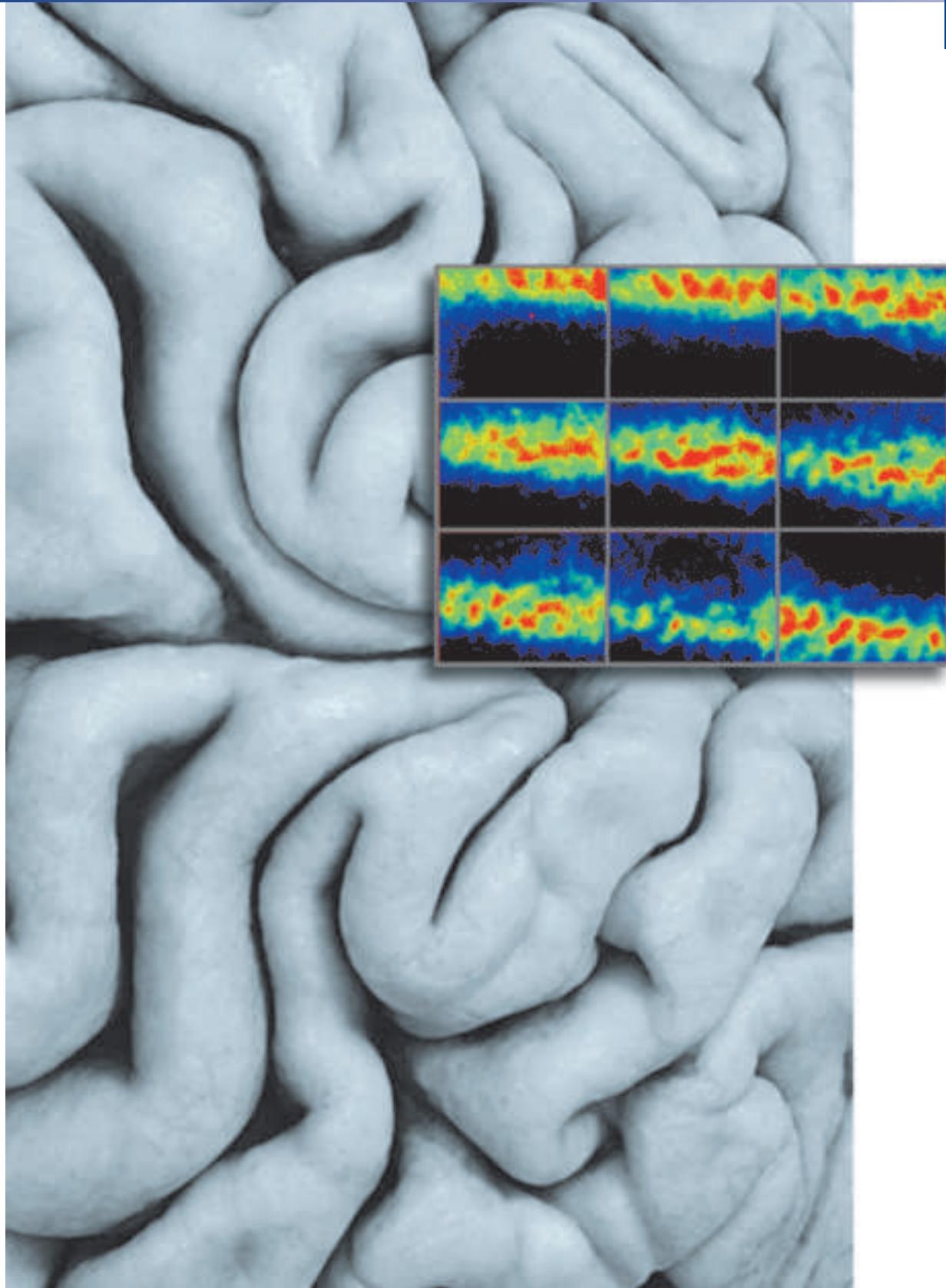
Important Original Papers

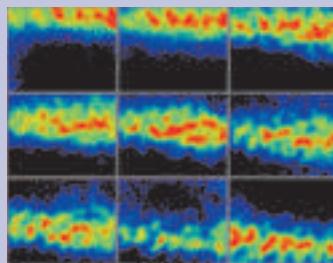
- BACSKAI, B. J. AND 6 OTHERS (1993) Spatially resolved dynamics of cAMP and protein kinase A subunits in *Aplysia* sensory neurons. *Science* 260: 222–226.
- BURGESS, G. M., P. P. GODFREY, J. S. MCKINNEY, M. J. BERRIDGE, R. F. IRVINE AND J. W. PUTNEY JR. (1984) The second messenger linking receptor activation to internal Ca²⁺ release in liver. *Nature* 309: 63–66.
- CONNOR, J. A. (1986) Digital imaging of free calcium changes and of spatial gradients in growing processes in single, mammalian central nervous system cells. *Proc. Natl. Acad. Sci. USA* 83: 6179–6183.
- DE KONINCK, P. AND H. SCHULMAN (1998) Sensitivity of CaM kinase II to the frequency of Ca²⁺ oscillations. *Science* 279: 227–230.
- FINCH, E. A. AND G. J. AUGUSTINE (1998) Local calcium signaling by IP₃ in Purkinje cell dendrites. *Nature* 396: 753–756.
- HARRIS, B. A., J. D. ROBISHAW, S. M. MUMBY AND A. G. GILMAN (1985) Molecular cloning of complementary DNA for the alpha subunit of the G protein that stimulates adenylate cyclase. *Science* 229: 1274–1277.
- KAMMERMEIER, P. J. AND S. R. IKEDA (1999) Expression of RGS2 alters the coupling of metabotropic glutamate receptor 1a to M-type K⁺ and N-type Ca²⁺ channels. *Neuron* 22: 819–829.
- KRAFT, A. S. AND W. B. ANDERSON (1983) Phorbol esters increase the amount of Ca²⁺-phospholipid-dependent protein kinase associated with plasma membrane. *Nature* 301: 621–623.
- LINDGREN, N. AND 8 OTHERS (2000) Regulation of tyrosine hydroxylase activity and phosphorylation at ser(19) and ser(40) via activation of glutamate NMDA receptors in rat striatum. *J. Neurochem.* 74: 2470–2477.
- MILLER, S. G. AND M. B. KENNEDY (1986) Regulation of brain type II Ca²⁺/calmodulin-dependent protein kinase by autophosphorylation: A Ca²⁺-triggered molecular switch. *Cell* 44: 861–870.
- NORTHUP, J. K., P. C. STERNWEIS, M. D. SMIGEL, L. S. SCHLEIFER, E. M. ROSS AND A. G. GILMAN (1980) Purification of the regulatory component of adenylate cyclase. *Proc. Natl. Acad. Sci. USA* 77: 6516–6520.
- SAITO, T. AND J. H. SCHWARTZ (1985) Phosphorylation-dependent subcellular translocation of a Ca²⁺/calmodulin-dependent protein kinase produces an autonomous enzyme in *Aplysia* neurons. *J. Cell Biol.* 100: 835–842.
- SHEN, K. AND T. MEYER (1999) Dynamic control of CaMKII translocation and localization in hippocampal neurons by NMDA receptor stimulation. *Science* 284: 162–166.
- SU, Y. AND 7 OTHERS (1995) Regulatory subunit of protein kinase A: Structure of deletion mutant with cAMP binding domains. *Science* 269: 807–813.
- TAO, X., S. FINKBEINER, D. B. ARNOLD, A. J. SHAYWITZ AND M. E. GREENBERG (1998) Ca²⁺ influx regulates BDNF transcription by a CREB family transcription factor-dependent mechanism. *Neuron* 20: 709–726.
- TESMER, J. J., R. K. SUNAHARA, A. G. GILMAN AND S. R. SPRANG (1997) Crystal structure of the catalytic domains of adenylyl cyclase in a complex with G_{sa}-GTP_{γS}. *Science* 278: 1907–1916.
- ZHANG, G., M. G. KAZANIEZ, P. M. BLUMBERG AND J. H. HURLEY (1995) Crystal structure of the cys2 activator-binding domain of protein kinase C delta in complex with phorbol ester. *Cell* 81: 917–924.

Books

- ALBERTS, B., A. JOHNSON, J. LEWIS, M. RAFF, K. ROBERTS AND P. WALTER (2002) *Molecular Biology of the Cell*, 4th Ed. New York: Garland Science.
- CARAFOLI, E. AND C. KLEE (1999) *Calcium as a Cellular Regulator*. New York: Oxford University Press.

Sensation and Sensory Processing





Surface view of the primary visual cortex illustrating patterns of neural activity visualized with intrinsic signal optical imaging techniques (see Box C in Chapter 11). Each panel illustrates the activity evoked by viewing a single thin vertical line. The smooth progression of the activated region from the upper left to the lower right panel illustrates the orderly mapping of visual space. The patchy appearance of the activated region in each panel reflects the columnar mapping of orientation preference. Red regions are the most active, black the least. (Courtesy of Bill Bosking, Justin Crowley, Tom Tucker, and David Fitzpatrick.)

UNIT II

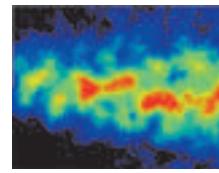
SENSATION AND SENSORY PROCESSING

- 8 *The Somatic Sensory System***
- 9 *Pain***
- 10 *Vision: The Eye***
- 11 *Central Visual Pathways***
- 12 *The Auditory System***
- 13 *The Vestibular System***
- 14 *The Chemical Senses***

Sensation entails the ability to transduce, encode, and ultimately perceive information generated by stimuli arising from both the external and internal environments. Much of the brain is devoted to these tasks. Although the basic senses—somatic sensation, vision, audition, vestibular sensation, and the chemical senses—are very different from one another, a few fundamental rules govern the way the nervous system deals with each of these diverse modalities. Highly specialized nerve cells called receptors convert the energy associated with mechanical forces, light, sound waves, odorant molecules, or ingested chemicals into neural signals—afferent sensory signals—that convey information about the stimulus to the brain. Afferent sensory signals activate central neurons capable of representing both the qualitative and quantitative aspects of the stimulus (what it is and how strong it is) and, in some modalities (somatic sensation, vision, and audition) the location of the stimulus in space (where it is).

The clinical evaluation of patients routinely requires an assessment of the sensory systems to infer the nature and location of potential neurological problems. Knowledge of where and how the different sensory modalities are transduced, relayed, represented, and further processed to generate appropriate behavioral responses is therefore essential to understanding and treating a wide variety of diseases. Accordingly, these chapters on the neurobiology of sensation also introduce some of the major structure/function relationships in the sensory components of the nervous system.

Chapter 8



The Somatic Sensory System

Overview

The somatic sensory system has two major components: a subsystem for the detection of mechanical stimuli (e.g., light touch, vibration, pressure, and cutaneous tension), and a subsystem for the detection of painful stimuli and temperature. Together, these two subsystems give humans and other animals the ability to identify the shapes and textures of objects, to monitor the internal and external forces acting on the body at any moment, and to detect potentially harmful circumstances. This chapter focuses on the mechanosensory subsystem; the pain and temperature subsystem is taken up in the following chapter.

Mechanosensory processing of external stimuli is initiated by the activation of a diverse population of cutaneous and subcutaneous mechanoreceptors at the body surface that relays information to the central nervous system for interpretation and ultimately action. Additional receptors located in muscles, joints, and other deep structures monitor mechanical forces generated by the musculoskeletal system and are called proprioceptors. Mechanosensory information is carried to the brain by several ascending pathways that run in parallel through the spinal cord, brainstem, and thalamus to reach the primary somatic sensory cortex in the postcentral gyrus of the parietal lobe. The primary somatic sensory cortex projects in turn to higher-order association cortices in the parietal lobe, and back to the subcortical structures involved in mechanosensory information processing.

Cutaneous and Subcutaneous Somatic Sensory Receptors

The specialized sensory receptors in the cutaneous and subcutaneous tissues are dauntingly diverse (Table 8.1). They include free nerve endings in the skin, nerve endings associated with specializations that act as amplifiers or filters, and sensory terminals associated with specialized transducing cells that influence the ending by virtue of synapse-like contacts. Based on function, this variety of receptors can be divided into three groups: **mechanoreceptors**, **nociceptors**, and **thermoceptors**. On the basis of their morphology, the receptors near the body surface can also be divided into **free** and **encapsulated** types. Nociceptor and thermoceptor specializations are referred to as **free nerve endings** because the unmyelinated terminal branches of these neurons ramify widely in the upper regions of the dermis and epidermis (as well as in some deeper tissues); their role in pain and temperature sensation is discussed in Chapter 9. Most other cutaneous receptors show some degree of **encapsulation**, which helps determine the nature of the stimuli to which they respond.

Despite their variety, all somatic sensory receptors work in fundamentally the same way: Stimuli applied to the skin deform or otherwise change the

TABLE 8.1
The Major Classes of Somatic Sensory Receptors

Receptor type	Anatomical characteristics	Associated axons ^a (and diameters)	Axonal conduction velocities	Location	Function	Rate of adaptation	Threshold of activation
Free nerve endings	Minimally specialized nerve endings	C, A δ	2–20 m/s	All skin	Pain, temperature, crude touch	Slow	High
Meissner's corpuscles	Encapsulated; between dermal papillae	A β 6–12 μm		Principally glabrous skin	Touch, pressure (dynamic)	Rapid	Low
Pacinian corpuscles	Encapsulated; onionlike covering	A β 6–12 μm		Subcutaneous tissue, interosseous membranes, viscera	Deep pressure, vibration (dynamic)	Rapid	Low
Merkel's disks	Encapsulated; associated with peptide-releasing cells	A β		All skin, hair follicles	Touch, pressure (static)	Slow	Low
Ruffini's corpuscles	Encapsulated; oriented along stretch lines	A β 6–12 μm		All skin	Stretching of skin	Slow	Low
Muscle spindles	Highly specialized (see Figure 8.5 and Chapter 15)	Ia and II		Muscles	Muscle length	Both slow and rapid	Low
Golgi tendon organs	Highly specialized (see Chapter 15)	Ib		Tendons	Muscle tension	Slow	Low
Joint receptors	Minimally specialized	—		Joints	Joint position	Rapid	Low

^aIn the 1920s and 1930s, there was a virtual cottage industry classifying axons according to their conduction velocity. Three main categories were discerned, called A, B, and C. A comprises the largest and fastest axons, C the smallest and slowest. Mechanoreceptor axons generally fall into category A. The A group is further broken down into subgroups designated α (the fastest), β , and δ (the slowest). To make matters even more confusing, muscle afferent axons are usually classified into four additional groups—I (the fastest), II, III, and IV (the slowest)—with subgroups designated by lowercase roman letters!

nerve endings, which in turn affects the ionic permeability of the receptor cell membrane. Changes in permeability generate a depolarizing current in the nerve ending, thus producing a **receptor (or generator) potential** that triggers action potentials, as described in Chapters 2 and 3. This overall process, in which the energy of a stimulus is converted into an electrical signal in the sensory neuron, is called **sensory transduction** and is the critical first step in all sensory processing.

The *quality* of a mechanosensory (or any other) stimulus (i.e., what it represents and where it is) is determined by the properties of the relevant receptors and the location of their central targets (Figure 8.1). The quantity or strength of the stimulus is conveyed by the rate of action potential discharge triggered by the receptor potential (although this relationship is nonlinear and often quite complex). Some receptors fire rapidly when a stimulus is first presented and then fall silent in the presence of continued stimulation (which is to say they “adapt” to the stimulus), whereas others generate a sustained discharge in the presence of an ongoing stimulus (Figure 8.2). The usefulness of having some receptors that adapt quickly and others that do not is to provide information about both the *dynamic* and *static* qualities of a stimulus. Receptors that initially fire in the presence of a stimulus and then

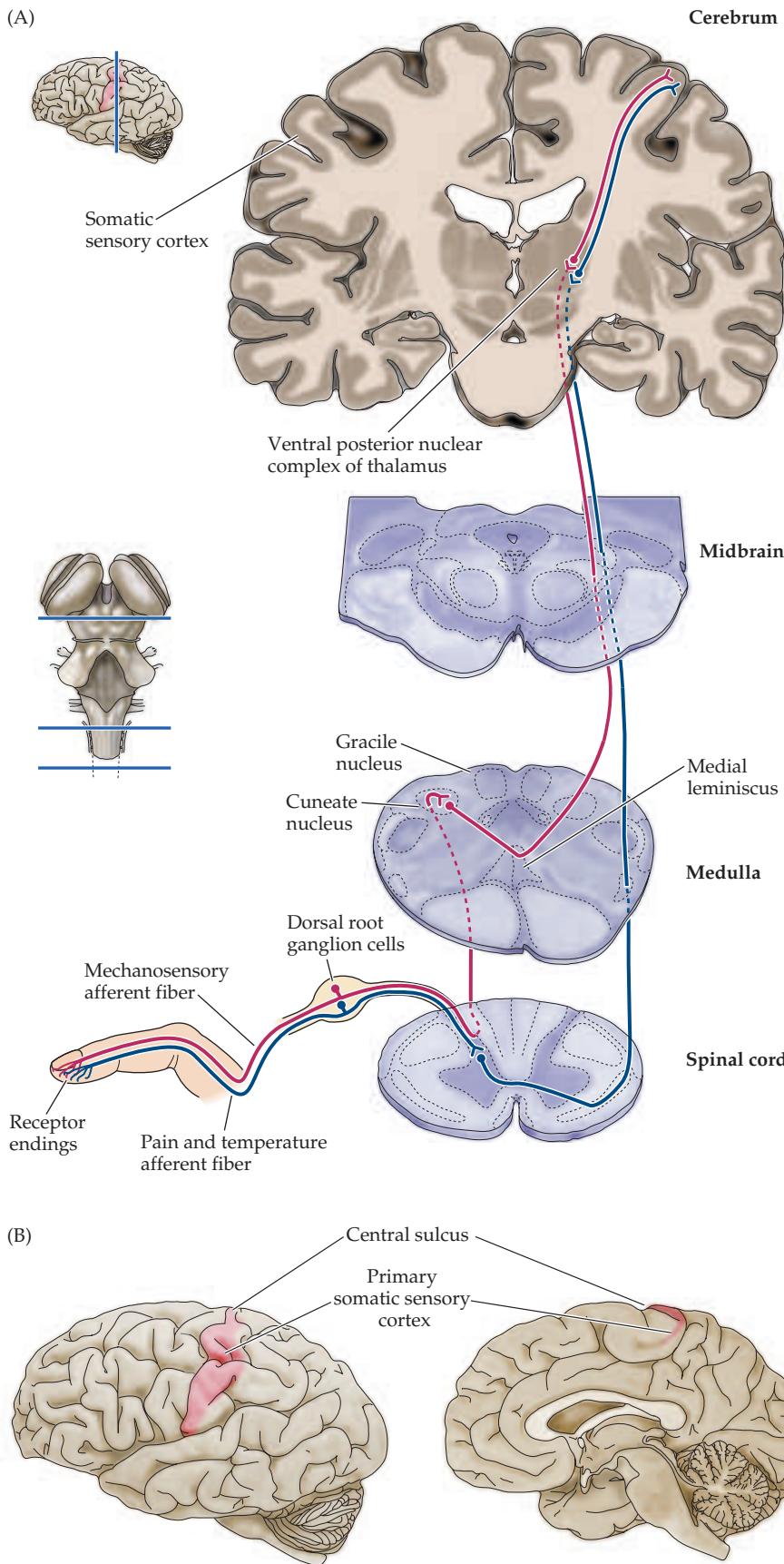


Figure 8.1 General organization of the somatic sensory system. (A) Mechanosensory information about the body reaches the brain by way of a three-neuron relay (shown in red). The first synapse is made by the terminals of the centrally projecting axons of dorsal root ganglion cells onto neurons in the brainstem nuclei (the local branches involved in segmental spinal reflexes are not shown here). The axons of these second-order neurons synapse on third-order neurons of the ventral posterior nuclear complex of the thalamus, which in turn send their axons to the primary somatic sensory cortex (red). Information about pain and temperature takes a different course (shown in blue; the anterolateral system), and is discussed in the following chapter. (B) Lateral and midsagittal views of the human brain, illustrating the approximate location of the primary somatic sensory cortex in the anterior parietal lobe, just posterior to the central sulcus.

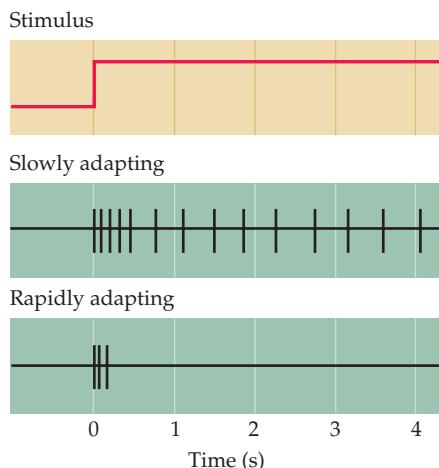


Figure 8.2 Slowly adapting mechanoreceptors continue responding to a stimulus, whereas rapidly adapting receptors respond only at the onset (and often the offset) of stimulation. These functional differences allow the mechanoreceptors to provide information about both the static (via slowly adapting receptors) and dynamic (via rapidly adapting receptors) qualities of a stimulus.

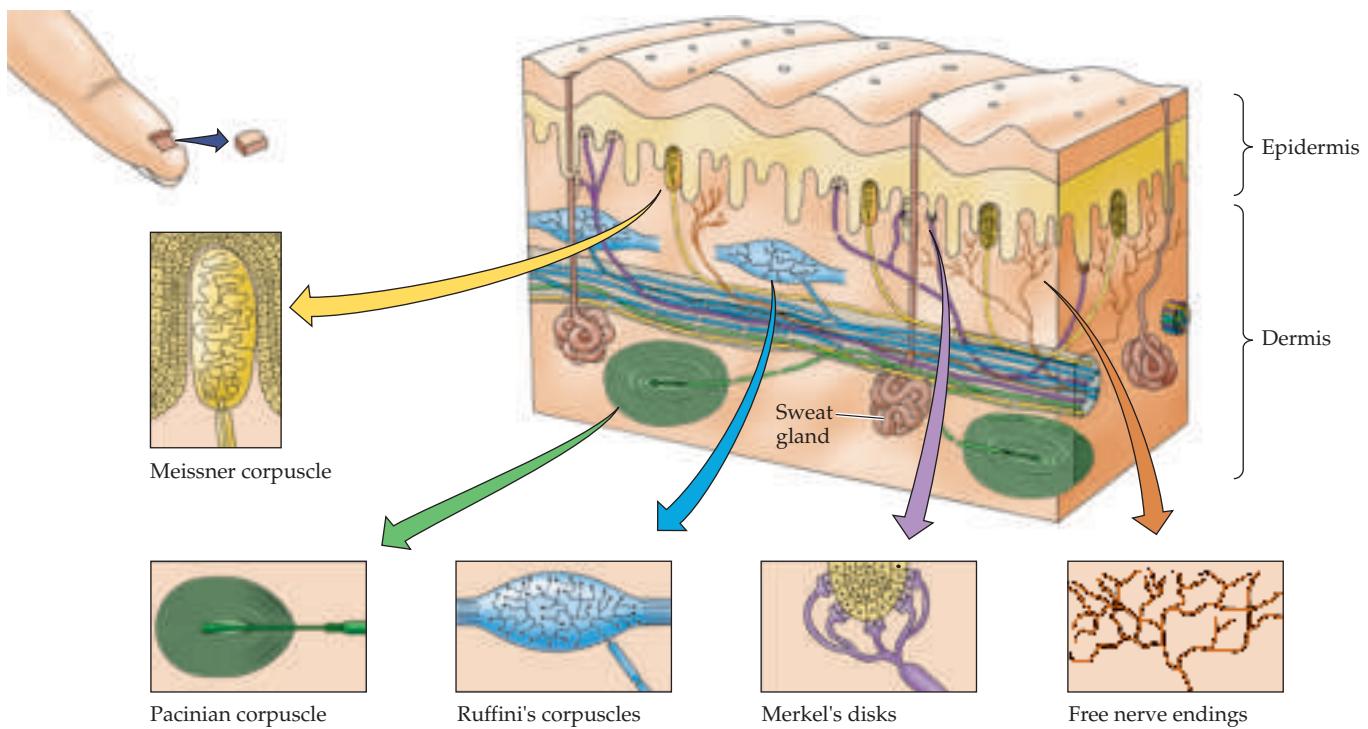
become quiescent are particularly effective in conveying information about changes in the information the receptor reports; conversely, receptors that continue to fire convey information about the persistence of a stimulus. Accordingly, somatic sensory receptors and the neurons that give rise to them are usually classified into rapidly or slowly adapting types (see Table 8.1). **Rapidly adapting**, or **phasic**, receptors respond maximally but briefly to stimuli; their response decreases if the stimulus is maintained. Conversely, **slowly adapting**, or **tonic**, receptors keep firing as long as the stimulus is present.

Mechanoreceptors Specialized to Receive Tactile Information

Four major types of encapsulated mechanoreceptors are specialized to provide information to the central nervous system about touch, pressure, vibration, and cutaneous tension: Meissner's corpuscles, Pacinian corpuscles, Merkel's disks, and Ruffini's corpuscles (Figure 8.3 and Table 8.1). These receptors are referred to collectively as **low-threshold** (or high-sensitivity) mechanoreceptors because even weak mechanical stimulation of the skin induces them to produce action potentials. All low-threshold mechanoreceptors are innervated by relatively large myelinated axons (type A β ; see Table 8.1), ensuring the rapid central transmission of tactile information.

Meissner's corpuscles, which lie between the dermal papillae just beneath the epidermis of the fingers, palms, and soles, are elongated receptors formed by a connective tissue capsule that comprises several lamellae of Schwann cells. The center of the capsule contains one or more afferent nerve fibers that generate rapidly adapting action potentials following minimal skin depression. Meissner's corpuscles are the most common mechanoreceptors of "glabrous" (smooth, hairless) skin (the fingertips, for instance), and their afferent fibers account for about 40% of the sensory innervation of the human hand. These corpuscles are particularly efficient in transducing information about the relatively low-frequency vibrations (30–50 Hz) that occur when textured objects are moved across the skin.

Pacinian corpuscles are large encapsulated endings located in the subcutaneous tissue (and more deeply in interosseous membranes and mesenteries of the gut). These receptors differ from Meissner's corpuscles in their morphology, distribution, and response threshold. The Pacinian corpuscle has an onion-like capsule in which the inner core of membrane lamellae is separated from an outer lamella by a fluid-filled space. One or more rapidly adapting afferent axons lie at the center of this structure. The capsule again acts as a filter, in this case allowing only transient disturbances at high frequencies (250–350 Hz) to activate the nerve endings. Pacinian corpuscles adapt more rapidly than Meissner's corpuscles and have a lower response threshold. These attributes suggest that Pacinian corpuscles are involved in the discrimination of fine surface textures or other moving stimuli that produce high-frequency vibration of the skin. In corroboration of this supposition, stimulation of Pacinian corpuscle afferent fibers in humans induces a sensation of vibration or tickle. They make up 10–15% of the cutaneous receptors in the hand. Pacinian corpuscles located in interosseous membranes probably detect vibrations transmitted to the skeleton. Structurally similar endings found in the bills of ducks and geese and in the legs of cranes and herons detect vibrations in water; such endings in the wings of soaring birds detect vibrations produced by air currents. Because they are rapidly adapting, Pacinian corpuscles, like Meissner's corpuscles, provide information primarily about the dynamic qualities of mechanical stimuli.



Slowly adapting cutaneous mechanoreceptors include **Merkel's disks** and **Ruffini's corpuscles** (see Figure 8.3 and Table 8.1). Merkel's disks are located in the epidermis, where they are precisely aligned with the papillae that lie beneath the dermal ridges. They account for about 25% of the mechanoreceptors of the hand and are particularly dense in the fingertips, lips, and external genitalia. The slowly adapting nerve fiber associated with each Merkel's disk enlarges into a saucer-shaped ending that is closely applied to another specialized cell containing vesicles that apparently release peptides that modulate the nerve terminal. Selective stimulation of these receptors in humans produces a sensation of light pressure. These several properties have led to the supposition that Merkel's disks play a major role in the static discrimination of shapes, edges, and rough textures.

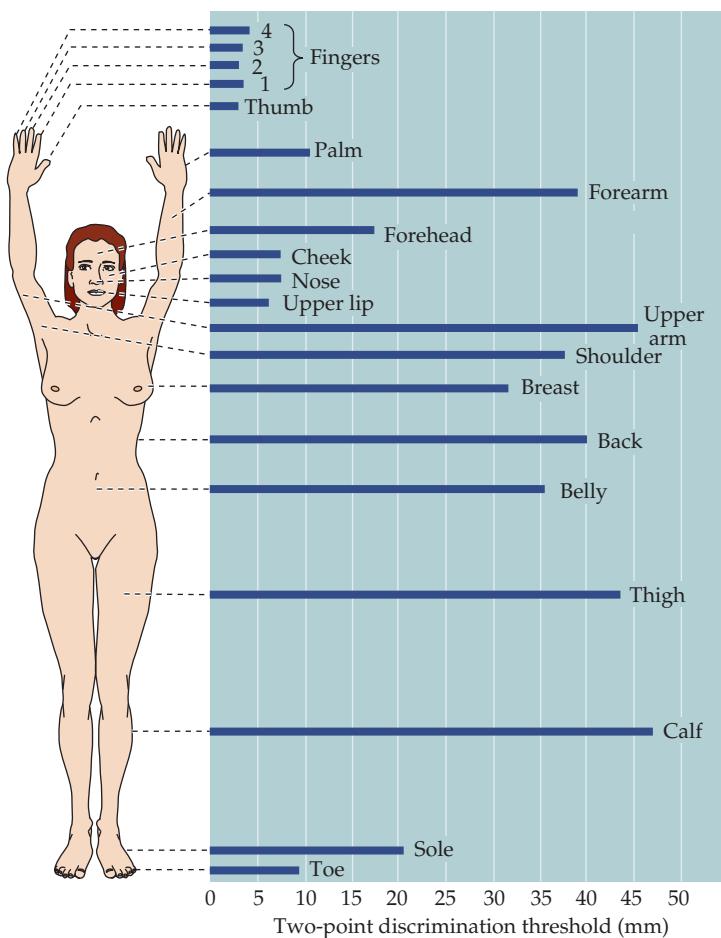
Ruffini's corpuscles, although structurally similar to other tactile receptors, are not well understood. These elongated, spindle-shaped capsular specializations are located deep in the skin, as well as in ligaments and tendons. The long axis of the corpuscle is usually oriented parallel to the stretch lines in skin; thus, Ruffini's corpuscles are particularly sensitive to the cutaneous stretching produced by digit or limb movements. They account for about 20% of the receptors in the human hand and do not elicit any particular tactile sensation when stimulated electrically. Although there is still some question as to their function, they probably respond primarily to internally generated stimuli (see the section on proprioception, below).

Differences in Mechanosensory Discrimination across the Body Surface

The accuracy with which tactile stimuli can be sensed varies from one region of the body to another, a phenomenon that illustrates some further principles

Figure 8.3 The skin harbors a variety of morphologically distinct mechanoreceptors. This diagram represents the smooth, hairless (also called glabrous) skin of the fingertip. The major characteristics of the various receptor types are summarized in Table 8.1. (After Darian-Smith, 1984.)

Figure 8.4 Variation in the sensitivity of tactile discrimination as a function of location on the body surface, measured here by two-point discrimination. (After Weinstein, 1968.)



of somatic sensation. Figure 8.4 shows the results of an experiment in which variation in tactile ability across the body surface was measured by **two-point discrimination**. This technique measures the minimal interstimulus distance required to perceive two simultaneously applied stimuli as distinct (the indentations of the points of a pair of calipers, for example). When applied to the skin, such stimuli of the fingertips are discretely perceived if they are only 2 mm apart. In contrast, the same stimuli applied to the forearm are not perceived as distinct until they are at least 40 mm apart! This marked regional difference in tactile ability is explained by the fact that the encapsulated mechanoreceptors that respond to the stimuli are three to four times more numerous in the fingertips than in other areas of the hand, and many times more dense than in the forearm. Equally important in this regional difference are the sizes of the neuronal receptive fields. The **receptive** field of a somatic sensory neuron is the region of the skin within which a tactile stimulus evokes a sensory response in the cell or its axon (Boxes A and B). Analysis of the human hand shows that the receptive fields of mechanosensory neurons are 1–2 mm in diameter on the fingertips but 5–10 mm on the palms. The receptive fields on the arm are larger still. The importance of receptive field size is easy to envision. If, for instance, the receptive fields of all cutaneous receptor neurons covered the entire digital pad, it would be impossible to discriminate two spatially separate stimuli applied to the fingertip (since all the receptive fields would be returning the same spatial information).

Box A

Receptive Fields and Sensory Maps in the Cricket

Two principles of somatosensory organization have emerged from studies of the mammalian brain: (1) individual neurons are tuned to particular aspects of complex stimuli; and (2) these stimulus qualities are represented in an orderly fashion in relevant regions of the nervous system. These principles apply equally well to invertebrates, including the equivalent of the somatic sensory system in insects such as crickets, grasshoppers, and cockroaches.

In the cricket, the salient tactile stimulation for the animal comes from air currents that displace sensory hairs of bilaterally symmetric sensory structures called cerci (sing. *cercus*). The location and structure of specific cercal hairs

allow them to be displaced by air currents having different directions and speeds (Figure A). Accordingly, the peripheral sensory neurons associated with the hairs represent the full range of air current directions and velocities impinging on the animal. This information is carried centrally and is systematically represented in a region of the cricket central nervous system called the terminal ganglion.

Individual neurons in this ganglion correspond to the cercal hairs, and have receptive fields and response properties that represent a full range of directions and speeds for extrinsic mechanical forces, including air currents (Figure B). For the cricket, the significance of this

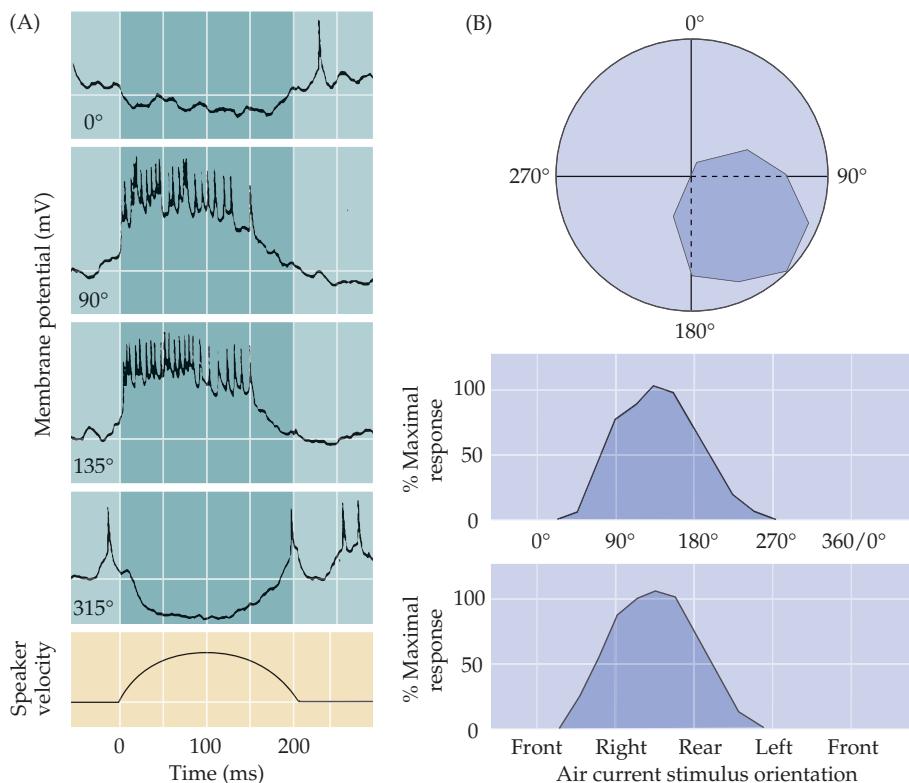
information is, among other things, detecting the direction and speed of oncoming objects to then execute motor programs for escape. (This is also the likely significance of this representation for cockroaches, which can therefore escape the consequences of a descending human foot.)

Much like the somatic sensory system in mammals, the primary sensory afferents project to the terminal ganglion in an orderly fashion, such that there is a somatotopic map of air current directions. And, like mammals, individual neurons within this representation are tuned to specific aspects of the mechanical forces acting on the cricket.

These facts about insects' mechanosensory system emphasize that somatic sensory functions are basically similar across a wide range of animals. Indeed, regardless of sensory modality, nervous system organization, or the identity of the organism, it is likely that stimulus specificity will be reflected in receptive fields of individual neurons and there will be orderly mapping of those receptive fields into either a topographic or computational map in the animal's brain.

References

- JACOBS, G. A. AND F. E. THEUNISSEN (1996) Functional organization of a neural map in the cricket cercal sensory system. *J. Neurosci.* 16: 769–784.
- MILLER, J. P., G. A. JACOBS AND F. E. THEUNISSEN (1991) Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary interneurons. *J. Neurophys.* 66: 1680–1688.
- MURPHEY, R. K. (1981) The structure and development of somatotopic map in crickets: The cercal afferent projection. *Dev. Biol.* 88: 236–246.
- MURPHEY, R. K. AND H. V. B. HIRSCH (1982) From cat to cricket: The genesis of response selectivity of interneurons. *Curr. Topics Dev. Biol.* 17: 241–256.

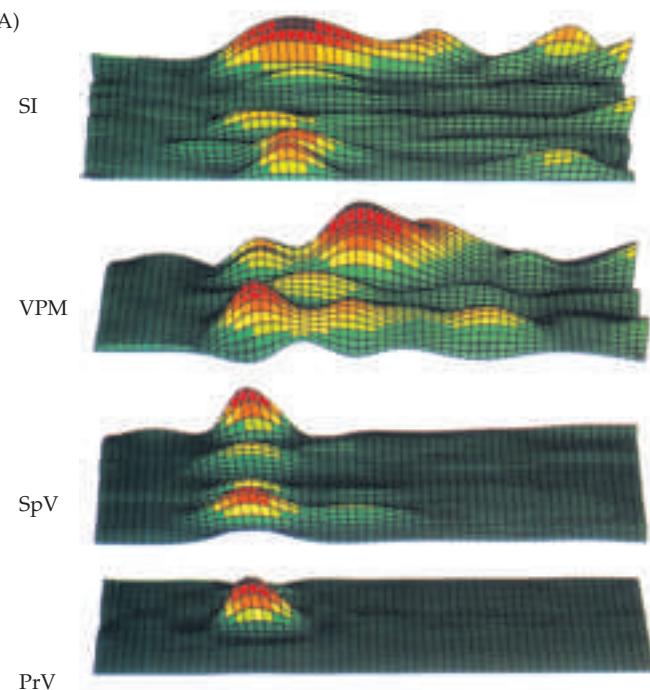


(A) Intracellular recording of action potential activity of an individual sensory neuron's responses to different directions of wind current. (B) The plots indicate this neuron's receptive field for wind direction (top) and the tuning curve for the neuron's selective firing to its preferred direction. (After Miller et al., 1991.)

Box B**Dynamic Aspects of Somatic Sensory Receptive Fields**

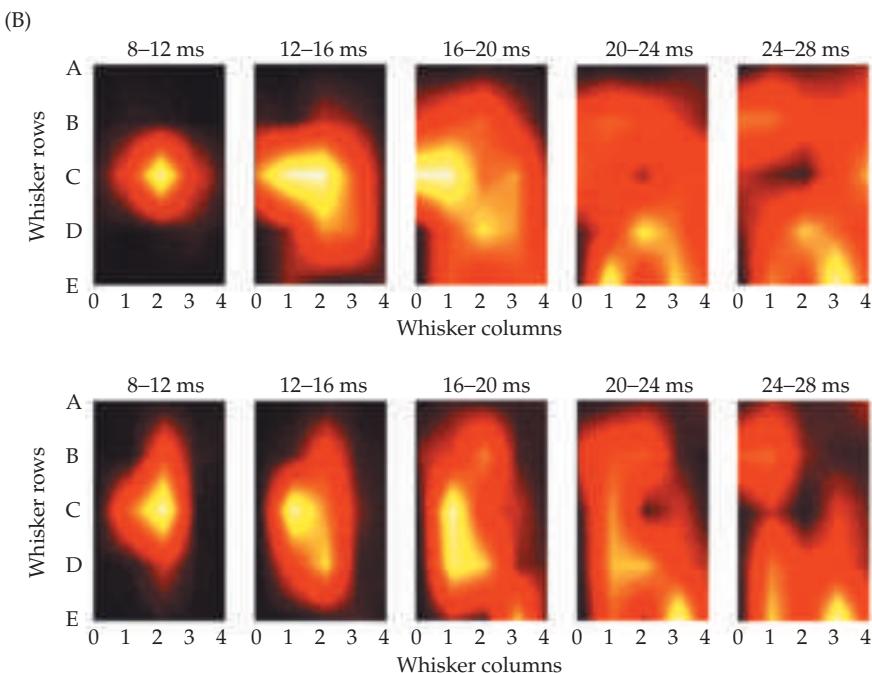
When humans explore objects with their hands, multiple contacts between the skin and the object surface generate extraordinarily complex patterns of tactile stimuli. As a consequence, the somatic sensory system must process signals that change continuously in time. Nonetheless, we routinely discriminate the size, texture and shape of objects with great accuracy. Until recently, the temporal structure of such stimuli was not considered a major variable in characterizing the physiological properties of somatic sensory neurons. For instance, the classical definition of the receptive field of a somatic sensory neuron takes into account only the overall area of the body surface that elicits significant variation in the neuron's firing rate. By the same token, the topographic maps in the somatic sensory system have been interpreted as evidence that tactile information processing involves primarily spatial criteria.

The advent of multiple electrode recording to simultaneously monitor the activity of large populations of single neurons has begun to change this "static" view of the somatic sensory system. In both primates and rodents, this approach has shown that the receptive fields of cortical and subcortical neurons



(A) Simultaneous electrode recordings in behaving rats allow monitoring of the spatiotemporal spread of neuronal activation across several levels of the somatic sensory system following stimulation (of a single facial whisker, in this example). These 3-D graphs represent patterns of neuronal ensemble activity at each level of the pathway. The *x* axis represents the poststimulus time in ms, the *y* axis the number of neurons recorded at each level; the color-coded gradient in the *z* axis shows the response of the neurons, with red the highest firing and green the lowest. SI, somatic sensory cortex; VPM, ventral posterior medial nucleus of the thalamus; SpV, spinal nucleus of the trigeminal brainstem complex; PrV, principal nucleus of the brainstem trigeminal complex. (From Nicholelis et al., 1997.)

Receptor density and receptive field sizes in different regions are not the only factors determining somatic sensation. Psychophysical analysis of tactile performance suggests that something more than the cutaneous periphery is needed to explain variations in tactile perception. For instance, sensory thresholds in two-point discrimination tests vary with practice, fatigue, and stress. The contextual significance of stimuli is also important in determining what we actually feel; even though we spend most of the day wearing clothes, we usually ignore the tactile stimulation that they produce. Some aspect of the mechanosensory system allows us to filter out this information and pay attention to it only when necessary. The fascinating phenomenon of "phantom limb" sensations after amputation (see Box C in Chapter 9) provides further evidence that tactile perception is not fully explained by the



(B) Receptive fields of two cortical neurons from two different animals. Each panel represents the matrix of whiskers on the animals' snout (whisker columns are on the *x* axis and whisker rows on the *y* axis) for a 4-ms epoch of poststimulus time. Within a particular time period, the center of the receptive field is defined as the whisker eliciting the greatest response magnitude (yellow). Note that the receptive field centers shift as a function of time. (From Ghazanfar and Nicolelis, 1998.)

vary as a function of time: The neuron responds differently to a spatially defined stimulus as the period of stimulation proceeds (see Figures A and B).

This coupling of space and time can

also be demonstrated at level of somatotopic maps. By recording the activity of single neurons located in different regions of the map simultaneously, it is apparent that the stimulation of a small

area of the skin tends to excite more and more neurons as time goes by. Thus, many more neurons than those located in the area of the map directly representing the stimulated skin actually respond to the stimulus, albeit at longer latencies.

The end result of these more complex neuronal responses is the emergence of spatiotemporal representations at all levels of the somatic sensory system. Thus, contrary to the classical notion of receptive fields, the somatic sensory system processes information in a dynamic way. Such processing is not only relevant for the normal operation of the system, but may also account for some aspects of adult plasticity (see Chapter 24).

References

- GHAZANFAR, A. A. AND M. A. L. NICOLELIS (1999) Spatiotemporal properties of layer V neurons of the rat primary somatosensory cortex. *Cereb. Cortex* 4: 348–361.
- NICOLELIS, M. A. L., A. A. GHAZANFAR, B. FAGGIN, S. VOTAW AND L. M. O. OLIVEIRA (1997) Reconstructing the engram: Simultaneous, multiple site, many single neuron recordings. *Neuron* 18: 529–537.
- NICOLELIS, M. A. L. AND 7 OTHERS (1998) Simultaneous encoding of tactile information by three primate cortical areas. *Nature Neurosci.* 1: 621–630.

peripheral information that travels centrally. The central nervous system clearly plays an active role in determining the perception of the mechanical forces that act on us.

Mechanoreceptors Specialized for Proprioception

Whereas cutaneous mechanoreceptors provide information derived from external stimuli, another major class of receptors provides information about mechanical forces arising from the body itself, the musculoskeletal system in particular. These are called **proprioceptors**, roughly meaning “receptors for self.” The purpose of proprioceptors is primarily to give detailed and continuous information about the position of the limbs and other body parts in

space (specialized mechanoreceptors also exist in the heart and major vessels to provide information about blood pressure, but these neurons are considered to be part of the visceral motor system; see Chapter 20). Low-threshold mechanoreceptors, including muscle spindles, Golgi tendon organs, and joint receptors, provide this kind of sensory information, which is essential to the accurate performance of complex movements. Information about the position and motion of the head is particularly important; in this case, proprioceptors are integrated with the highly specialized vestibular system, which is considered separately in Chapter 13.

The most detailed knowledge about proprioception derives from studies of **muscle spindles**, which are found in all but a few striated (skeletal) muscles. Muscle spindles consist of four to eight specialized **intrafusal muscle fibers** surrounded by a capsule of connective tissue. The intrafusal fibers are distributed among the ordinary (extrafusal) fibers of skeletal muscle in a parallel arrangement (Figure 8.5). In the largest of the several intrafusal fibers, the nuclei are collected in an expanded region in the center of the fiber called a bag; hence the name *nuclear bag fibers*. The nuclei in the remaining two to six smaller intrafusal fibers are lined up single file, with the result that these fibers are called *nuclear chain fibers*. Myelinated sensory axons belonging to group Ia innervate muscle spindles by encircling the middle portion of both types of intrafusal fibers (see Figure 8.5 and Table 8.1). The Ia axon terminal is known as the **primary sensory ending** of the spindle. Secondary innervation is provided by group II axons that innervate the nuclear chain fibers and give off a minor branch to the nuclear bag fibers. The intrafusal muscle fibers contract when commanded to do so by motor axons derived from a pool of specialized motor neurons in the spinal cord (called **γ motor neurons**). The major function of muscle spindles is to provide information about muscle length (that is, the degree to which they are being stretched). A detailed account of how these important receptors function during movement is given in Chapters 15 and 16.

The density of spindles in human muscles varies. Large muscles that generate coarse movements have relatively few spindles; in contrast, extraocular muscles and the intrinsic muscles of the hand and neck are richly supplied with spindles, reflecting the importance of accurate eye movements, the need to manipulate objects with great finesse, and the continuous demand for precise positioning of the head. This relationship between receptor den-

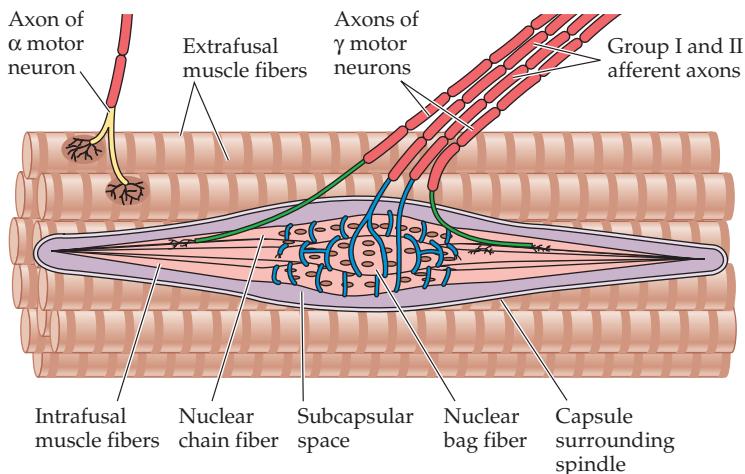


Figure 8.5 A muscle spindle and several extrafusal muscle fibers. See text for description. (After Matthews, 1964.)

sity and muscle size is consistent with the generalization that the sensory motor apparatus at all levels of the nervous system is much richer for the hands, head, speech organs, and other parts of the body that are used to perform especially important and demanding tasks. Spindles are lacking altogether in a few muscles, such as those of the middle ear, which do not require the kind of feedback that these receptors provide.

Whereas muscle spindles are specialized to signal changes in muscle *length*, low-threshold mechanoreceptors in tendons inform the central nervous system about changes in muscle *tension*. These mechanoreceptors, called **Golgi tendon organs**, are innervated by branches of group Ib afferents and are distributed among the collagen fibers that form the tendons (see Chapter 15).

Finally, rapidly adapting mechanoreceptors in and around joints gather dynamic information about limb position and joint movement. The function of these **joint receptors** is not well understood.

Active Tactile Exploration

Tactile discrimination—that is, perceiving the detailed shape or texture of an object—normally entails active exploration. In humans, this is typically accomplished by using the hands to grasp and manipulate objects, or by moving the fingers across a surface so that a sequence of contacts between the skin and the object of interest is established. Psychophysical evidence indicates that relative movement between the skin and a surface is the single most important requirement for accurate discrimination of texture. Animal experiments confirm the dependence of tactile discrimination on active exploration. Rats, for instance, discriminate the details of texture by rhythmically brushing their facial whiskers across surfaces. Active touching, which is called **haptics**, involves the interpretation of complex spatiotemporal patterns of stimuli that are likely to activate many classes of mechanoreceptors. Haptics also requires dynamic interactions between motor and sensory signals, which presumably induce sensory responses in central neurons that differ from the responses of the same cells during passive stimulation of the skin (see Box B).

The Major Afferent Pathway for Mechanosensory Information: The Dorsal Column–Medial Lemniscus System

The action potentials generated by tactile and other mechanosensory stimuli are transmitted to the spinal cord by afferent sensory axons traveling in the peripheral nerves. The neuronal cell bodies that give rise to these first-order axons are located in the **dorsal root (or sensory) ganglia** associated with each segmental spinal nerve (see Figure 8.1 and Box C). Dorsal root ganglion cells are also known as **first-order neurons** because they initiate the sensory process. The ganglion cells thus give rise to long peripheral axons that end in the somatic receptor specializations already described, and shorter central axons that reach the dorsolateral region of the spinal cord via the **dorsal (sensory) roots** of each spinal cord segment. The large myelinated fibers that innervate low-threshold mechanoreceptors are derived from the largest neurons in these ganglia, whereas the smaller ganglion cells give rise to smaller afferent nerve fibers that end in the high-threshold nociceptors and thermoreceptors (see Table 8.1).

Depending on whether they belong to the mechanosensory system or to the pain and temperature system, the first-order axons carrying information

from somatic receptors have different patterns of termination in the spinal cord and define distinct somatic sensory pathways within the central nervous system (see Figure 8.1). The **dorsal column-medial lemniscus pathway** carries the majority of information from the mechanoreceptors that mediate tactile discrimination and proprioception (Figure 8.6); the **spinothalamic (anterolateral) pathway** mediates pain and temperature sensation and is described in Chapter 9. This difference in the afferent pathways of these modalities is one of the reasons that pain and temperature sensation is treated separately here.

Upon entering the spinal cord, the first-order axons carrying information from peripheral mechanoreceptors bifurcate into ascending and descending branches, which in turn send collateral branches to several spinal segments. Some collateral branches penetrate the dorsal horn of the cord and synapse on neurons located mainly in a region called Rexed's laminae III–V. These synapses mediate, among other things, segmental reflexes such as the "knee-jerk" or myotatic reflex described in Chapter 1, and are further considered in Chapters 15 and 16. The major branch of the incoming axons, however, ascends ipsilaterally through the **dorsal columns** (also called the *posterior funiculi*) of the cord, all the way to the lower medulla, where it terminates by contacting **second-order neurons** in the **gracile** and **cuneate nuclei** (together referred to as the **dorsal column nuclei**; see Figures 8.1 and 8.6A). Axons in the dorsal columns are topographically organized such that the fibers that convey information from lower limbs are in the medial subdivision of the dorsal columns, called the **gracile tract**, a fact of some significance in the clinical localization of neural injury. The lateral subdivision, called the **cuneate tract**, contains axons conveying information from the upper limbs, trunk, and neck. At the level of the upper thorax, the dorsal columns account for more than a third of the cross-sectional area of the human spinal cord.

Despite their size, lesions limited to the dorsal columns of the spinal cord in both humans and monkeys have only a modest effect on the performance of simple tactile tasks. Such lesions, however, do impede the ability to detect the direction and speed of tactile stimuli, as well as degrading the ability to sense the position of the limbs in space. Dorsal column lesions may also reduce a patient's ability to initiate active movements related to tactile exploration. For instance, such individuals have difficulty recognizing numbers and letters drawn on their skin. The relatively mild deficit that follows dorsal column lesions is presumably explained by the fact that some axons responsible for cutaneous mechanoreception also run in the spinothalamic (pain and temperature) pathway, as described in Chapter 9.

The second-order relay neurons in the dorsal column nuclei send their axons to the somatic sensory portion of the thalamus (see Figure 8.6A). The axons from dorsal column nuclei project in the dorsal portion of each side of the lower brainstem, where they form the **internal arcuate tract**. The internal arcuate axons subsequently cross the midline to form another named tract that is elongated dorsoventrally, the **medial lemniscus**. (The crossing of these fibers is called the *decussation* of the medial lemniscus, from the roman numeral "X," or *decem*; the word *lemniscus* means "ribbon.")

In a cross-section through the medulla, such as the one shown in Figure 8.6A, the medial lemniscal axons carrying information from the lower limbs are located ventrally, whereas the axons related to the upper limbs are located dorsally (again, a fact of some clinical importance). As the medial lemniscus ascends through the pons and midbrain, it rotates 90° laterally, so that the upper body is eventually represented in the medial portion of the tract, and the lower body in the lateral portion. The axons of the medial lem-

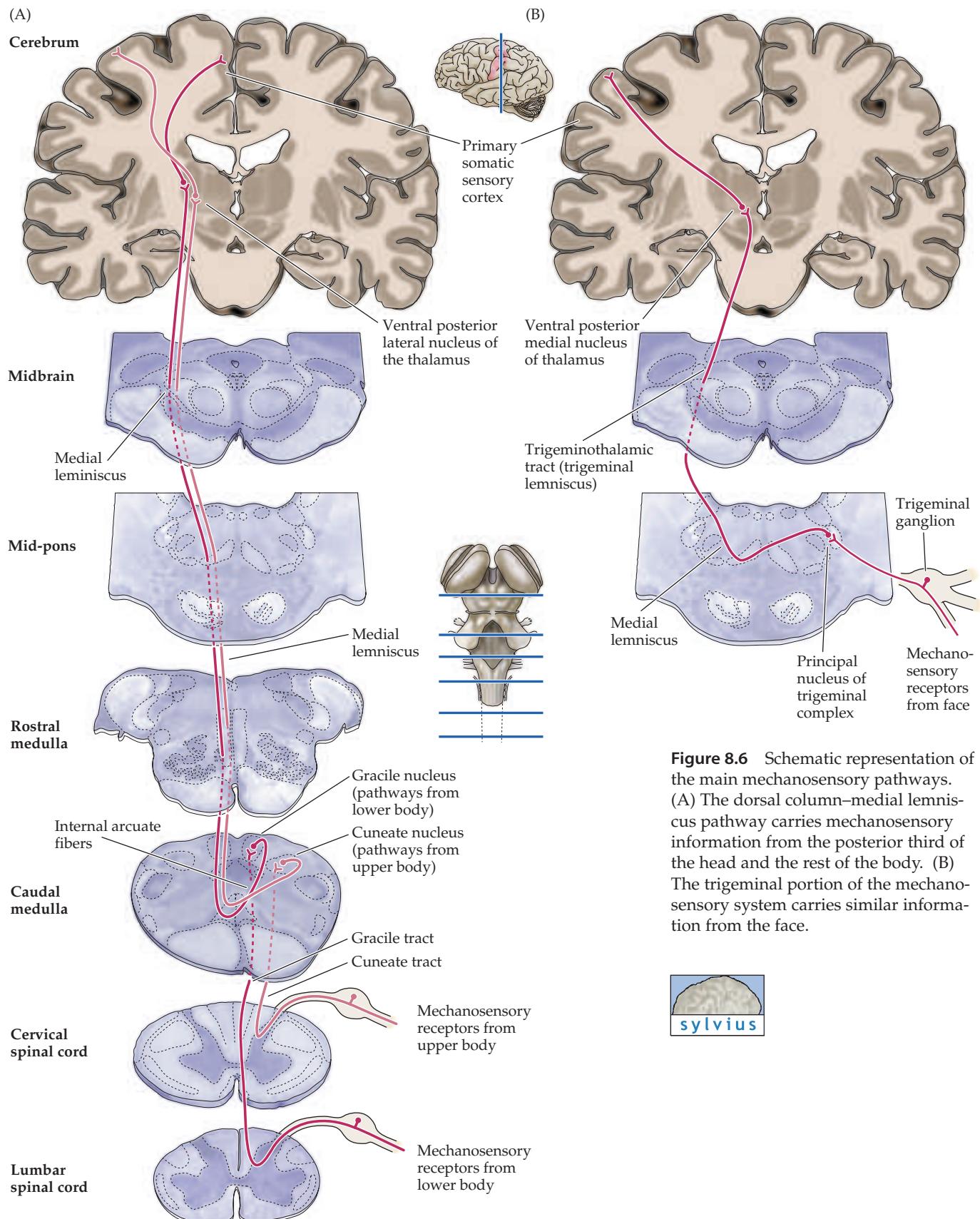


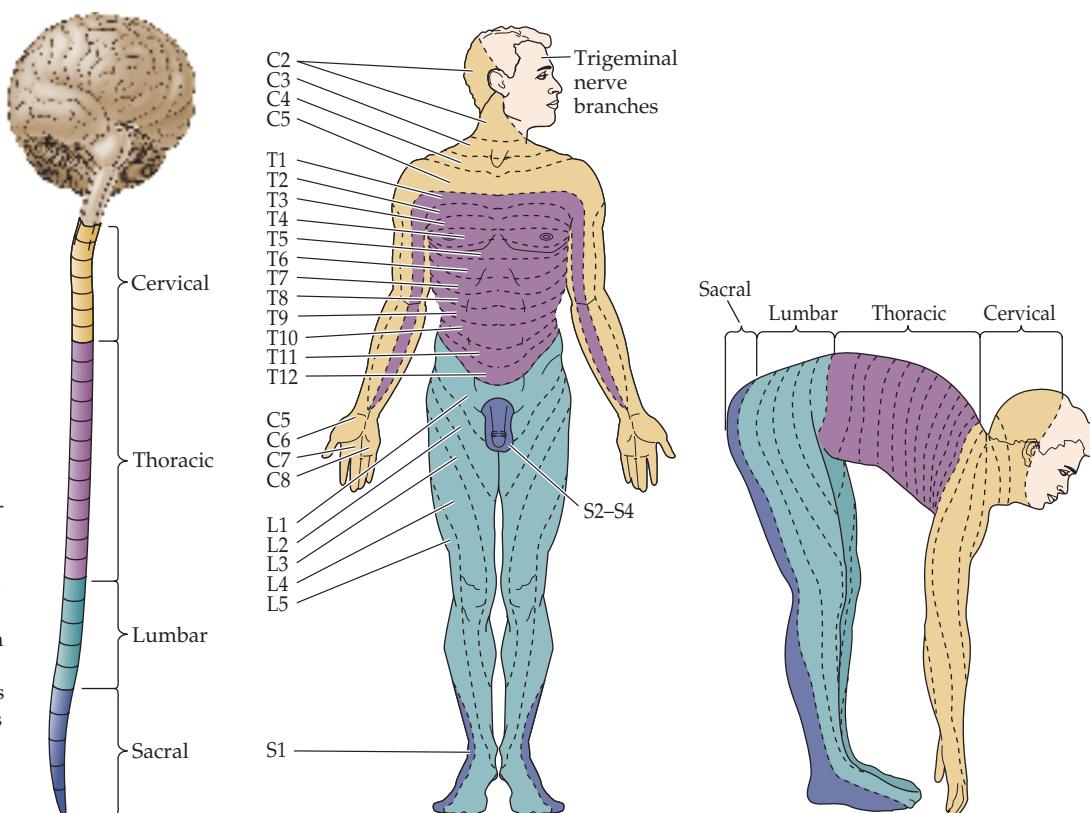
Figure 8.6 Schematic representation of the main mechanosensory pathways. (A) The dorsal column-medial lemniscus pathway carries mechanosensory information from the posterior third of the head and the rest of the body. (B) The trigeminal portion of the mechanosensory system carries similar information from the face.



Box C

Dermatomes

The innervation arising from a single dorsal root ganglion and its spinal nerve is called a dermatome. The full set of sensory dermatomes is shown here for a typical adult. Knowledge of this arrangement is particularly important in defining the location of suspected spinal (and other) lesions. The numbers refer to the spinal segments by which each nerve is named. (After Rosenzweig et al., 2002.)



Each dorsal root (or sensory) ganglion and associated spinal nerve arises from an iterated series of embryonic tissue masses called somites. This fact of development explains the overall segmental arrangement of somatic nerves (and the targets they innervate) in the adult (see figure). The territory innervated by each spinal nerve is called a dermatome. In humans, the cutaneous area of each dermatome has been defined in patients in whom specific dorsal roots were affected

(as in herpes zoster, or “shingles”) or after surgical interruption (for relief of pain or other reasons). Such studies show that dermatomal maps vary among individuals. Moreover, dermatomes overlap substantially, so that injury to an individual dorsal root does not lead to complete loss of sensation in the relevant skin region, the overlap being more extensive for touch, pressure, and vibration than for pain and temperature. Thus, testing for pain sensation provides

a more precise assessment of a segmental nerve injury than does testing for responses to touch, pressure, or vibration. Finally, the segmental distribution of proprioceptors does not follow the dermatomal map but is more closely allied with the pattern of muscle innervation. Despite these limitations, knowledge of dermatomes is essential in the clinical evaluation of neurological patients, particularly in determining the level of a spinal lesion.

niscus thus reach the ventral posterior lateral (VPL) nucleus of the thalamus, whose cells are the **third-order neurons** of the dorsal column-medial lemniscus system (see Figure 8.7).

The Trigeminal Portion of the Mechanosensory System

As noted, the dorsal column-medial lemniscus pathway described in the preceding section carries somatic information from only the upper and lower body and from the posterior third of the head. Tactile and propriocep-

tive information from the face is conveyed from the periphery to the thalamus by a different route. Information derived from the face is transmitted to the central nervous system via the **trigeminal somatic sensory system** (Figure 8.6B). Low-threshold mechanoreception in the face is mediated by first-order neurons in the trigeminal (cranial nerve V) ganglion. The peripheral processes of these neurons form the three main subdivisions of the **trigeminal nerve** (the **ophthalmic**, **maxillary**, and **mandibular branches**), each of which innervates a well-defined territory on the face and head, including the teeth and the mucosa of the oral and nasal cavities. The central processes of trigeminal ganglion cells form the sensory roots of the trigeminal nerve; they enter the brainstem at the level of the pons to terminate on neurons in the subdivisions of the **trigeminal brainstem complex**.

The trigeminal complex has two major components: the **principal nucleus** (responsible for processing mechanosensory stimuli), and the **spinal nucleus** (responsible for processing painful and thermal stimuli). Thus, most of the axons carrying information from low-threshold cutaneous mechanoreceptors in the face terminate in the principal nucleus. In effect, this nucleus corresponds to the dorsal column nuclei that relay mechanosensory information from the rest of the body. The spinal nucleus corresponds to a portion of the spinal cord that contains the second-order neurons in the pain and temperature system for the rest of the body (see Chapter 9). The second-order neurons of the trigeminal brainstem nuclei give off axons that cross the midline and ascend to the ventral posterior medial (VPM) nucleus of the thalamus by way of the **trigeminothalamic tract** (also called the trigeminal lemniscus).

The Somatic Sensory Components of the Thalamus

Each of the several ascending somatic sensory pathways originating in the spinal cord and brainstem converge on the thalamus (Figure 8.7). The **ventral posterior complex** of the thalamus, which comprises a lateral and a medial nucleus, is the main target of these ascending pathways. As already mentioned, the more laterally located **ventral posterior lateral (VPL) nucleus** receives projections from the medial lemniscus carrying all somatosensory information from the body and posterior head, whereas the more medially located **ventral posterior medial (VPM) nucleus** receives axons from the trigeminal lemniscus (that is, mechanosensory and nociceptive information from the face). Accordingly, the ventral posterior complex of the thalamus contains a complete representation of the somatic sensory periphery.

The Somatic Sensory Cortex

The axons arising from neurons in the ventral posterior complex of the thalamus project to cortical neurons located primarily in layer IV of the somatic sensory cortex (see Figure 8.7; also see Box A in Chapter 25 for a more detailed description of cortical lamination). The **primary somatic sensory cortex** in humans (also called **SI**), which is located in the postcentral gyrus of the parietal lobe, comprises four distinct regions, or fields, known as **Brodmann's areas 3a, 3b, 1, and 2**. Experiments carried out in nonhuman primates indicate that neurons in areas 3b and 1 respond primarily to cutaneous stimuli, whereas neurons in 3a respond mainly to stimulation of proprioceptors; area 2 neurons process both tactile and proprioceptive stimuli. Mapping studies in humans and other primates show further that each

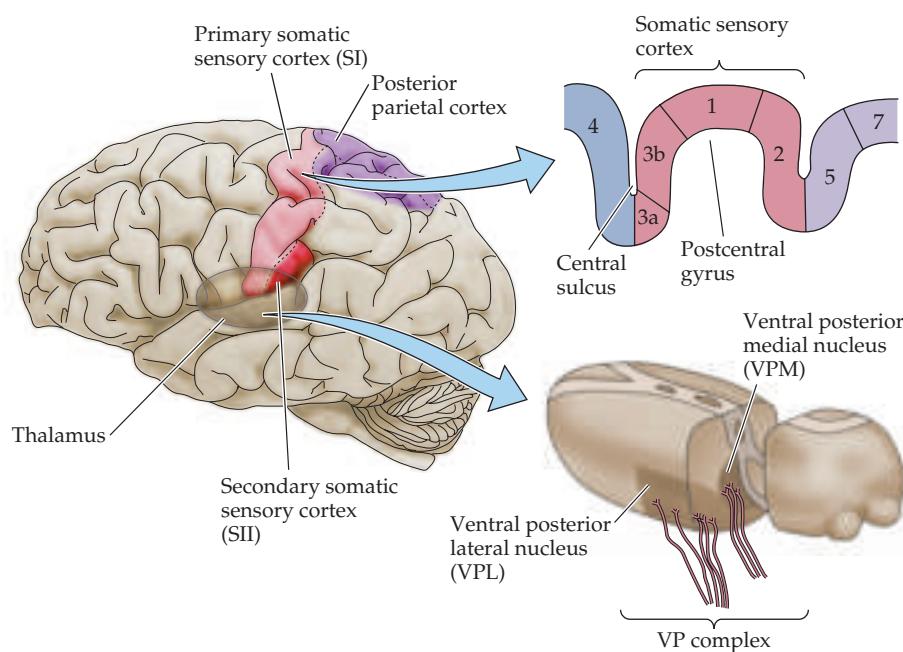


Figure 8.7 Diagram of the somatic sensory portions of the thalamus and their cortical targets in the postcentral gyrus. The ventral posterior nuclear complex comprises the VPM, which relays somatic sensory information carried by the trigeminal system from the face, and the VPL, which relays somatic sensory information from the rest of the body. Inset above shows organization of the primary somatosensory cortex in the postcentral gyrus, shown here in a section cutting across the gyrus from anterior to posterior. (After Brodal, 1992, and Jones et al., 1982.)

of these four cortical areas contains a separate and complete representation of the body. In these **somatotopic maps**, the foot, leg, trunk, forelimbs, and face are represented in a medial to lateral arrangement, as shown in Figures 8.8A,B and 8.9.

Although the topographic organization of the several somatic sensory areas is similar, the functional properties of the neurons in each region and their organization are distinct (Box D). For instance, the neuronal receptive fields are relatively simple in area 3b; the responses elicited in this region are generally to stimulation of a single finger. In areas 1 and 2, however, the majority of the receptive fields respond to stimulation of multiple fingers. Furthermore, neurons in area 1 respond preferentially to particular directions of skin stimulation, whereas many area 2 neurons require complex stimuli to activate them (such as a particular shape). Lesions restricted to area 3b produce a severe deficit in both texture and shape discrimination. In contrast, damage confined to area 1 affects the ability of monkeys to perform accurate texture discrimination. Area 2 lesions tend to produce deficits in finger coordination, and in shape and size discrimination.

A salient feature of cortical maps, recognized soon after their discovery, is their failure to represent the body in actual proportion. When neurosurgeons determined the representation of the human body in the primary sensory (and motor) cortex, the homunculus (literally, “little man”) defined by such mapping procedures had a grossly enlarged face and hands compared to the torso and proximal limbs (Figure 8.8C). These anomalies arise because

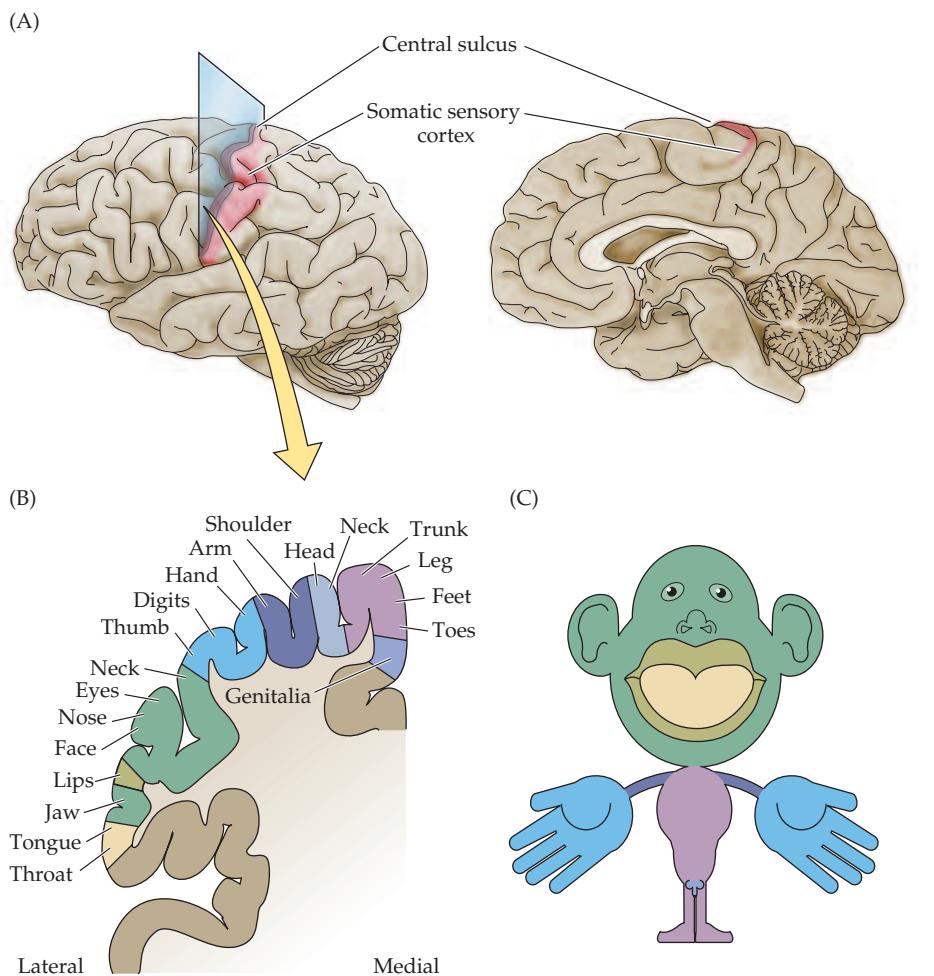


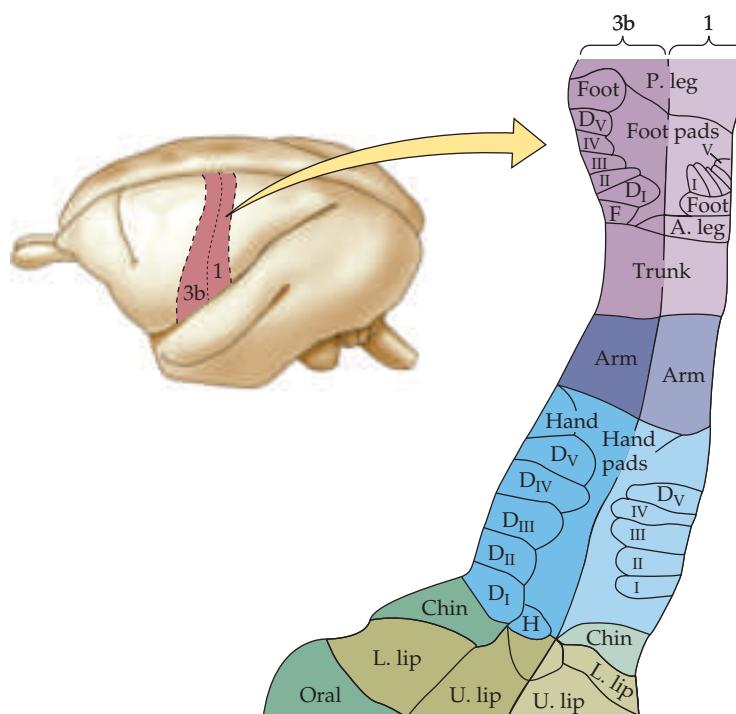
Figure 8.8 Somatotopic order in the human primary somatic sensory cortex. (A) Diagram showing the region of the human cortex from which electrical activity is recorded following mechanosensory stimulation of different parts of the body. The patients in the study were undergoing neurosurgical procedures for which such mapping was required. Although modern imaging methods are now refining these classical data, the human somatotopic map first defined in the 1930s has remained generally valid. (B) Diagram along the plane in (A) showing the somatotopic representation of body parts from medial to lateral. (C) Cartoon of the homunculus constructed on the basis of such mapping. Note that the amount of somatic sensory cortex devoted to the hands and face is much larger than the relative amount of body surface in these regions. A similar disproportion is apparent in the primary motor cortex, for much the same reasons (see Chapter 17). (After Penfield and Rasmussen, 1950, and Corsi, 1991.)



manipulation, facial expression, and speaking are extraordinarily important for humans, requiring more central (and peripheral) circuitry to govern them. Thus, in humans, the cervical spinal cord is enlarged to accommodate the extra circuitry related to the hand and upper limb, and as stated earlier, the density of receptors is greater in regions such as the hands and lips. Such distortions are also apparent when topographical maps are compared across species. In the rat brain, for example, an inordinate amount of the somatic sensory cortex is devoted to representing the large facial whiskers that pro-

Figure 8.9 The primary somatic sensory map in the owl monkey based, as in Figure 8.8, on the electrical responsiveness of the cortex to peripheral stimulation.

Much more detailed mapping is possible in experimental animals than in neuro-surgical patients. The enlargement on the right shows areas 3b and 1, which process most cutaneous mechanosensory information. The arrangement is generally similar to that determined in humans. (After Kaas, 1983.)



vide a key component of the somatic sensory input for rats and mice (see Boxes B and D), while raccoons overrepresent their paws and the platypus its bill. In short, the sensory input (or motor output) that is particularly significant to a given species gets relatively more cortical representation.

Higher-Order Cortical Representations

Somatic sensory information is distributed from the primary somatic sensory cortex to “higher-order” cortical fields (as well as to subcortical structures). One of these higher-order cortical centers, the secondary somatosensory cortex (sometimes called SII and adjacent to the primary cortex; see Figure 8.7), receives convergent projections from the primary somatic sensory cortex and sends projections in turn to limbic structures such as the amygdala and hippocampus (see Chapters 28 and 30). This latter pathway is believed to play an important role in tactile learning and memory. Neurons in motor cortical areas in the frontal lobe also receive tactile information from the anterior parietal cortex and, in turn, provide feedback projections to several cortical somatic sensory regions. Such integration of sensory and motor information is considered in Chapters 19 and 25, where the role of these “association” regions of the cerebral cortex are discussed in more detail.

Finally, a fundamental but often neglected feature of the somatic sensory system is the presence of massive descending projections. These pathways originate in sensory cortical fields and run to the thalamus, brainstem, and spinal cord. Indeed, descending projections from the somatic sensory cortex outnumber ascending somatic sensory pathways! Although their physiological role is not well understood, it is generally assumed (with some experimental support) that descending projections modulate the ascending flow of sensory information at the level of the thalamus and brainstem.

Box D

Patterns of Organization within the Sensory Cortices: Brain Modules

Observations over the last 40 years have made it clear that there is an iterated substructure within the somatic sensory (and many other) cortical maps. This substructure takes the form of units called *modules*, each involving hundreds or thousands of nerve cells in repeating patterns. The advantages of these iterated patterns for brain function remain largely mysterious; for the neurobiologist, however, such iterated arrangements have provided important clues about cortical connectivity and the mechanisms by which neural activity influences brain development (see Chapters 22 and 23).

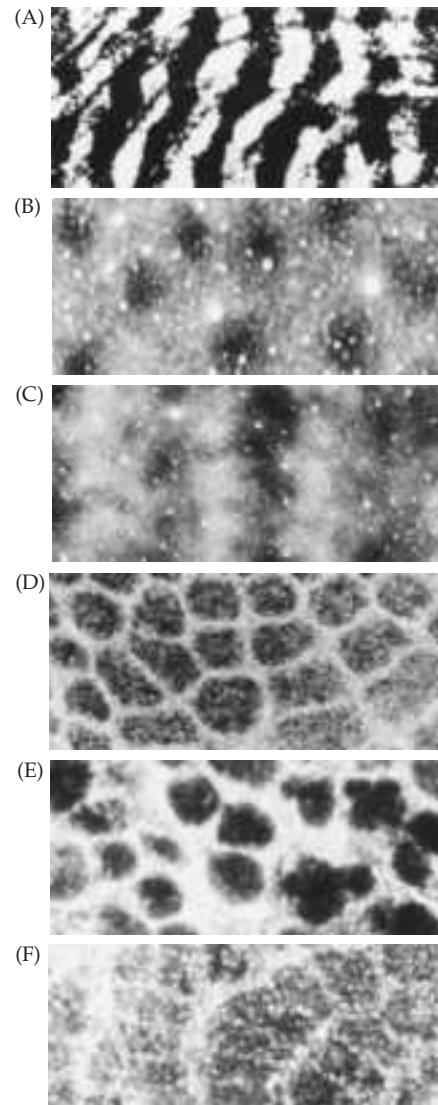
The observation that the somatic sensory cortex comprises elementary units of vertically linked cells was first noted in the 1920s by the Spanish neuroanatomist Rafael Lorente de Nô, based on his studies in the rat. The potential importance of cortical modularity remained largely unexplored until the 1950s, however, when electrophysiological experiments indicated an arrangement of repeating units in the brains of cats and, later, monkeys. Vernon Mountcastle, a neurophysiologist at Johns Hopkins, found that vertical microelectrode penetrations in the primary somatosensory cortex of these animals encountered cells that responded to the same sort of mechanical stimulus presented at the same location on the body surface. Soon after Mountcastle's pioneering work, David Hubel and Torsten Wiesel discovered a similar arrangement in the cat primary visual cortex. These and other observations led Mountcastle to the general view that "the elementary pattern of organization of the cerebral cortex is a vertically oriented column or cylinder of cells capable of input-output functions of considerable complexity." Since these discoveries in the late 1950s and early 1960s, the view that modular circuits represent a fundamental feature of the mammalian cerebral cortex has gained wide acceptance, and many such entities

have now been described in various cortical regions (see figure).

This wealth of evidence for such patterned circuits has led many neuroscientists to conclude, like Mountcastle, that modules are a fundamental feature of the cerebral cortex, essential for perception, cognition, and perhaps even consciousness. Despite the prevalence of iterated modules, there are some problems with the view that modular units are universally important in cortical function. First, although modular circuits of a given class are readily seen in the brains of some species, they have not been found in the same brain regions of other, sometimes closely related, animals. Second, not all regions of the mammalian cortex are organized in a modular fashion. And third, no clear function of such modules has been discerned, much effort and speculation notwithstanding. This salient feature of the organization of the somatic sensory cortex and other cortical (and some subcortical) regions therefore remains a tantalizing puzzle.

References

- HUBEL, D. H. (1988) *Eye, Brain, and Vision*. Scientific American Library. New York: W. H. Freeman.
- LORENTE DE NÔ, R. (1949) The structure of the cerebral cortex. *Physiology of the Nervous System*, 3rd Ed. New York: Oxford University Press.
- MOUNTCASTLE, V. B. (1957) Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. Neurophysiol.* 20: 408–434.
- MOUNTCASTLE, V. B. (1998) *Perceptual Neuroscience: The Cerebral Cortex*. Cambridge: Harvard University Press.
- PURVES, D., D. RIDDLE AND A. LAMANTIA (1992) Iterated patterns of brain circuitry (or how the cortex gets its spots). *Trends Neurosci.* 15: 362–368.
- WOOLSEY, T. A. AND H. VAN DER LOOS (1970) The structural organization of layer IV in the somatosensory region (SI) of mouse cerebral cortex. The description of a cortical field composed of discrete cytoarchitectonic units. *Brain Res.* 17: 205–242.



Examples of iterated, modular substructures in the mammalian brain. (A) Ocular dominance columns in layer IV in the primary visual cortex (V1) of a rhesus monkey. (B) Repeating units called "blobs" in layers II and III in V1 of a squirrel monkey. (C) Stripes in layers II and III in V2 of a squirrel monkey. (D) Barrels in layer IV in primary somatic sensory cortex of a rat. (E) Glomeruli in the olfactory bulb of a mouse. (F) Iterated units called "barreloids" in the thalamus of a rat. These and other examples indicate that modular organization is commonplace in the brain. These units are on the order of one hundred to several hundred microns across. (From Purves et al., 1992.)

Summary

The components of the somatic sensory system considered in this chapter process information conveyed by mechanical stimuli that impinge upon the body surface or that are generated within the body itself (proprioception). This processing is performed by neurons distributed across several brain structures that are connected by both ascending and descending pathways. Transmission of afferent mechanosensory information from the periphery to the brain begins with a variety of receptor types that initiate action potentials. This activity is conveyed centrally via a chain of neurons, referred to as the first-, second-, and third-order cells. First-order neurons are located in the dorsal root and cranial nerve ganglia. Second-order neurons are located in brainstem nuclei. Third-order neurons are found in the thalamus, from whence they project to the cerebral cortex. These pathways are topographically arranged throughout the system, the amount of cortical and subcortical space allocated to various body parts being proportional to the density of peripheral receptors. Studies of non-human primates show that specific cortical regions correspond to each functional submodality; area 3b, for example, processes information from low-threshold cutaneous receptors, and area 3a from proprioceptors. Thus, at least two broad criteria operate in the organization of the somatic sensory system: modality and somatotopy. The end result of this complex interaction is the unified perceptual representation of the body and its ongoing interaction with the environment.

Additional Reading

Reviews

- CHAPIN, J. K. (1987) Modulation of cutaneous sensory transmission during movement: Possible mechanisms and biological significance. In *Higher Brain Function: Recent Explorations of the Brain's Emergent Properties*. S. P. Wise (ed.). New York: John Wiley and Sons, pp. 181–208.
- DARIAN-SMITH, I. (1982) Touch in primates. *Annu. Rev. Psychol.* 33: 155–194.
- JOHANSSON, R. S. AND A. B. VALLBO (1983) Tactile sensory coding in the glabrous skin of the human. *Trends Neurosci.* 6: 27–32.
- KAAS, J. H. (1990) Somatosensory system. In *The Human Nervous System*. G Paxinos (ed.). San Diego: Academic Press, pp. 813–844.
- KAAS, J. H. (1993) The functional organization of somatosensory cortex in primates. *Ann. Anat.* 175: 509–518.
- KAAS, J. H. AND C. E. COLLINS (2003) The organization of somatosensory cortex in anthropoid primates. *Adv. Neurotol.* 2003: 93: 57–67.
- MOUNTCASTLE, V. B. (1975) The view from within: Pathways to the study of perception. *Johns Hopkins Med. J.* 136: 109–131.
- NICOLELIS, M. A. AND E. E. FANSELLOW (2002) Thalamocortical optimization of tactile processing according to behavioral state. *Nat. Neurosci.* 5(6): 517–523.
- PETERSEN, R. S., S. PANZERI AND M. E. DIAMOND (2002) Population coding in somatosensory

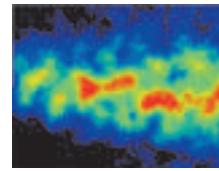
cortex. *Curr. Opin. Neurobiol.* 12(4): 441–447.

WOOLSEY, C. (1958) Organization of somatic sensory and motor areas of the cerebral cortex. In *Biological and Biochemical Bases of Behavior*. H. F. Harlow and C. N. Woolsey (eds.). Madison, WI: University of Wisconsin Press, pp. 63–82.

Important Original Papers

- ADRIAN, E. D. AND Y. ZOTTERMAN (1926) The impulses produced by sensory nerve endings. II. The response of a single end organ. *J. Physiol.* 61: 151–171.
- JOHANSSON, R. S. (1978) Tactile sensibility of the human hand: Receptive field characteristics of mechanoreceptive units in the glabrous skin. *J. Physiol. (Lond.)* 281: 101–123.
- JOHNSON, K. O. AND G. D. LAMB (1981) Neural mechanisms of spatial tactile discrimination: Neural patterns evoked by Braille-like dot patterns in the monkey. *J. Physiol. (London)* 310: 117–144.
- JONES, E. G. AND D. P. FRIEDMAN (1982) Projection pattern of functional components of thalamic ventrobasal complex on monkey somatosensory cortex. *J. Neurophysiol.* 48: 521–544.
- JONES, E. G. AND T. P. S. POWELL (1969) Connexions of the somatic sensory cortex of the rhesus monkey. I. Ipsilateral connexions. *Brain* 92: 477–502.
- LAMOTTE, R. H. AND M. A. SRINIVASAN (1987) Tactile discrimination of shape: Responses of rapidly adapting mechanoreceptive afferents to a step stroked across the monkey finger-
- pad. *J. Neurosci.* 7: 1672–1681.
- LAUBACH, M., J. WESSBER AND M. A. L. NICOLELIS (2000) Cortical ensemble activity increasingly predicts behavior outcomes during learning of a motor task. *Nature* 405: 567–571.
- MOORE, C. I. AND S. B. NELSON (1998) Spatio-temporal subthreshold receptive fields in the vibrissa representation of rat primary somatosensory cortex. *J. Neurophysiol.* 80: 2882–2892.
- MOORE, C. I., S. B. NELSON AND M. SUR (1999) Dynamics of neuronal processing in rat somatosensory cortex. *TINS* 22: 513–520.
- NICOLELIS, M. A. L., L. A. BACCALA, R. C. S. LIN AND J. K. CHAPIN (1995) Sensorimotor encoding by synchronous neural ensemble activity at multiple levels of the somatosensory system. *Science* 268: 1353–1358.
- SUR, M. (1980) Receptive fields of neurons in areas 3b and 1 of somatosensory cortex in monkeys. *Brain Res.* 198: 465–471.
- WALL, P. D. AND W. NOORDENHOS (1977) Sensory functions which remain in man after complete transection of dorsal columns. *Brain* 100: 641–653.
- ZHU, J. J. AND B. CONNORS (1999) Intrinsic firing patterns and whisker-evoked synaptic responses of neurons in the rat barrel cortex. *J. Neurophysiol.* 81: 1171–1183.

Chapter 9



Pain

Overview

A natural assumption is that the sensation of pain arises from excessive stimulation of the same receptors that generate other somatic sensations (i.e., those discussed in Chapter 8). This is not the case. Although similar in some ways to the sensory processing of ordinary mechanical stimulation, the perception of pain, called nociception, depends on specifically dedicated receptors and pathways. Since alerting the brain to the dangers implied by noxious stimuli differs substantially from informing it about innocuous somatic sensory stimuli, it makes good sense that a special subsystem be devoted to the perception of potentially threatening circumstances. The overriding importance of pain in clinical practice, as well as the many aspects of pain physiology and pharmacology that remain imperfectly understood, continue to make nociception an extremely active area of research.

Nociceptors

The relatively unspecialized nerve cell endings that initiate the sensation of pain are called **nociceptors** (*noci* is derived from the Latin *nocere*, “to hurt”). Like other cutaneous and subcutaneous receptors, they transduce a variety of stimuli into receptor potentials, which in turn trigger afferent action potentials (see Figure 8.2). Moreover, nociceptors, like other somatic sensory receptors, arise from cell bodies in dorsal root ganglia (or in the trigeminal ganglion) that send one axonal process to the periphery and the other into the spinal cord or brainstem (see Figure 8.1).

Because peripheral nociceptive axons terminate in unspecialized “free endings,” it is conventional to categorize nociceptors according to the properties of the axons associated with them (see Table 8.1). As described in the previous chapter, the somatic sensory receptors responsible for the perception of innocuous mechanical stimuli are associated with myelinated axons that have relatively rapid conduction velocities. The axons associated with nociceptors, in contrast, conduct relatively slowly, being only lightly myelinated or, more commonly, unmyelinated. Accordingly, axons conveying information about pain fall into either the A δ group of myelinated axons, which conduct at about 20 m/s, or into the C fiber group of unmyelinated axons, which conduct at velocities generally less than 2 m/s. Thus, even though the conduction of all nociceptive information is relatively slow, there are fast and slow pain pathways.

In general, the faster-conducting A δ nociceptors respond either to dangerously intense mechanical or to mechanothermal stimuli, and have receptive fields that consist of clusters of sensitive spots. Other unmyelinated nociceptors tend to respond to thermal, mechanical, and chemical stimuli, and are

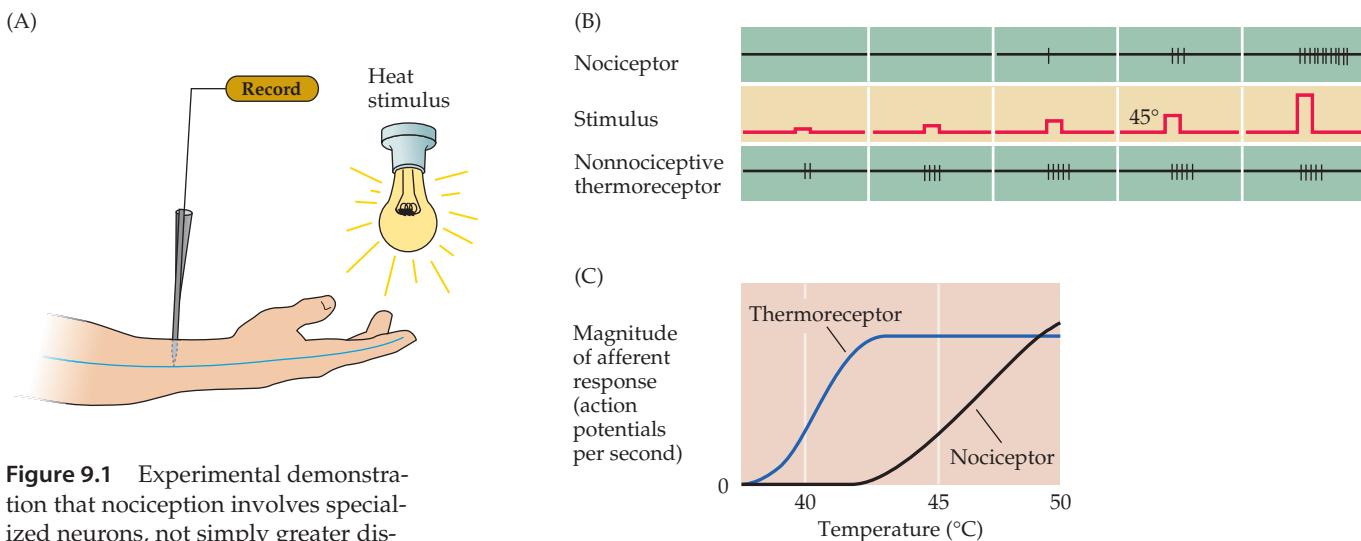
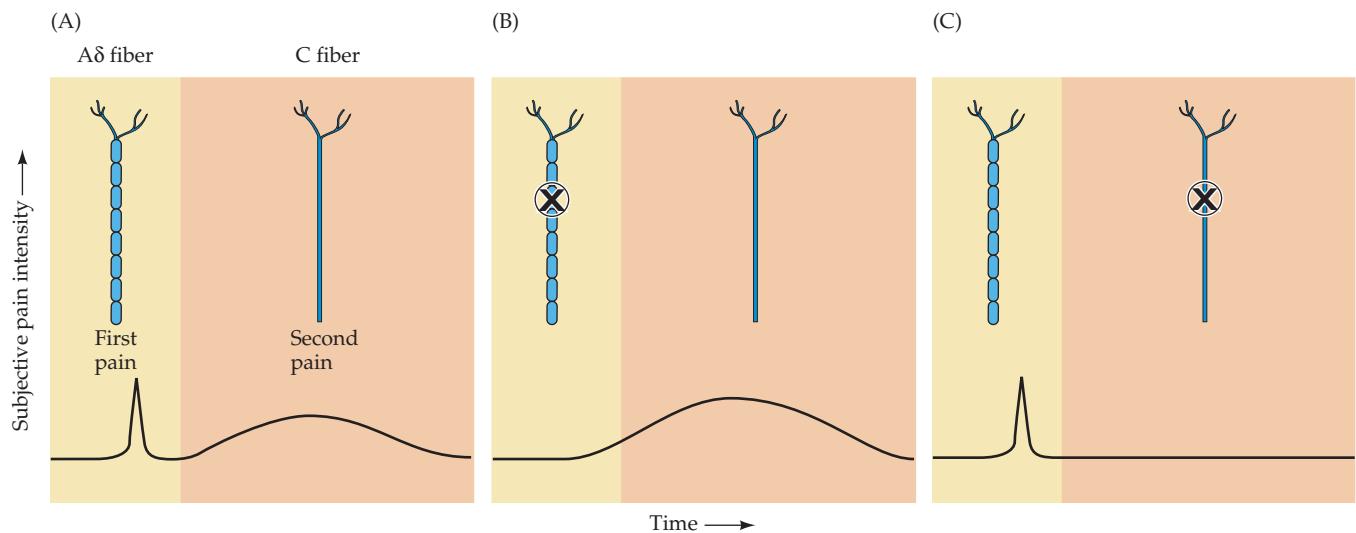


Figure 9.1 Experimental demonstration that nociception involves specialized neurons, not simply greater discharge of the neurons that respond to normal stimulus intensities. (A) Arrangement for transcutaneous nerve recording. (B) In the painful stimulus range, the axons of thermoreceptors fire action potentials at the same rate as at lower temperatures; the number and frequency of action potential discharge in the nociceptive axon, however, continues to increase. (Note that 45°C is the approximate threshold for pain.) (C) Summary of results. (After Fields, 1987.)

therefore said to be *polymodal*. In short, there are three major classes of nociceptors in the skin: **A δ mechanosensitive nociceptors**; **A δ mechanothermal nociceptors**; and **polymodal nociceptors**, the latter being specifically associated with C fibers. The receptive fields of all pain-sensitive neurons are relatively large, particularly at the level of the thalamus and cortex, presumably because the detection of pain is more important than its precise localization.

Studies carried out in both humans and experimental animals demonstrated some time ago that the rapidly conducting axons that subserve somatic sensory sensation are not involved in the transmission of pain. A typical experiment of this sort is illustrated in Figure 9.1. The peripheral axons responsive to nonpainful mechanical or thermal stimuli do not discharge at a greater rate when painful stimuli are delivered to the same region of the skin surface. The nociceptive axons, on the other hand, begin to discharge only when the strength of the stimulus (a thermal stimulus in the example in Figure 9.1) reaches high levels; at this same stimulus intensity, other thermoreceptors discharge at a rate no different from the maximum rate already achieved within the nonpainful temperature range, indicating that there are both nociceptive and nonnociceptive thermoreceptors. Equally important, direct stimulation of the large-diameter somatic sensory afferents at any frequency in humans does not produce sensations that are described as painful. In contrast, the smaller-diameter, more slowly conducting A δ and C fibers are active when painful stimuli are delivered; and when stimulated electrically in human subjects, they produce pain.

How, then, do these different classes of nociceptors lead to the perception of pain? As mentioned, one way of determining the answer has been to stimulate different nociceptors in human volunteers while noting the sensations reported. In general, two categories of pain perception have been described: a sharp **first pain** and a more delayed, diffuse, and longer-lasting sensation that is generally called **second pain** (Figure 9.2A). Stimulation of the large, rapidly conducting A α and A β axons in peripheral nerves does not elicit the sensation of pain. When the stimulus intensity is raised to a level that activates a subset of A δ fibers, however, a tingling sensation or, if the stimulation is intense enough, a feeling of sharp pain is reported. If the stimulus intensity is increased still further, so that the small-diameter, slowly conducting C fiber axons are brought into play, then a duller, longer-lasting



sensation of pain is experienced. It is also possible to selectively anesthetize C fibers and A δ fibers; in general, these selective blocking experiments confirm that the A δ fibers are responsible for first pain, and that C fibers are responsible for the duller, longer-lasting second pain (Figure 9.2B,C).

Transduction of Nociceptive Signals

Given the variety of stimuli (mechanical, thermal, and chemical) that can give rise to painful sensations, the transduction of nociceptive signals is a complex task. While many puzzles remain, some insights have come from the identification of specific receptors associated with nociceptive afferent endings. These receptors are sensitive to both heat and to capsaicin, the ingredient in chili peppers that is responsible for the familiar tingling or burning sensation produced by spicy foods (Box A). The so-called vanilloid receptor (VR-1 or TRPV1) is found in C and A δ fibers and is activated by moderate heat (45°C—a temperature that is perceived as uncomfortable) as well as by capsaicin. Another type of receptor (vanilloid-like receptor, VRL-1 or TRPV2) has a higher threshold response to heat (52°C), is not sensitive to capsaicin, and is found in A δ fibers. Both are members of the larger family of *transient receptor potential* (TRP) channels, first identified in studies of the phototransduction pathway in fruit flies and now known to comprise a large number of receptors sensitive to different ranges of heat and cold. Structurally, TRP channels resemble voltage-gated potassium or cyclic nucleotide-gated channels, having six transmembrane domains with a pore between domains 5 and 6. Under resting conditions the pore of the channel is closed. In the open, activated state, these receptors allow an influx of sodium and calcium that initiates the generation of action potentials in the nociceptive fibers.

Since the same receptor is responsive to heat as well as capsaicin, it is not surprising that chili peppers seem "hot." A puzzle, however, is why the nervous system has evolved receptors that are sensitive to a chemical in chili peppers. As with the case of other plant compounds that selectively activate neural receptors (see the discussion of opiates below), it seems likely that TRPV1 receptors detect endogenous substances whose chemical structure resembles that of capsaicin. In fact, there is now some evidence that 'endovanilloids' that are produced by peripheral tissues in response to injury,

Figure 9.2 Pain can be separated into an early perception of sharp pain and a later sensation that is described as having a duller, burning quality. (A) First and second pain, as these sensations are called, are carried by different axons, as can be shown by (B) the selective blockade of the more rapidly conducting myelinated axons that carry the sensation of first pain, or (C) blockade of the more slowly conducting C fibers that carry the sensation of second pain. (After Fields, 1990.)

Box A

Capsaicin

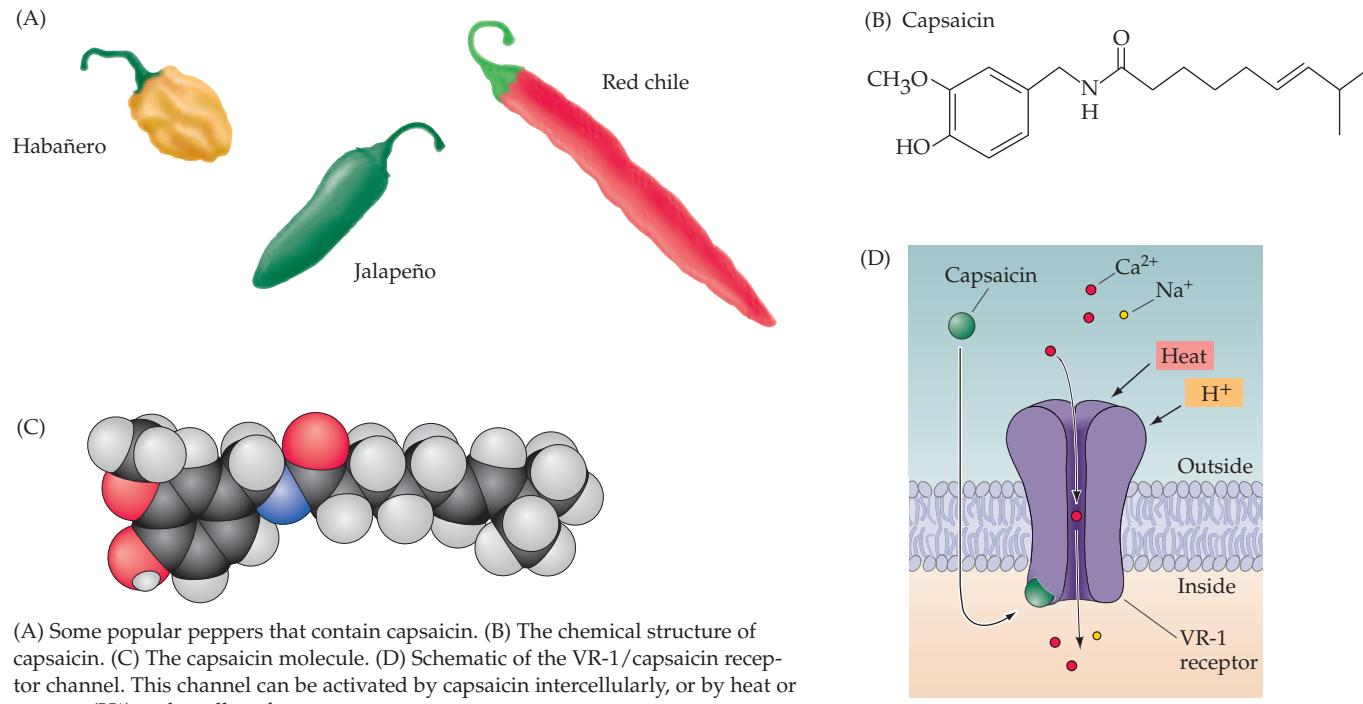
Capsaicin, the principle ingredient responsible for the pungency of hot peppers, is eaten daily by over a third of the world's population. Capsaicin activates responses in a subset of nociceptive C fibers (polymodal nociceptors; see Chapter 9) by opening ligand-gated ion channels that permit the entry of Na^+ and Ca^{2+} . One of these channels (VR-1) has been cloned and has been found to be activated by capsaicin, acid, and anandamide (an endogenous compound that also activates cannabinoid receptors), and by heating the tissue to about 43°C. It follows that anandamide and temperature are probably the endogenous activators of these channels. Mice whose VR-1 receptors have been knocked out drink capsaicin solutions as if they were water. Receptors for capsaicin have been found in polymodal nociceptors of all mammals, but are not present in birds (leading to the produc-

tion of squirrel-proof birdseed laced with capsaicin!).

When applied to the mucus membranes of the oral cavity, capsaicin acts as an irritant, producing protective reactions. When injected into skin, it produces a burning pain and elicits hyperalgesia to thermal and mechanical stimuli. Repeated applications of capsaicin also desensitize pain fibers and prevent neuromodulators such as substance P, VIP, and somatostatin from being released by peripheral and central nerve terminals. Consequently, capsaicin is used clinically as an analgesic and anti-inflammatory agent; it is usually applied topically in a cream (0.075%) to relieve the pain associated with arthritis, postherpetic neuralgia, mastectomy, and trigeminal neuralgia. Thus, this remarkable chemical irritant not only gives gustatory pleasure on an enormous scale, but is also a useful pain reliever!

References

- CATERINA, M. J., M. A. SCHUMACHER, M. TOMINAGA, T. A. ROSEN, J. D. LEVINE AND D. JULIUS (1997) The capsaicin receptor: A heat-activated ion channel in the pain pathway. *Nature* 389: 816–826.
- CATERINA, M. J. AND 8 OTHERS (2000) Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science* 288: 306–313.
- SZALLASI, A. AND P. M. BLUMBERG (1999) Vanilloid (capsaicin) receptors and mechanisms. *Pharm. Reviews* 51: 159–212.
- TOMINAGA, M. AND 8 OTHERS (1998) The cloned capsaicin receptor integrates multiple pain-producing stimuli. *Neuron* 21: 531–543.
- ZYGMUNT, P. M. AND 7 OTHERS (1999) Vanilloid receptors on sensory nerves mediate the vasodilator action of anandamide. *Nature* 400: 452–457.



and that these substances, along with other factors, contribute to the nociceptive response to injury.

Central Pain Pathways

The pathways that carry information about noxious stimuli to the brain, as might be expected for such an important and multifaceted system, are also complex (see Boxes B and C). It helps in understanding this complexity to distinguish two components of pain: the sensory discriminative component, which signals the location, intensity, and quality of the noxious stimulation, and the affective-motivational component of pain—which signals the unpleasant quality of the experience, and enables the autonomic activation that follows a noxious stimulus (the classic fight-or-flight reaction; see Chapter 20). The discriminative component is thought to depend on pathways that target the traditional somatosensory areas of cortex, while the affective-motivational component is thought to depend on additional cortical and brainstem pathways. The major pathways are summarized in Figure 9.3.

Pathways responsible for the discriminative component of pain originate with other sensory neurons, in dorsal root ganglia and, like other sensory nerve cells the central axons of nociceptive nerve cells enter the spinal cord via the dorsal roots (Figure 9.3A). When these centrally projecting axons reach the dorsal horn of the spinal cord, they branch into ascending and descending collaterals, forming the **dorsolateral tract of Lissauer** (named after the German neurologist who first described this pathway in the late nineteenth century). Axons in Lissauer's tract typically run up and down for one or two spinal cord segments before they penetrate the gray matter of the dorsal horn. Once within the dorsal horn, the axons give off branches that contact neurons located in several of Rexed's laminae (these laminae are the descriptive divisions of the spinal gray matter in cross section, again named after the neuroanatomist who described these details in the 1950s).

The axons of these second-order neurons in the dorsal horn of the spinal cord cross the midline and ascend all the way to the brainstem and thalamus in the anterolateral (also called ventrolateral) quadrant of the contralateral half of the spinal cord. These fibers form the **spinothalamic tract**, the major ascending pathway for information about pain and temperature. This overall pathway is also referred to as the **anterolateral system**, much as the mechanosensory pathway is referred to as the dorsal column-medial lemniscus system.

The location of the spinothalamic tract is particularly important clinically because of the characteristic sensory deficits that follow certain spinal cord injuries. Since the mechanosensory pathway ascends ipsilaterally in the cord, a unilateral spinal lesion will produce sensory loss of touch, pressure, vibration, and proprioception below the lesion on the same side. The pathways for pain and temperature, however, cross the midline to ascend on the opposite side of the cord. Therefore, diminished sensation of pain below the lesion will be observed on the side *opposite* the mechanosensory loss (and the lesion). This pattern is referred to as a **dissociated sensory loss** and (together with local dermatomal signs; see Box C in Chapter 8) helps define the level of the lesion (Figure 9.4).

As is the case of the mechanosensory pathway, information about noxious and thermal stimulation of the face follows a separate route to the thalamus (see Figure 9.3B). First-order axons originating from the trigeminal ganglion cells and from ganglia associated with nerves VII, IX, and X carry information from facial nociceptors and thermoreceptors into the brainstem. After

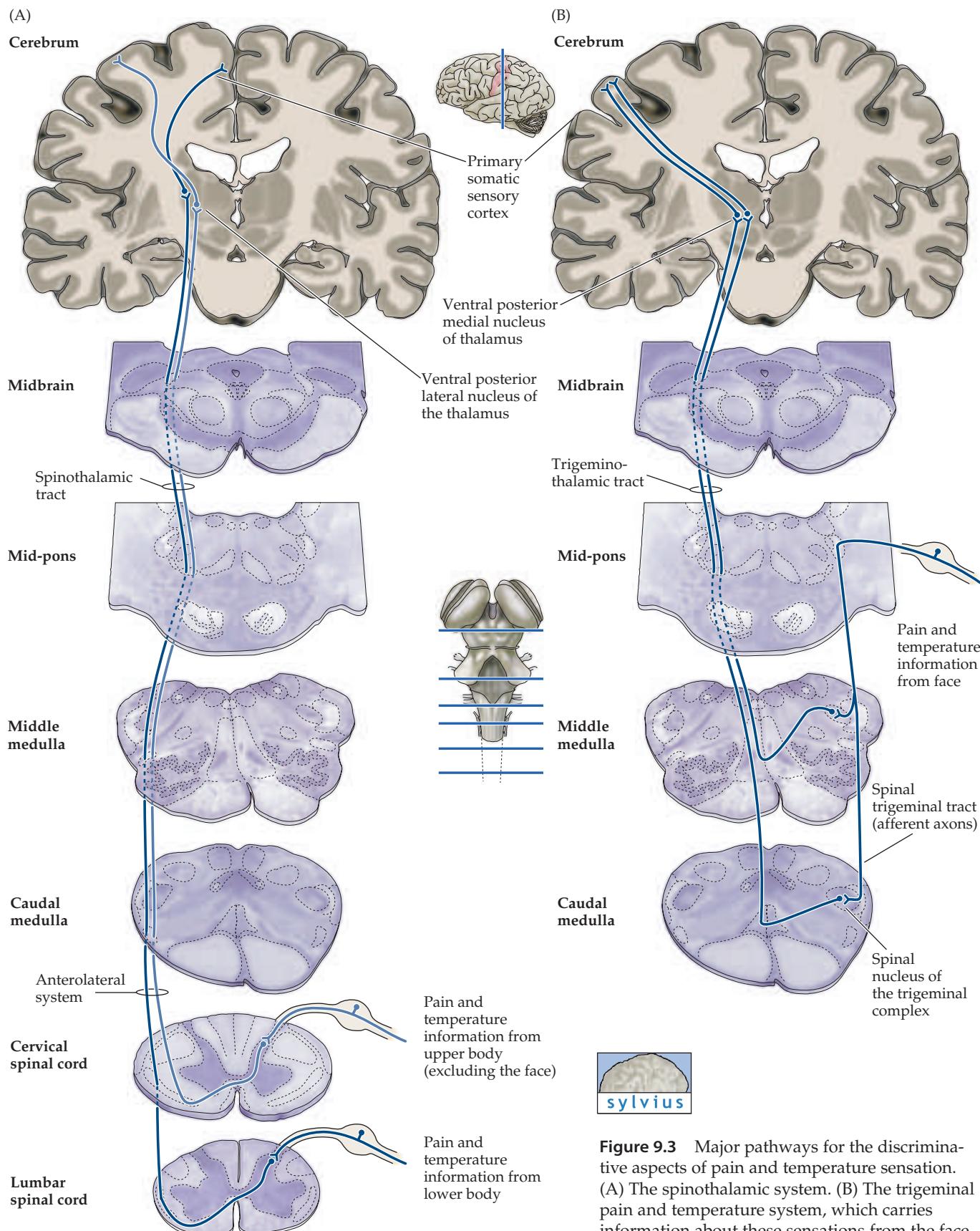


Figure 9.3 Major pathways for the discriminative aspects of pain and temperature sensation. (A) The spinothalamic system. (B) The trigeminal pain and temperature system, which carries information about these sensations from the face.

Box B

Referred Pain

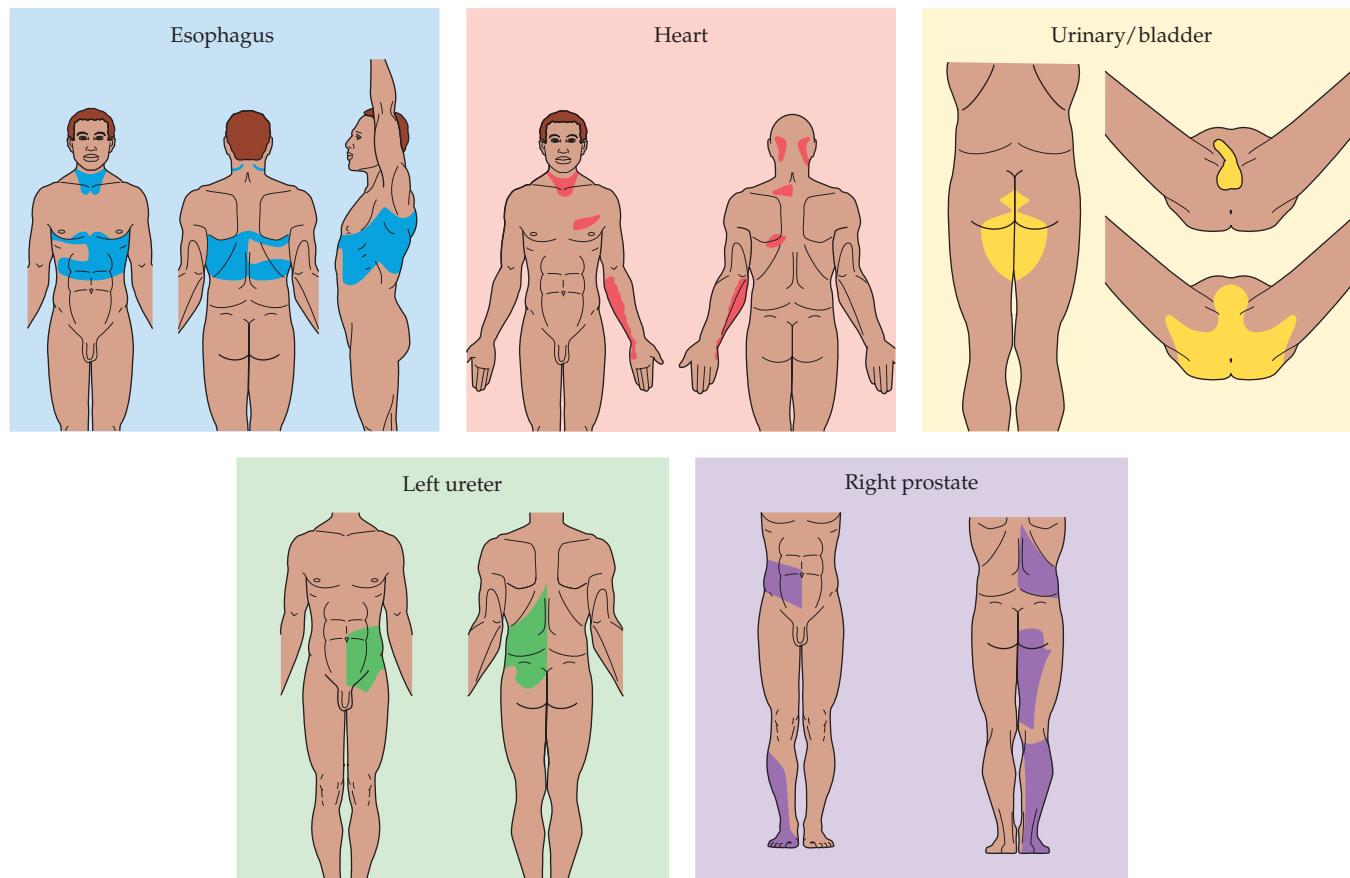
Surprisingly, there are few, if any, neurons in the dorsal horn of the spinal cord that are specialized solely for the transmission of *visceral* pain. Obviously, we recognize such pain, but it is conveyed centrally via dorsal horn neurons that are also concerned with *cutaneous* pain. As a result of this economical arrangement, the disorder of an internal organ is sometimes perceived as cutaneous pain. A patient may therefore present to the physician with the complaint of pain at a site other than its actual source, a potentially confusing phenomenon called referred pain. The most common clinical example is anginal pain (pain

arising from heart muscle that is not being adequately perfused with blood) referred to the upper chest wall, with radiation into the left arm and hand. Other important examples are gallbladder pain referred to the scapular region, esophageal pain referred to the chest wall, ureteral pain (e.g., from passing a kidney stone) referred to the lower abdominal wall, bladder pain referred to the perineum, and the pain from an inflamed appendix referred to the anterior abdominal wall around the umbilicus. Understanding referred pain can lead to an astute diagnosis that might otherwise be missed.

References

- CAPPS, J. A. AND G. H. COLEMAN (1932) *An Experimental and Clinical Study of Pain in the Pleura, Pericardium, and Peritoneum*. New York: Macmillan.
- HEAD, H. (1893) On disturbances of sensation with special reference to the pain of visceral disease. *Brain* 16: 1–32.
- KELLGREW, J. H. (1939–1942) On the distribution of pain arising from deep somatic structures with charts of segmental pain areas. *Clin. Sci.* 4: 35–46.

Examples of pain arising from a visceral disorder referred to a cutaneous region (color).



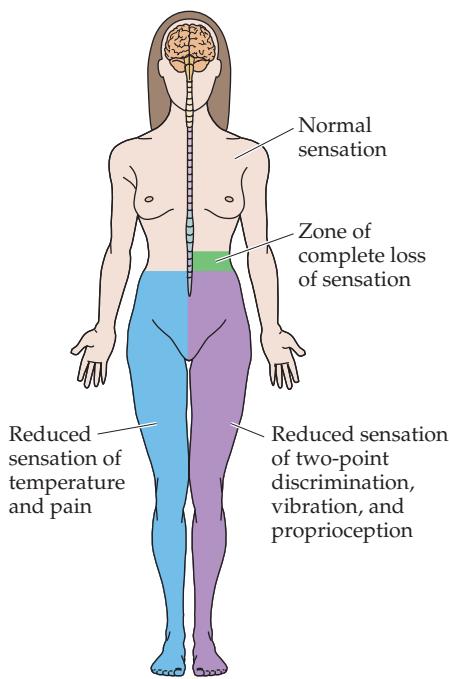


Figure 9.4 Pattern of “dissociated” sensory loss following a spinal cord hemisection at the 10th thoracic level on the left side. This pattern, together with motor weakness on the same side as the lesion, is sometimes referred to as Brown-Séquard syndrome.

entering the pons, these small myelinated and unmyelinated trigeminal fibers *descend* to the medulla, forming the **spinal trigeminal tract** (or spinal tract of cranial nerve V), and terminate in two subdivisions of the spinal trigeminal complex: the pars interpolaris and pars caudalis. Axons from the second-order neurons in these two trigeminal nuclei, like their counterparts in the spinal cord, cross the midline and ascend to the contralateral thalamus in the **trigeminothalamic tract**.

The principal target of the spinothalamic and trigeminothalamic pathway is the ventral posterior nucleus of the thalamus. Similar to the organization of the mechanosensory pathways, information from the body terminates in the VPL, while information from the face terminate in the VPM. These nuclei send their axons to primary and secondary somatosensory cortex. The nociceptive information transmitted to these cortical areas is thought to be responsible for the discriminative component of pain: identifying the location, the intensity, and quality of the stimulation. Consistent with this interpretation, electrophysiological recordings from nociceptive neurons in S1, show that these neurons have small localized receptive fields, properties commensurate with behavioral measures of pain localization.

The affective-motivational aspect of pain is evidently mediated by separate projections of the anterolateral system to the reticular formation of the midbrain (in particular the parabrachial nucleus), and to thalamic nuclei that lie medial to the ventral posterior nucleus (including the so-called intralaminar nuclei; see Figure 9.5). Studies in rodents show that neurons in the parabrachial nucleus respond to most types of noxious stimuli, and have large receptive fields that can include the whole surface of the body. Neurons in the parabrachial nucleus project in turn to the hypothalamus and the amygdala, thus providing nociceptive information to circuits known to be concerned with motivation and affect (see Chapter 28). These parabrachial targets are also the source of projections to the periaqueductal grey of the midbrain, a structure that plays an important role in the descending control of activity in the pain pathway. Nociceptive inputs to the parabrachial nucleus and to the ventral posterior nucleus arise from separate populations of neurons in the dorsal horn of the spinal cord. Parabrachial inputs arise from neurons in the most superficial part of the dorsal horn (lamina I), while ventral posterior inputs arise from deeper parts of the dorsal horn (e.g., lamina V). By taking advantage of the unique molecular signature of these two sets of neurons, it has been possible to selectively eliminate the nociceptive inputs to the parabrachial nucleus in rodents. In these animals, the behavioral responses to the presentation of noxious stimulation (capsaicin, for example) are substantially attenuated.

Projections from the anterolateral system to the medial thalamic nuclei provide nociceptive signals to areas in the frontal lobe, the insula and the cingulate cortex (Figure 9.5). In accord with this anatomy, functional imaging studies in humans have shown a strong correlation between activity in the anterior cingulate cortex and the experience of a painful stimulus. Moreover, experiments using hypnosis have been able to tease apart the neural response to changes in the intensity of a painful stimulus from changes in its unpleasantness. Changes in intensity are accompanied by changes in the activity of neurons in somatosensory cortex, with little change in the activity of cingulate cortex, whereas changes in unpleasantness are correlated with changes in the activity of neurons in cingulate cortex.

From this description, it should be evident that the full experience of pain involves the cooperative action of an extensive network of brain regions whose properties are only beginning to be understood (Box C). The cortical

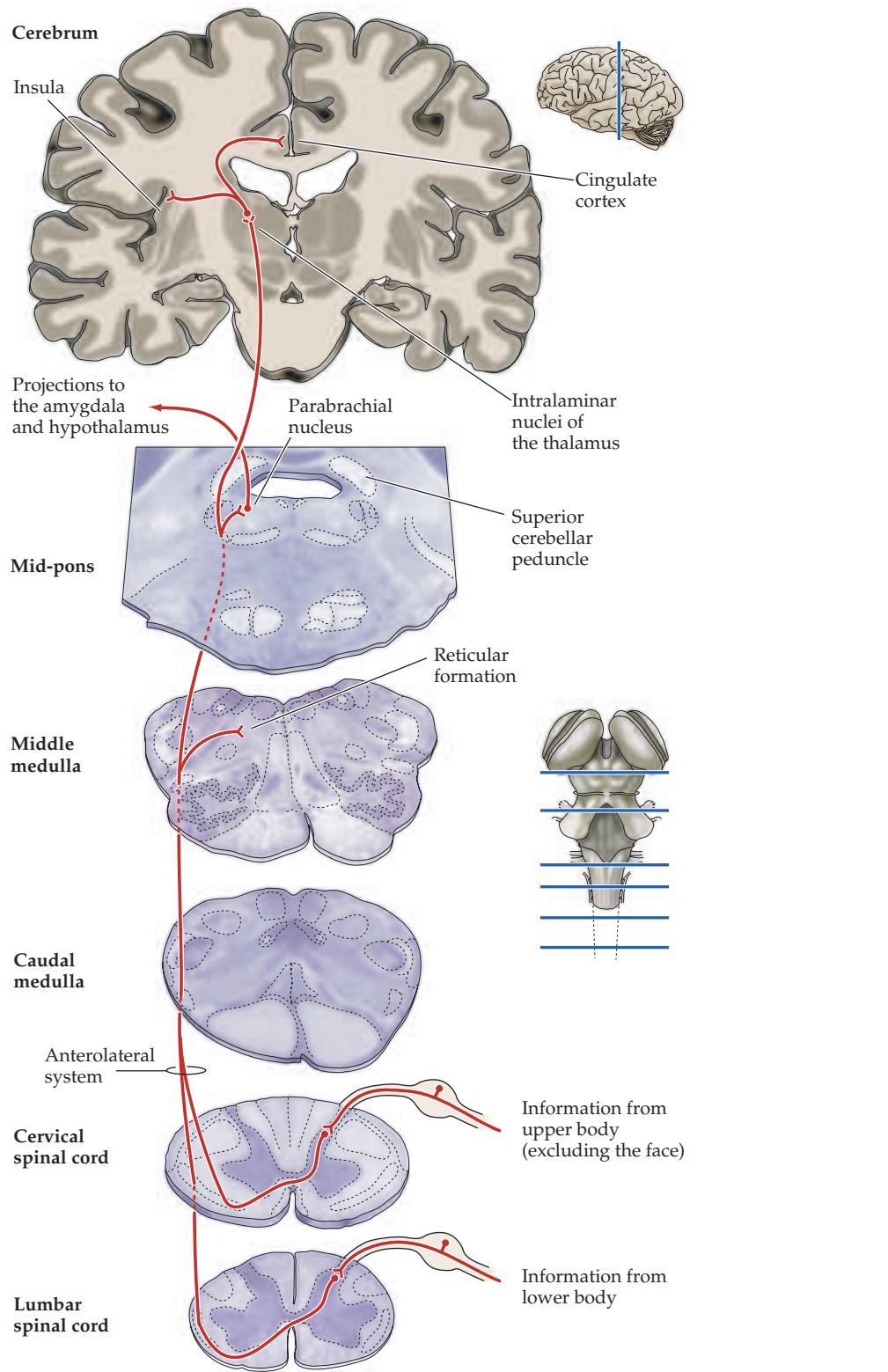


Figure 9.5 Affective-motivational pain pathways. Nociceptive information critical for signaling the unpleasant quality of pain is mediated by projections to the reticular formation (including the parabrachial nucleus) and to the intralaminar nuclei of the thalamus.



Box C

A Dorsal Column Pathway for Visceral Pain

Chapters 8 and 9 present a framework for considering the central neural pathways that convey innocuous mechanosensory signals and painful signals from cutaneous and deep somatic sources. Considering just the signals derived from the body below the head, discriminative mechanosensory and proprioceptive information travels to the ventral posterior thalamus via the dorsal-column medial lemniscal system (see Figure 8.6A), while nociceptive information travels to the same (and additional) thalamic relays via the anterolateral systems (see Figure 9.3A). But how do painful signals that arise in the visceral organs of the pelvis, abdomen, and thorax enter the central nervous system and ultimately reach consciousness?

The answer is via a newly discovered component of the dorsal column medial lemniscal pathway that conveys visceral nociception. Although Chapter 20 will present more information on the systems that receive and process visceral sensory information, at this juncture it is worth considering this component of the pain pathways and how this particular pathway has begun to impact clinical medicine.

Primary visceral afferents from the pelvic and abdominal viscera enter the spinal cord and synapse on second-order neurons in the dorsal horn of the lumbar-sacral spinal cord. As discussed in Box A and Chapter 20, some of these second-order neurons are cells that give rise to the anterolateral systems and contribute to referred visceral pain patterns. However, other neurons—perhaps primarily those that give rise to nociceptive signals—synapse upon neurons in the intermediate gray region of the spinal cord near the central canal. These neurons, in turn, send their axons not through the anterolateral white matter of the spinal cord (as might be expected for a pain pathway) but through the dorsal

columns in a position very near the midline (see Figure A). Similarly, second-order neurons in the thoracic spinal cord that convey nociceptive signals from thoracic viscera send their axons rostrally through the dorsal columns along the dorsal intermediate septum, near the division of the gracile and cuneate fasciculi. These second order axons then synapse in the dorsal column nuclei of the caudal medulla, where neurons give rise to arcuate fibers that form the contralateral medial lemniscus and eventually synapse on thalamocortical projection neurons in the ventral-posterior thalamus.

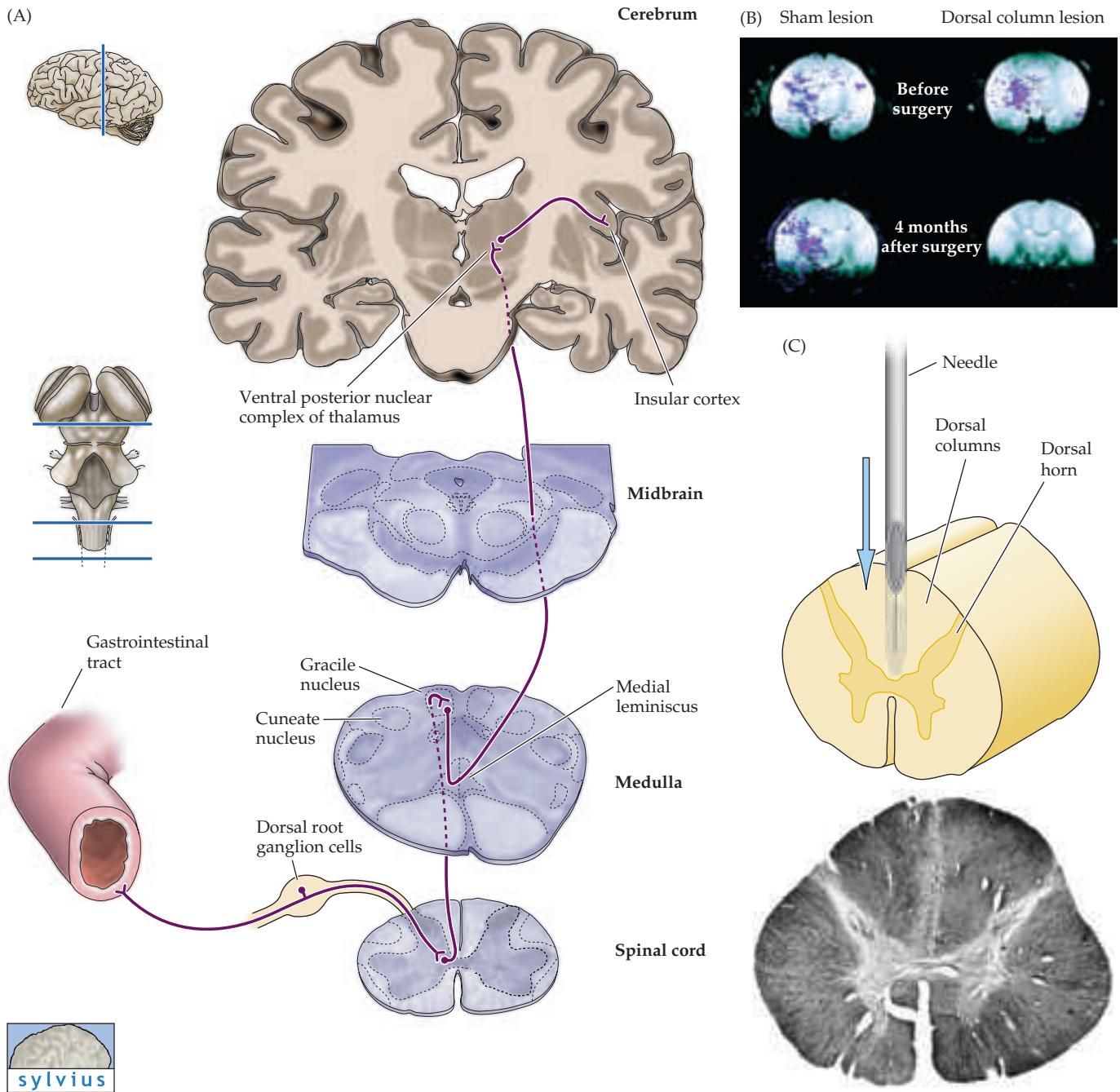
This dorsal column visceral sensory projection now appears to be the principal pathway by which painful sensations arising in the viscera are detected and discriminated. Several observations support this conclusion: (1) neurons in the ventral posterior lateral nucleus, nucleus gracilis and near the central canal of the spinal cord all respond to noxious visceral stimulation; (2) responses of neurons in the ventral posterior lateral nucleus and nucleus gracilis to such stimulation are greatly reduced by spinal lesions of the dorsal columns (see Figure B), but not lesions of the anterolateral white matter; and (3) infusion of drugs that block nociceptive synaptic transmission into the intermediate gray region of the sacral spinal cord blocks the responses of neurons in the nucleus gracilis to noxious visceral stimulation, but not to innocuous cutaneous stimulation.

The discovery of this visceral sensory component in the dorsal-column medial lemniscal system has helped to explain why surgical transection of the axons that run in the medial part of the dorsal columns (a procedure termed *midline myelotomy*) generates significant relief from the debilitating pain that can result from visceral cancers in the abdomen and pelvis. Although the initial develop-

ment of this surgical procedure preceded the elucidation of this visceral pain pathway, these new discoveries have renewed interest in midline myelotomy as a palliative neurosurgical intervention for cancer patients whose pain is otherwise unmanageable. Indeed, precise knowledge of the visceral sensory pathway in the dorsal columns has led to further refinements that permit a minimally invasive (“punctate”) surgical procedure that attempts to interrupt the second-order axons of this pathway within just a single spinal segment (typically, a mid- or lower-thoracic level; see Figure C). In so doing, this procedure offers some hope to patients who struggle to maintain a reasonable quality of life in extraordinarily difficult circumstances.

References

- AL-CHAER, E. D., N. B. LAWAND, K. N. WESTLUND AND W. D. WILLIS (1996) Visceral nociceptive input into the ventral posterolateral nucleus of the thalamus: a new function for the dorsal column pathway. *J. Neurophys.* 76: 2661–2674.
- AL-CHAER, E. D., N. B. LAWAND, K. N. WESTLUND AND W. D. WILLIS (1996) Pelvic visceral input into the nucleus gracilis is largely mediated by the postsynaptic dorsal column pathway. *J. Neurophys.* 76: 2675–2690.
- BECKER, R., S. GATSCHER, U. SURE AND H. BERTLANFFY (2001) The punctate midline myelotomy concept for visceral cancer pain control – case report and review of the literature. *Acta Neurochir. [Suppl.]* 79: 77–78.
- HITCHCOCK, E. R. (1970) Stereotactic cervical myelotomy. *J. Neurol. Neurosurg. Psychiatry* 33: 224–230.
- KIM, Y. S. AND S. J. KWON (2000) High thoracic midline dorsal column myelotomy for severe visceral pain due to advanced stomach cancer. *Neurosurg.* 46:85–90.
- NAUTA, H. AND 8 OTHERS (2000) Punctate midline myelotomy for the relief of visceral cancer pain. *J. Neurosurg. (Spine 2)* 92: 125–130.
- WILLIS, W. D., E. D. AL-CHAER, M. J. QUAST AND K. N. WESTLUND (1999) A visceral pain pathway in the dorsal column of the spinal cord. *Proc. Natl. Acad. Sci. USA* 96: 7675–7679.



(A) A visceral pain pathway in the dorsal-column medial lemniscal system. For simplicity, only the pathways that mediate visceral pain from the pelvis and lower abdomen are illustrated. The mechanosensory component of this system for the discrimination of tactile stimuli and the anterolateral system for the detection of painful and thermal cutaneous stimuli are also shown for comparison (see also Figures 8.6A and 9.3A). (B) Empirical evidence supporting the existence of the visceral pain pathway shown in (A). Increased neural activity was observed with functional MRI techniques in the thalamus of monkeys that were subjected to noxious distention of the colon and rectum,

indicating the processing of visceral pain. This activity was abolished by lesion of the dorsal columns at T10, but not by "sham" surgery. (From Willis et al., 1999.) (C) Top, one method of punctate midline myelotomy for the relief of severe visceral pain. Bottom, myelin-stained section of the thoracic spinal cord (T10) from a patient who underwent midline myelotomy for the treatment of colon cancer pain that was not controlled by analgesics. After surgery, the patient experienced relief from pain during the remaining three months of his life. (From Hirshberg et al., 1996; drawing after Nauta et al., 1997.)

representation of pain is the least well documented aspect of the central pathways for nociception, and further studies will be needed to elucidate the contribution of regions outside the somatosensory areas of the parietal lobe. Nevertheless, a prominent role for these areas in the perception of pain is suggested by the fact that ablations of the relevant regions of the parietal cortex do not generally alleviate chronic pain (although they impair contralateral mechanosensory perception, as expected).

Sensitization

Following a painful stimulus associated with tissue damage (e.g., cuts, scrapes, and bruises), stimuli in the area of the injury and the surrounding region that would ordinarily be perceived as slightly painful are perceived as significantly more so, a phenomenon referred to as **hyperalgesia**. A good example of hyperalgesia is the increased sensitivity to temperature that occurs after a sunburn. This effect is due to changes in neuronal sensitivity that occur at the level of peripheral receptors as well as their central targets.

Peripheral sensitization results from the interaction of nociceptors with the “inflammatory soup” (Figure 9.6) of substances released when tissue is damaged. These products of tissue damage include extracellular protons, arachidonic acid and other lipid metabolites, bradykinin, histamine, serotonin, prostaglandins, nucleotides, and nerve growth factor (NGF), all of which can interact with receptors or ion channels of nociceptive fibers, augmenting their response. For example, the responses of the TRPV1 receptor to heat can be potentiated by direct interaction of the channel with extracellular protons or lipid metabolites. NGF and bradykinin also potentiate the

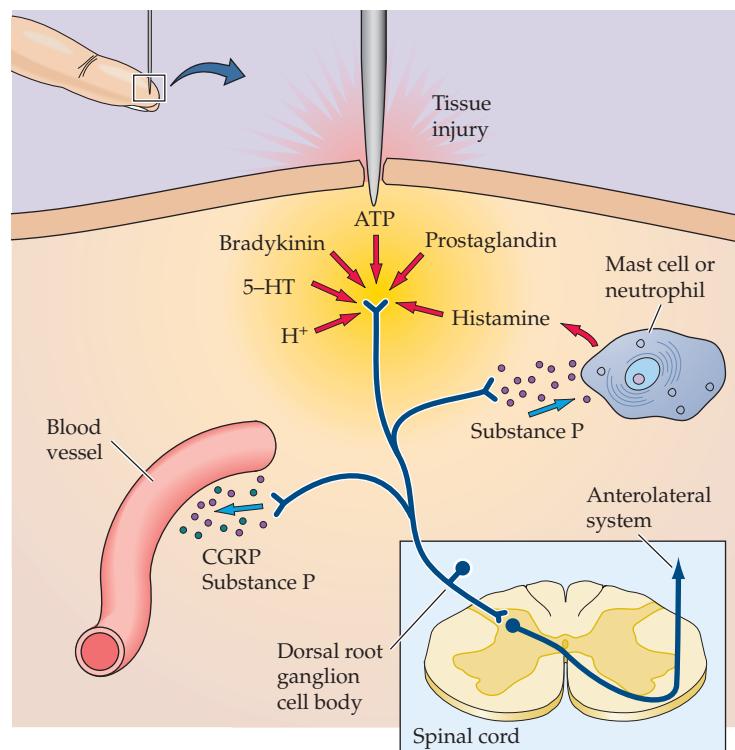


Figure 9.6 Inflammatory response to tissue damage. Substances released by damaged tissues augment the response of nociceptive fibers. In addition, electrical activation of nociceptors causes the release of peptides and neurotransmitters that further contribute to the inflammatory response.

activity of the TRPV1 receptors, but do so indirectly through the actions of separate cell-surface receptors (TrkA and bradykinin receptors respectively) and their associated intracellular signalling pathways. The prostaglandins are thought to contribute to peripheral sensitization by binding to G-protein-coupled receptors that increase levels of cyclic AMP within nociceptors. Prostaglandins also reduce the threshold depolarization required for generating action potentials via phosphorylation of a specific class of TTX-resistant Na channels that are expressed in nociceptors. In addition, electrical activity in the nociceptors causes them to release peptides and neurotransmitters such as substance P, calcitonin-gene-related peptide (CGRP) and ATP which further contribute to the inflammatory response, causing vasodilation, swelling, and the release of histamine from mast cells. The presumed purpose of the complex chemical signaling arising from local damage is not only to protect the injured area (as a result of the painful perceptions produced by ordinary stimuli close to the site of damage), but also to promote healing and guard against infection by means of local effects such as increased blood flow and the migration of white blood cells to the site. Obviously the identification of the components of the inflammatory soup and their mechanisms of action is a fertile area to explore for potential analgesics (i.e., compounds that reduce pain intensity). For example, so-called nonsteroidal anti-inflammatory drugs (NSAIDs), which include aspirin and ibuprofen, act by inhibiting cyclooxygenase (COX), an enzyme important in the biosynthesis of prostaglandins.

Central sensitization refers to an immediate onset, activity dependent increase in the excitability of neurons in the dorsal horn of the spinal cord following high levels of activity in the nociceptive afferents. As a result, activity levels in nociceptive afferents that were subthreshold prior to the sensitizing event, become sufficient to generate action potentials in dorsal horn neurons, contributing to an increase in pain sensitivity. Although central sensitization is triggered in dorsal horn neurons by activity in nociceptors, the effects generalize to other inputs that arise from low threshold mechanoreceptors. Thus, stimuli that under normal conditions would be innocuous (such as brushing the surface of the skin) activate second-order neurons in the dorsal horn that receive nociceptive inputs, and give rise to a sensation of pain. The induction of pain by what is normally an innocuous stimulus is referred to as **allodynia**. This phenomenon typically occurs immediately after the painful event and can outlast the stimulus by several hours.

Like its peripheral counterpart, a number of different mechanisms contribute to central sensitization, and these can be divided broadly into transcription independent and dependent processes. One form of transcription independent central sensitization called “windup” involves a progressive increase in the discharge rate of dorsal horn neurons in response to repeated low frequency activation of nociceptive afferents. A behavioral correlate of the windup phenomenon has been studied by examining the perceived intensity of pain in response to multiple presentations of a noxious stimulus. Although the intensity of the stimulation is constant, the perceived intensity increases with each stimulus presentation. Windup lasts only during the period of stimulation and arises from the summation of the slow synaptic potentials that are evoked in dorsal horn neurons by nociceptive inputs. The sustained depolarization of the dorsal horn neurons results in part from the activation of voltage dependent L-type calcium channels, and from the removal of the Mg block of NMDA receptors, increasing the sensitivity of the

Box D

Phantom Limbs and Phantom Pain

Following the amputation of an extremity, nearly all patients have an illusion that the missing limb is still present. Although this illusion usually diminishes over time, it persists in some degree throughout the amputee's life and can often be reactivated by injury to the stump or other perturbations. Such phantom sensations are not limited to amputated limbs; phantom breasts following mastectomy, phantom genitalia following castration, and phantoms of the entire lower body following spinal cord transection have all been reported. Phantoms are also common after local nerve block for surgery. During recovery from brachial plexus anesthesia, for example, it is not unusual for the patient to experience a phantom arm, perceived as whole and intact, but displaced from the real arm. When the real arm is viewed, the phantom appears to "jump into" the arm and may emerge and reenter intermittently as the anesthesia wears off. These sensory phantoms demonstrate that the central machinery for processing somatic sensory information is

not idle in the absence of peripheral stimuli; apparently, the central sensory processing apparatus continues to operate independently of the periphery, giving rise to these bizarre sensations.

Interestingly, considerable functional reorganization of somatotopic maps in the primary somatosensory cortex occurs in amputees (see Chapter 24). This reorganization starts immediately after the amputation and tends to evolve for several years. One of the effects of this process is that neurons that have lost their original inputs (i.e., from the removed limb) respond to tactile stimulation of other body parts. A surprising consequence is that stimulation of the face, for example, can be experienced as if the missing limb had been touched.

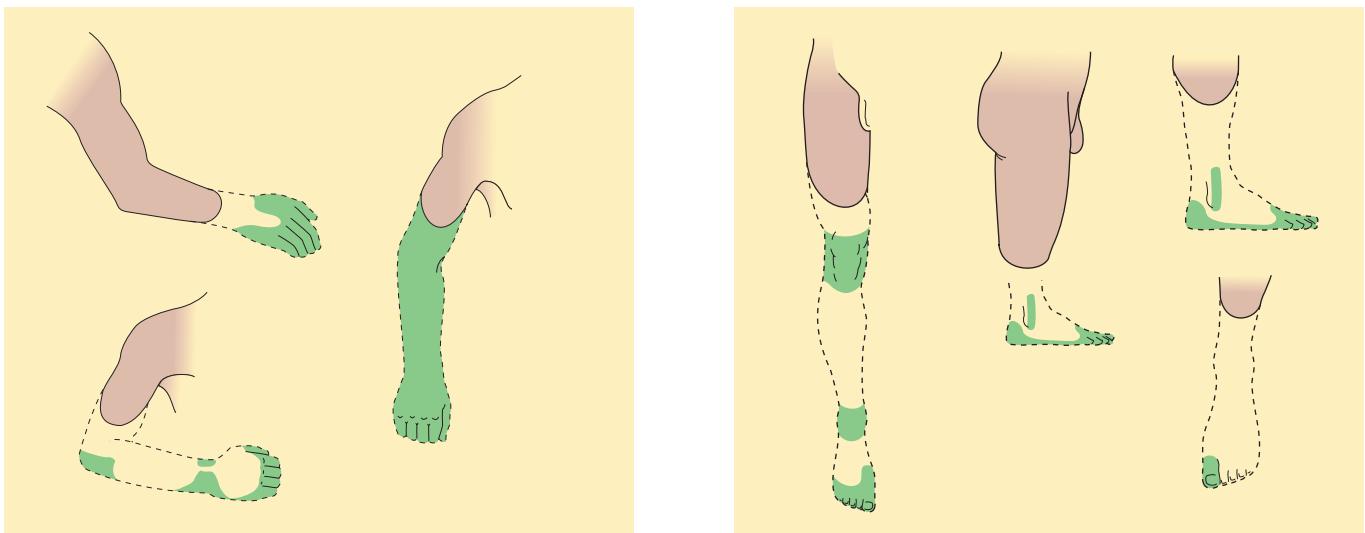
Further evidence that the phenomenon of phantom limb is the result of a central representation is the experience of children born without limbs. Such individuals have rich phantom sensations, despite the fact that a limb never developed. This observation suggests that a full representation of the body

exists independently of the peripheral elements that are mapped. Based on these results, Ronald Melzack proposed that the loss of a limb generates an internal mismatch between the brain's representation of the body and the pattern of peripheral tactile input that reaches the neocortex. The consequence would be an illusory sensation that the missing body part is still present and functional. With time, the brain may adapt to this loss and alter its intrinsic somatic representation to better accord with the new configuration of the body. This change could explain why the phantom sensation appears almost immediately after limb loss, but usually decreases in intensity over time.

Phantoms might simply be a curiosity—or a provocative clue about higher-order somatic sensory processing—were it not for the fact that a substantial number of amputees also develop phantom pain. This common problem is usually described as a tingling or burning sensation in the missing part. Sometimes, however, the sensation becomes a more seri-

the dorsal horn neuron to glutamate, the transmitter in nociceptive afferents. Other forms of central sensitization that last longer than the period of sensory stimulation (such as allodynia) are thought to involve an LTP-like enhancement of postsynaptic potentials (see Chapter 24). The longest lasting forms, resulting from transcription dependent processes, can be elicited by changes in neuronal activity or by humoral signals. Those elicited by neuronal activity are localized to the site of the injury, while humoral activation can lead to more widespread changes. For example, cytokines released from microglia and from other sources promote the widespread transcription of COX-2 and the production of prostaglandins in dorsal horn neurons. As described for nociceptive afferents, increased levels of prostaglandins in CNS neurons augments neuronal excitability. Thus the analgesic effects of drugs that inhibit COX are due to actions in both the periphery and within the dorsal horn.

As injured tissue heals, the sensitization induced by peripheral and central mechanisms typically declines and the threshold for pain returns to



Drawings of phantom arms and legs, based on patients' reports. The phantom is indicated by a dashed line, with the colored regions showing the most vividly experienced parts. Note that some phantoms are telescoped into the stump. (After Solonen, 1962.)

ous pain that patients find increasingly debilitating. Phantom pain is, in fact, one of the more common causes of chronic pain syndromes and is extraordinarily difficult to treat. Because of the widespread nature of central pain processing, ablation of the spinothalamic tract, portions of the thalamus, or even primary

sensory cortex does not generally relieve the discomfort felt by these patients.

References

- MELZACK, R. (1989) Phantom limbs, the self, and the brain. The D.O. Hebb Memorial Lecture. *Canad. Psychol.* 30: 1–14.
- MELZACK, R. (1990) Phantom limbs and the concept of a neuromatrix. *TINS* 13: 88–92.
- NASHOLD, B. S., JR. (1991) Paraplegia and pain. In *Deafferentation Pain Syndromes: Pathophysiology and Treatment*, B. S. Nashold, Jr. and J. Ovelmen-Levitt (eds.). New York: Raven Press, pp. 301–319.
- RAMACHANDRAN, V. S. AND S. BLAKESLEE (1998) *Phantoms in the Brain*. New York: William Morrow & Co.
- SOLONEN, K. A. (1962) The phantom phenomenon in amputated Finnish war veterans. *Acta Orthop. Scand. Suppl.* 54: 1–37.

preinjury levels. However, when the afferent fibers or central pathways themselves are damaged—a frequent complication in pathological conditions that include diabetes, shingles, AIDS, multiple sclerosis, and stroke—these processes can persist. The resulting condition is referred to as **neuropathic pain**, a chronic, intensely painful experience that is difficult to treat with conventional analgesic medications. (See Box D for a description of neuropathic pain associated with amputation of an extremity.) The pain can arise spontaneously (without a stimulus) or can be produced by mild forms of stimulation that are common to everyday experience, such as the gentle touch and pressure of clothing, or warm and cool temperatures. Patients often describe their experience as a constant burning sensation interrupted by episodes of shooting, stabbing, or electric shocklike jolts. Because the disability and psychological stress associated with chronic neuropathic pain can be severe, much present research is being devoted to better understanding of the mechanisms of peripheral and central sensitization with the hope of more effective therapies for this debilitating syndrome.

Descending Control of Pain Perception

With respect to the *interpretation* of pain, observers have long commented on the difference between the objective reality of a painful stimulus and the subjective response to it. Modern studies of this discrepancy have provided considerable insight into how circumstances affect pain perception and, ultimately, into the anatomy and pharmacology of the pain system.

During World War II, Henry Beecher and his colleagues at Harvard Medical School made a fundamental observation. In the first systematic study of its kind, they found that soldiers suffering from severe battle wounds often experienced little or no pain. Indeed, many of the wounded expressed surprise at this odd dissociation. Beecher, an anesthesiologist, concluded that the perception of pain depends on its context. For instance, the pain of an injured soldier on the battlefield would presumably be mitigated by the imagined benefits of being removed from danger, whereas a similar injury in a domestic setting would present quite a different set of circumstances that could exacerbate the pain (loss of work, financial liability, and so on). Such observations, together with the well-known placebo effect (discussed in the next section), make clear that the perception of pain is subject to central modulation (all sensations are subject to at least some degree of this kind of modification). This statement, incidentally, should not be taken as a vague acknowledgment about the importance of psychological or “top-down” influences on sensory experience. On the contrary, there has been a gradual realization among neuroscientists and neurologists that such “psychological” effects are as real and important as any other neural phenomenon. This appreciation has provided a much more rational view of psychosomatic problems in general, and pain in particular.

The Placebo Effect

The placebo effect is defined as a physiological response following the administration of a pharmacologically inert “remedy.” The word *placebo* means “I will please,” and the placebo effect has a long history of use (and abuse) in medicine. The reality of the effect is undisputed. In one classic study, medical students were given one of two different pills, one said to be a sedative and the other a stimulant. In fact, both pills contained only inert ingredients. Of the students who received the “sedative,” more than two-thirds reported that they felt drowsy, and students who took two such pills felt sleepier than those who had taken only one. Conversely, a large fraction of the students who took the “stimulant” reported that they felt less tired. Moreover, about a third of the entire group reported side effects ranging from headaches and dizziness to tingling extremities and a staggering gait! Only 3 of the 56 students studied reported that the pills they took had no appreciable effect.

In another study of this general sort, 75% of patients suffering from postoperative wound pain reported satisfactory relief after an injection of sterile saline. The researchers who carried out this work noted that the responders were indistinguishable from the nonresponders, both in the apparent severity of their pain and psychological makeup. Most tellingly, this placebo effect in postoperative patients could be blocked by naloxone, a competitive antagonist of opiate receptors, indicating a substantial pharmacological basis for the pain relief experienced (see the next section). A common misunderstanding about the placebo effect is the view that patients who

respond to a therapeutically meaningless reagent are not suffering real pain, but only “imagining” it; this is certainly not the case.

Among other things, the placebo effect probably explains the efficacy of acupuncture anesthesia and the analgesia that can sometimes be achieved by hypnosis. In China, surgery has often been carried out under the effect of a needle (often carrying a small electrical current) inserted at locations dictated by ancient acupuncture charts. Before the advent of modern anesthetic techniques, operations such as thyroidectomies for goiter were commonly done without extraordinary discomfort, particularly among populations where stoicism was the cultural norm.

The mechanisms of pain amelioration on the battlefield, in acupuncture anesthesia, and with hypnosis are presumably related. Although the mechanisms by which the brain affects the perception of pain are only beginning to be understood, the effect is neither magical nor a sign of a suggestible intellect. In short, the placebo effect is quite real.

The Physiological Basis of Pain Modulation

Understanding the central modulation of pain perception (on which the placebo effect is presumably based) was greatly advanced by the finding that electrical or pharmacological stimulation of certain regions of the midbrain produces relief of pain. This analgesic effect arises from activation of descending pain-modulating pathways that project to the dorsal horn of the spinal cord (as well as to the spinal trigeminal nucleus) and regulate the transmission of information to higher centers (Figure 9.7A). One of the major brainstem regions that produce this effect is located in the periaqueductal gray of the midbrain. Electrical stimulation at this site in experimental animals not only produces analgesia by behavioral criteria, but also demonstrably inhibits the activity of nociceptive projection neurons in the dorsal horn of the spinal cord.

Further studies of descending pathways to the spinal cord that regulate the transmission of nociceptive information have shown that they arise from a number of brainstem sites, including the parabrachial nucleus, the dorsal raphe, and locus coeruleus and the medullary reticular formation (see Figure 9.7A). The analgesic effects of stimulating the periaqueductal gray are mediated through these brainstem sites. These centers employ a wealth of different neurotransmitters (noradrenaline, serotonin, dopamine, histamine, acetylcholine) and can exert both facilitatory and inhibitory effects on the activity of neurons in the dorsal horn. The complexity of these interactions is made even greater by the fact that descending projections can exert their effects on a variety of sites within the dorsal horn including the synaptic terminals of nociceptive afferents, excitatory and inhibitory interneurons, the synaptic terminals of the other descending pathways, as well as by contacting the projection neurons themselves. Although these descending projections were originally viewed as a mechanism that served primarily to inhibit the transmission of nociceptive signals, it is now evident that these projections provide a balance of facilitatory and inhibitory influences that ultimately determines the efficacy of nociceptive transmission.

In addition to descending projections, local interactions between mechanoreceptive afferents and neural circuits within the dorsal horn can modulate the transmission of nociceptive information to higher centers (Figure 9.7B). These interactions are thought to explain the ability to reduce the sensation of sharp pain by activating low-threshold mechanoreceptors: If you crack your shin or stub a toe, a natural (and effective) reaction is to vigor-

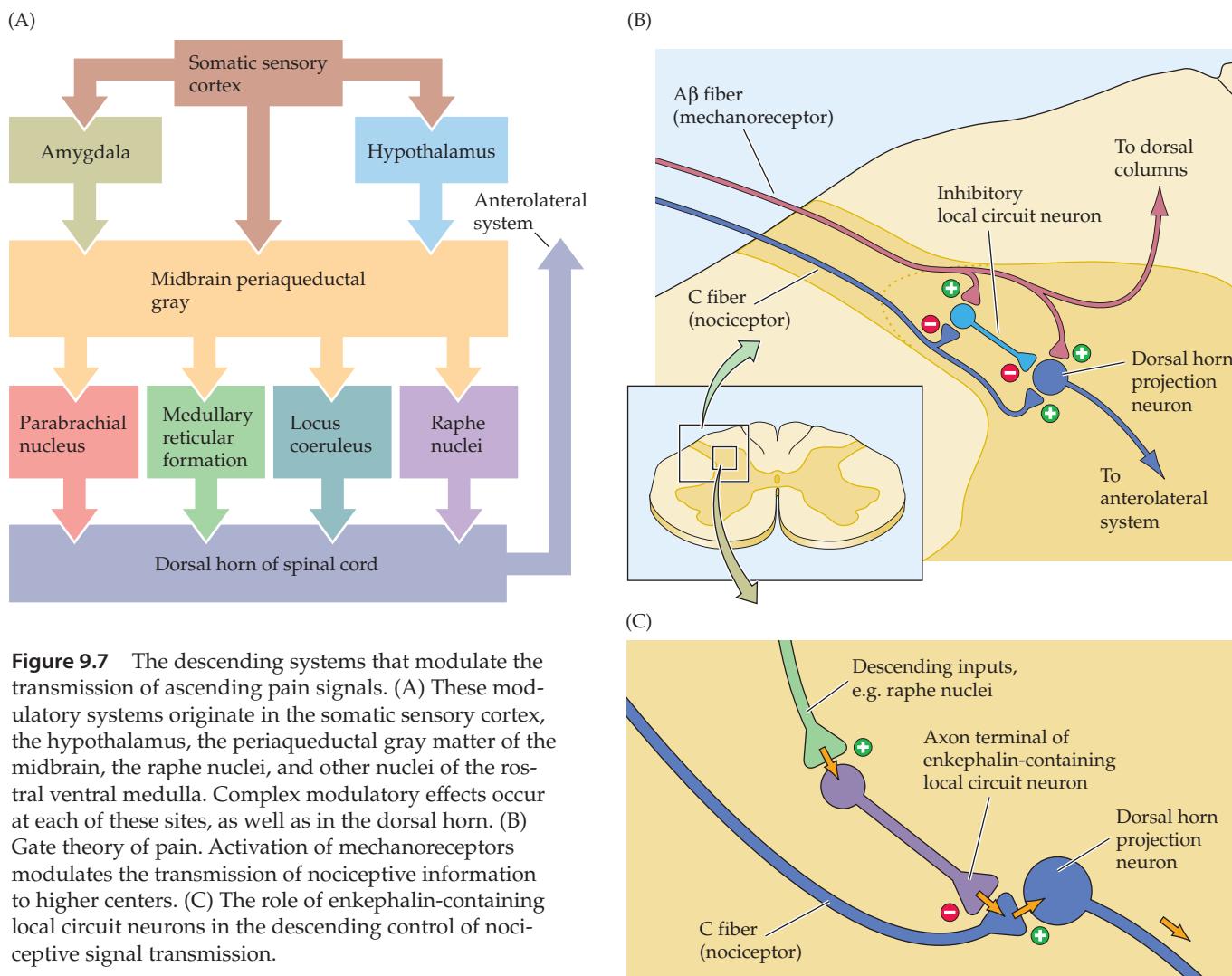


Figure 9.7 The descending systems that modulate the transmission of ascending pain signals. (A) These modulatory systems originate in the somatic sensory cortex, the hypothalamus, the periaqueductal gray matter of the midbrain, the raphe nuclei, and other nuclei of the rostral ventral medulla. Complex modulatory effects occur at each of these sites, as well as in the dorsal horn. (B) Gate theory of pain. Activation of mechanoreceptors modulates the transmission of nociceptive information to higher centers. (C) The role of enkephalin-containing local circuit neurons in the descending control of nociceptive signal transmission.



ously rub the site of injury for a minute or two. Such observations, buttressed by experiments in animals, led Ronald Melzack and Patrick Wall to propose that the flow of nociceptive information through the spinal cord is modulated by concomitant activation of the large myelinated fibers associated with low-threshold mechanoreceptors. Even though further investigation led to modification of some of the original propositions in Melzack and Wall's **gate theory of pain**, the idea stimulated a great deal of work on pain modulation and has emphasized the importance of synaptic interactions within the dorsal horn for modulating the perception of pain intensity.

The most exciting advance in this long-standing effort to understand central mechanisms of pain regulation has been the discovery of **endogenous opioids**. For centuries it had been apparent that opium derivatives such as morphine are powerful analgesics—indeed, they remain a mainstay of analgesic therapy today. Modern animal studies have shown that a variety of brain regions are susceptible to the action of opiate drugs, particularly—and significantly—the periaqueductal gray matter and other sources of descend-

ing projections. There are, in addition, opiate-sensitive neurons within the dorsal horn of the spinal cord. In other words, the areas that produce analgesia when stimulated are also responsive to exogenously administered opiates. It seems likely, then, that opiate drugs act at most or all of the sites shown in Figure 9.7 in producing their dramatic pain-relieving effects.

The analgesic action of opiates implied the existence of specific brain and spinal cord receptors for these drugs long before the receptors were actually found during the 1960s and 1970s. Since such receptors are unlikely to exist for the purpose of responding to the administration of opium and its derivatives, the conviction grew that there must be *endogenous* compounds for which these receptors had evolved (see Chapter 6). Several categories of endogenous opioids have now been isolated from the brain and intensively studied. These agents are found in the same regions that are involved in the modulation of nociceptive afferents, although each of the families of endogenous opioid peptides has a somewhat different distribution. All three of the major groups (**enkephalins**, **endorphins**, and **dynorphins**; see Table 6.2) are present in the periaqueductal gray matter. The enkephalins and dynorphins have also been found in the rostral ventral medulla and in the spinal cord regions involved in the modulation of pain.

One of the most compelling examples of the mechanism by which endogenous opiates modulate transmission of nociceptive information occurs at the first synapse in the pain pathway between nociceptive afferents and projection neurons in the dorsal horn of the spinal cord (see Figure 9.7B). A class of enkephalin-containing local circuit neurons within the dorsal horn synapses with the axon terminals of nociceptive afferents, which synapse in turn with dorsal horn projection neurons. The release of enkephalin onto the nociceptive terminals inhibits their release of neurotransmitter onto the projection neuron, reducing the level of activity that is passed on to higher centers. Enkephalin-containing local circuit neurons are themselves the targets of descending projections, thus providing a powerful mechanism by which higher centers can decrease the activity relayed by nociceptive afferents.

A particularly impressive aspect of this story is the wedding of physiology, pharmacology, and clinical research to yield a much richer understanding of the intrinsic modulation of pain. This information has finally begun to explain the subjective variability of painful stimuli and the striking dependence of pain perception on the context of the experience. Precisely how pain is modulated is being explored in many laboratories at present, motivated by the tremendous clinical (and economic) benefits that would accrue from still deeper knowledge of the pain system and its molecular underpinnings.

Summary

Whether from a structural or functional perspective, pain is an extraordinarily complex sensory modality. Because of the importance of warning an animal about dangerous circumstances, the mechanisms and pathways that subserve nociception are widespread and redundant. A distinct set of pain afferents with membrane receptors known as nociceptors transduces noxious stimulation and conveys this information to neurons in the dorsal horn of the spinal cord. The major central pathway responsible for transmitting the discriminative aspects of pain (location, intensity and quality) differs from the mechano-sensory pathway primarily in that the central axons of dorsal root ganglion cells synapse on second-order neurons in the dorsal horn; the axons of the second-order neurons then cross the midline in the spinal cord and ascend to

thalamic nuclei that relay information to the somatic sensory cortex of the postcentral gyrus. Additional pathways involving a number of centers in the brainstem, thalamus, and cortex mediate the affective and motivational responses to painful stimuli. Descending pathways interact with local circuits in the spinal cord to regulate the transmission of nociceptive signals to higher centers. Tremendous progress in understanding pain has been made in the last 25 years, and much more seems likely, given the importance of the problem. No patients are more distressed—or more difficult to treat—than those with chronic pain. Indeed, some aspects of pain seem much more destructive to the sufferer than required by any physiological purposes. Perhaps such seemingly excessive effects are a necessary but unfortunate by-product of the protective benefits of this vital sensory modality.

Additional Reading

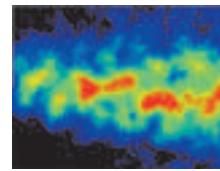
Reviews

- CATERINA, M. J. AND D. JULIUS (1999) Sense and specificity: A molecular identity for nociceptors. *Curr. Opin. Neurobiol.* 9: 525–530.
- DI MARZO, V., P. M. BLUMBERG AND A. SZALLASI (2002) Endovanilloid signaling in pain. *Curr. Opin. Neurobiol.* 12: 372–379.
- DUBNER, R. AND M. S. GOLD (1999) The neurobiology of pain. *Proc. Natl. Acad. Sci. USA* 96: 7627–7630.
- FIELDS, H. L. AND A. I. BASBAUM (1978) Brain-stem control of spinal pain transmission neurons. *Annu. Rev. Physiol.* 40: 217–248.
- HUNT, S. P. AND P. W. MANTYH (2001) The molecular dynamics of pain control. *Nat. Rev. Neurosci.* 2: 83–91.
- JI, R. R., T. KOHNO, K. A. MOORE AND C. J. WOOLF (2003) Central sensitization and LTP: Do pain and memory share similar mechanisms? *TINS* 26: 696–705.
- JULIUS, D. AND A. I. BASBAUM (2001) Molecular mechanisms of nociception. *Nature* 413: 203–209.
- MILLAN, M. J. (2002) Descending control of pain. *Prog. Neurobiol.* 66: 355–474.
- PATAPOUTIAN, A., A. M. PEIER, G. M. STORY AND V. VISWANATH (2003) ThermoTRP channels and beyond: Mechanisms of temperature sensation. *Nat. Rev. Neurosci.* 4: 529–539.
- RAINVILLE, P. (2002) Brain mechanisms of pain affect and pain modulation. *Curr. Opin. Neurobiol.* 12: 195–204.
- SCHOLZ, J. AND C. J. WOOLF (2002) Can we conquer pain? *Nat. Rev. Neurosci.* 5 (Suppl): 1062–1067.
- TREEDE, R. D., D. R. KENSHALO, R. H. GRACELY AND A. K. JONES (1999) The cortical representation of pain. *Pain* 79: 105–111.
- BASBAUM, A. I. AND H. L. FIELDS (1979) The origin of descending pathways in the dorsolateral funiculus of the spinal cord of the cat and rat: Further studies on the anatomy of pain modulation. *J. Comp. Neurol.* 187: 513–522.
- BEECHER, H. K. (1946) Pain in men wounded in battle. *Ann. Surg.* 123: 96.
- BLACKWELL, B., S. S. BLOOMFIELD AND C. R. BUNCHER (1972) Demonstration to medical students of placebo response and non-drug factors. *Lancet* 1: 1279–1282.
- CATERINA, M. J. AND 8 OTHERS (2000) Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science* 288: 306–313.
- CRAIG, A. D., M. C. BUSHNELL, E.-T. ZHANG AND A. BLOMQVIST (1994) A thalamic nucleus specific for pain and temperature sensation. *Nature* 372: 770–773.
- CRAIG, A. D., E. M. REIMAN, A. EVANS AND M. C. BUSHNELL (1996) Functional imaging of an illusion of pain. *Nature* 384: 258–260.
- LEVINE, J. D., H. L. FIELDS AND A. I. BASBAUM (1993) Peptides and the primary afferent nociceptor. *J. Neurosci.* 13: 2273–2286.
- MOGIL, J. S. AND J. E. GRISEL (1998) Transgenic studies of pain. *Pain* 77: 107–128.

Books

- FIELDS, H. L. (1987) *Pain*. New York: McGraw-Hill.
- FIELDS, H. L. (ed.) (1990) *Pain Syndromes in Neurology*. London: Butterworths.
- KOLB, L. C. (1954) *The Painful Phantom*. Springfield, IL: Charles C. Thomas.
- SKRABANEK, P. AND J. MCCORMICK (1990) *Follies and Fallacies in Medicine*. New York: Prometheus Books.
- WALL, P. D. AND R. MELZACK (1989) *Textbook of Pain*. New York: Churchill Livingstone.

Chapter 10



Vision: The Eye

Overview

The human visual system is extraordinary in the quantity and quality of information it supplies about the world. A glance is sufficient to describe the location, size, shape, color, and texture of objects and, if the objects are moving, their direction and speed. Equally remarkable is the fact that visual information can be discerned over a wide range of stimulus intensities, from the faint light of stars at night to bright sunlight. The next two chapters describe the molecular, cellular, and higher-order mechanisms that allow us to see. The first steps in the process of seeing involve transmission and refraction of light by the optics of the eye, the transduction of light energy into electrical signals by photoreceptors, and the refinement of these signals by synaptic interactions within the neural circuits of the retina.

Anatomy of the Eye

The eye is a fluid-filled sphere enclosed by three layers of tissue (Figure 10.1). Only the innermost layer of the eye, the **retina**, contains neurons that are sensitive to light and are capable of transmitting visual signals to central targets. The immediately adjacent layer of tissue includes three distinct but continuous structures collectively referred to as the **uveal tract**. The largest component of the uveal tract is the **choroid**, which is composed of a rich capillary bed (important for nourishing the photoreceptors of the retina) as well as a high concentration of the light absorbing pigment melanin. Extending from the choroid near the front of the eye is the **ciliary body**, a ring of tissue that encircles the lens and consists of a muscular component that is important for adjusting the refractive power of the lens, and a vascular component (the so-called ciliary processes) that produces the fluid that fills the front of the eye. The most anterior component of the uveal tract is the **iris**, the colored portion of the eye that can be seen through the cornea. It contains two sets of muscles with opposing actions, which allow the size of the **pupil** (the opening in its center) to be adjusted under neural control. The **sclera** forms the outermost tissue layer of the eye and is composed of a tough white fibrous tissue. At the front of the eye, however, this opaque outer layer is transformed into the **cornea**, a specialized transparent tissue that permits light rays to enter the eye.

Beyond the cornea, light rays pass through two distinct fluid environments before striking the retina. In the **anterior chamber**, just behind the cornea and in front of the lens, lies **aqueous humor**, a clear, watery liquid that supplies nutrients to both of these structures. Aqueous humor is produced by the ciliary processes in the **posterior chamber** (the region between

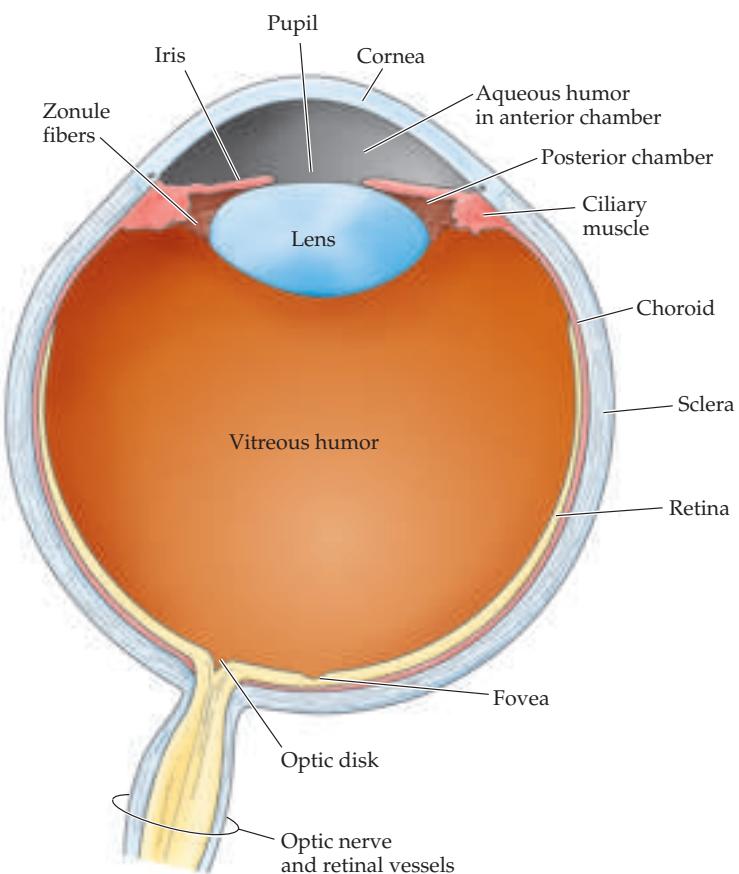


Figure 10.1 Anatomy of the human eye.

the lens and the iris) and flows into the anterior chamber through the pupil. The amount of fluid produced by the ciliary processes is substantial: it is estimated that the entire volume of fluid in the anterior chamber is replaced 12 times a day. Thus the rates of aqueous humor production must be balanced by comparable rates of drainage from the anterior chamber in order to ensure a constant intraocular pressure. A specialized meshwork of cells that lies at the junction of the iris and the cornea (a region called the **limbus**) is responsible for aqueous drainage. Failure of adequate drainage results in a disorder known as **glaucoma**, in which high levels of intraocular pressure can reduce the blood supply to the eye and eventually damage retinal neurons.

The space between the back of the lens and the surface of the retina is filled with a thick, gelatinous substance called the **vitreous humor**, which accounts for about 80% of the volume of the eye. In addition to maintaining the shape of the eye, the vitreous humor contains phagocytic cells that remove blood and other debris that might otherwise interfere with light transmission. The housekeeping abilities of the vitreous humor are limited, however, as a large number of middle-aged and elderly individuals with visual “floaters” will attest. Floaters are collections of debris too large for phagocytic consumption that therefore remain to cast annoying shadows on the retina; they typically arise when the aging vitreous membrane pulls away from the overly long eyeball of myopic individuals (Box A).

The Formation of Images on the Retina

Normal vision requires that the optical media of the eye be transparent, and both the **cornea** and the **lens** are remarkable examples of tissue specializations that achieve a level of transparency that rivals that found in inorganic materials such as glass. Not surprisingly, alterations in the composition of the cornea or the lens can significantly reduce their transparency and have serious consequences for visual perception. Indeed, **cataracts** (opacities in the lens) account for roughly half the cases of blindness in the world, and almost everyone over the age of 70 will experience some loss of transparency in the lens that ultimately degrades the quality of visual experience. Fortunately, there are successful surgical treatments for cataracts that can restore vision in most cases. Furthermore, the recognition that a major factor in the production of cataracts is exposure to ultraviolet (UV) solar radiation has heightened public awareness of the need to protect the lens (and the retina) by reducing UV exposure through the use of sunglasses.

Beyond efficiently transmitting light energy, the primary function of the optical components of the eye is to achieve a focused image on the surface of the retina. The cornea and the lens are primarily responsible for the refraction (bending) of light that is necessary for formation of focused images on the photoreceptors of the retina (Figure 10.2). The cornea contributes most of the necessary refraction, as can be appreciated by considering the hazy, out-of-focus images experienced when swimming underwater. Water, unlike air, has a refractive index close to that of the cornea; as a result, immersion in water virtually eliminates the refraction that normally occurs at the air/cornea interface; thus the image is no longer focused on the retina. The lens has considerably less refractive power than the cornea; however, the refraction supplied by the lens is adjustable, allowing objects at various distances from the observer to be brought into sharp focus.

Dynamic changes in the refractive power of the lens are referred to as **accommodation**. When viewing distant objects, the lens is made relatively thin and flat and has the least refractive power. For near vision, the lens becomes thicker and rounder and has the most refractive power (see Figure 10.2). These changes result from the activity of the **ciliary muscle** that surrounds the lens. The lens is held in place by radially arranged connective tissue bands (called zonule fibers) that are attached to the ciliary muscle. The shape of the lens is thus determined by two opposing forces: the elasticity of the lens, which tends to keep it rounded up (removed from the eye, the lens

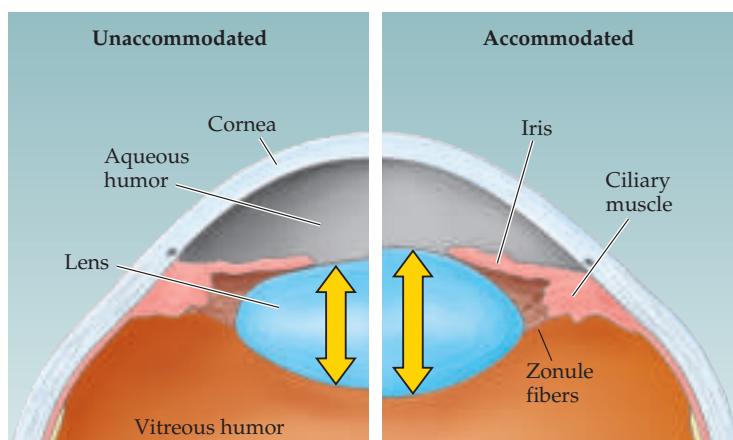


Figure 10.2 Diagram showing the anterior part of the human eye in the unaccommodated (left) and accommodated (right) state. Accommodation for focusing on near objects involves the contraction of the ciliary muscle, which reduces the tension in the zonule fibers and allows the elasticity of the lens to increase its curvature.

Box A

Myopia and Other Refractive Errors

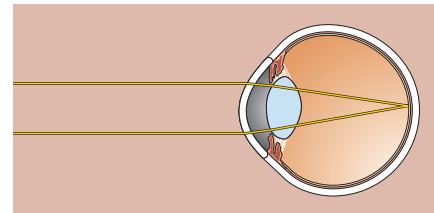
Optical discrepancies among the various components of the eye cause a majority of the human population to have some form of refractive error, called *ametropia*. People who are unable to bring distant objects into clear focus are said to be nearsighted, or myopic (Figures A and B). *Myopia* can be caused by the corneal surface being too curved, or by the eyeball being too long. In either case, with the lens as flat as it can be, the image of distant objects focuses in front of, rather than on, the retina. People who are unable to focus on near objects are said to be farsighted, or hyperopic. *Hyperopia* can be caused by the eyeball being too short or the refracting system too weak (Figure C). Even with the lens in its most rounded-up state, the image is out of focus on the retinal surface (focusing at some point behind it). Both myopia and hyperopia are correctable by appropriate lenses—concave (minus) and convex (plus), respectively—or by the increasingly popular technique of corneal surgery.

Myopia, or nearsightedness, is by far the most common ametropia; an estimated 50% of the population in the United States is affected. Given the large number of people who need glasses, contact lenses, or surgery to correct this refractive error, one naturally wonders how nearsighted people coped before spectacles were invented only a few centuries ago. From what is now known

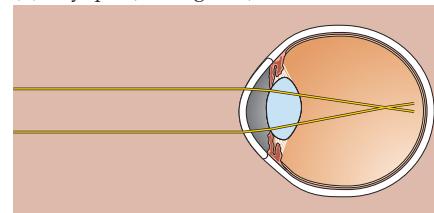
about myopia, most people's vision may have been considerably better in ancient times. The basis for this assertion is the surprising finding that the growth of the eyeball is strongly influenced by focused light falling on the retina. This phenomenon was first described in 1977 by Torsten Wiesel and Elio Raviola at Harvard Medical School, who studied monkeys reared with their lids sutured (the same approach used to demonstrate the effects of visual deprivation on cortical connections in the visual system; see Chapter 23), a procedure that deprives the eye of focused retinal images. They found that animals growing to maturity under these conditions show an elongation of the eyeball. The effect of focused light deprivation appears to be a local one, since the abnormal growth of the eye occurs in experimental animals even if the optic nerve is cut. Indeed, if only a portion of the retinal surface is deprived of focused light, then only that region of the eyeball grows abnormally.

Although the mechanism of light-mediated control of eye growth is not fully understood, many experts now believe that the prevalence of myopia is due to some aspect of modern civilization—perhaps learning to read and write at an early age—that interferes with the normal feedback control of vision on eye development, leading to abnormal elongation of the eyeball. A corollary of this hypothesis is that if children (or, more

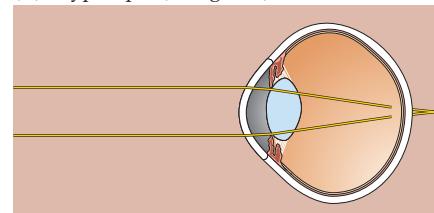
(A) Emmetropia (normal)



(B) Myopia (nearsighted)



(C) Hyperopia (farsighted)



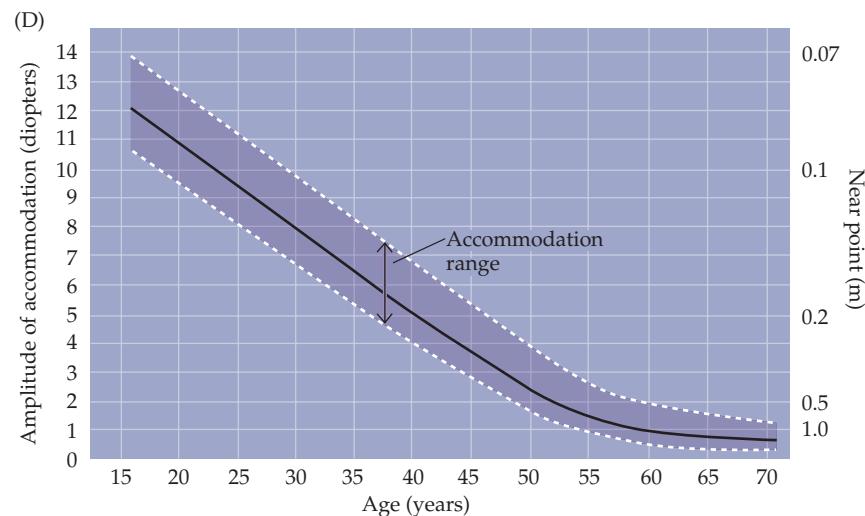
likely, their parents) wanted to improve their vision, they might be able to do so by practicing far vision to counterbalance the near work "overload." Practically, of

becomes spheroidal), and the tension exerted by the zonule fibers, which tends to flatten it. When viewing distant objects, the force from the zonule fibers is greater than the elasticity of the lens, and the lens assumes the flatter shape appropriate for distance viewing. Focusing on closer objects requires relaxing the tension in the zonule fibers, allowing the inherent elasticity of the lens to increase its curvature. This relaxation is accomplished by the sphincter-like contraction of the ciliary muscle. Because the ciliary muscle forms a ring around the lens, when the muscle contracts, the attachment points of the zonule fibers move toward the central axis of the eye, thus

(D) Changes in the ability of the lens to round up (accommodate) with age. The graph also shows how the near point (the closest point to the eye that can be brought into focus) changes. Accommodation, which is an optical measurement of the refractive power of the lens, is given in diopters. (After Westheimer, 1974.)

course, most people would probably choose wearing glasses or contacts or having corneal surgery rather than indulging in the onerous daily practice that would presumably be required. Not everyone agrees, however, that such a remedy would be effective, and a number of investigators (and drug companies) are exploring the possibility of pharmacological intervention during the period of childhood when abnormal eye growth is presumed to occur. In any event, it is a remarkable fact that deprivation of focused light on the retina causes a compensatory growth of the eye and that this feedback loop is so easily perturbed.

Even people with normal (emmetropic) vision as young adults eventually experience difficulty focusing on near objects. One of the many consequences of aging is that the lens loses its elasticity; as a result, the maximum curvature the lens can achieve when the ciliary muscle contracts is gradually reduced. The near point (the closest point that can be brought into clear focus) thus recedes, and objects (such as this book) must be farther and farther away from the eye in order to focus them on the retina. At some point, usually during



early middle age, the accommodative ability of the eye is so reduced that near vision tasks like reading become difficult or impossible (Figure D). This condition is referred to as presbyopia, and can be corrected by convex lenses for near-vision tasks, or by bifocal lenses if myopia is also present (which requires a negative correction). Bifocal correction presents a particular problem for those who prefer contact lenses. Because contact lenses float on the surface of the cornea, having the distance correction above and the near correction below (as in conventional bifocal glasses) doesn't work (although "omnifocal" contact lenses have recently been used with some success). A surprisingly effective solution to this problem for some contact lens wearers has been to put a near correcting lens in one eye and a distance correcting lens in the other! The success of this approach is another

testament to the remarkable ability of the visual system to adjust to a wide variety of unusual demands.

References

- BOCK, G. AND K. WIDDOWS (1990) *Myopia and the Control of Eye Growth*. Ciba Foundation Symposium 155. Chichester: Wiley.
- COSTER, D. J. (1994) *Physics for Ophthalmologists*. Edinburgh: Churchill Livingston.
- KAUFMAN, P. L. AND A. ALM (EDS.) (2002) *Adler's Physiology of the Eye: Clinical Application*, 10th Ed. St. Louis, MO: Mosby Year Book.
- SHERMAN, S. M., T. T. NORTON AND V. A. CASAGRANDE (1977) Myopia in the lid-sutured tree shrew. *Brain Res.* 124: 154–157.
- WALLMAN, J., J. TURKEL AND J. TRACTMAN (1978) Extreme myopia produced by modest changes in early visual experience. *Science* 201: 1249–1251.
- WIESEL, T. N. AND E. RAVIOLA (1977) Myopia and eye enlargement after neonatal lid fusion in monkeys. *Nature* 266: 66–68.

reducing the tension on the lens. Unfortunately, changes in the shape of the lens are not always able to produce a focused image on the retina, in which case a sharp image can be focused only with the help of additional corrective lenses (see Box A).

Adjustments in the size of the pupil also contribute to the clarity of images formed on the retina. Like the images formed by other optical instruments, those generated by the eye are affected by spherical and chromatic aberrations, which tend to blur the retinal image. Since these aberrations are greatest for light rays that pass farthest from the center of the lens, narrow-

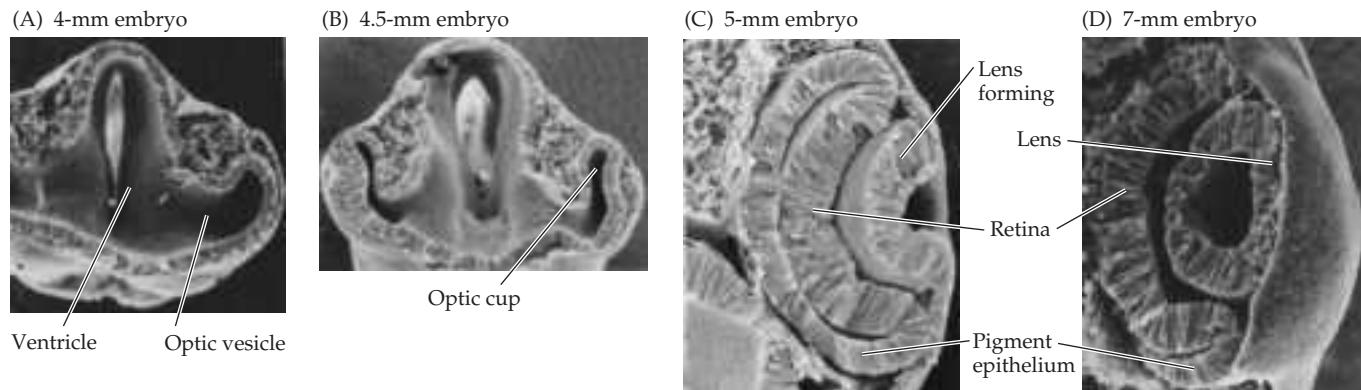
ing the pupil reduces both spherical and chromatic aberration, just as closing the iris diaphragm on a camera lens improves the sharpness of a photographic image. Reducing the size of the pupil also increases the depth of field—that is, the distance within which objects are seen without blurring. However, a small pupil also limits the amount of light that reaches the retina, and, under conditions of dim illumination, visual acuity becomes limited by the number of available photons rather than by optical aberrations. An adjustable pupil thus provides an effective means of reducing optical aberrations, while maximizing depth of field to the extent that different levels of illumination permit. The size of the pupil is controlled by innervation from both sympathetic and parasympathetic divisions of the visceral motor system, which are in turn modulated by several brainstem centers (see Chapters 19 and 20).

The Retina

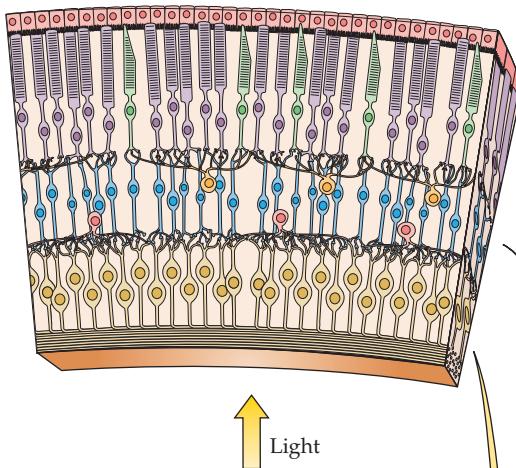
Despite its peripheral location, the **retina** or neural portion of the eye, is actually part of the central nervous system. During development, the retina forms as an outpocketing of the diencephalon, called the optic vesicle, which undergoes invagination to form the optic cup (Figure 10.3; see also Chapter 21). The inner wall of the optic cup gives rise to the retina, while the outer wall gives rise to the **retinal pigment epithelium**. This epithelium is a thin melanin-containing structure that reduces backscattering of light that enters the eye; it also plays a critical role in the maintenance of photoreceptors, renewing photopigments and phagocytosing the photoreceptor disks, whose turnover at a high rate is essential to vision.

Consistent with its status as a full-fledged part of the central nervous system, the retina comprises complex neural circuitry that converts the graded electrical activity of photoreceptors into action potentials that travel to the brain via axons in the optic nerve. Although it has the same types of functional elements and neurotransmitters found in other parts of the central nervous system, the retina comprises fewer classes of neurons, and these are arranged in a manner that has been less difficult to unravel than the circuits in other areas of the brain. There are five types of neurons in the retina: **photoreceptors, bipolar cells, ganglion cells, horizontal cells, and amacrine cells**. The cell bodies and processes of these neurons are stacked in alternating layers, with the cell bodies located in the inner nuclear, outer nuclear, and ganglion cell layers, and the processes and synaptic contacts located in the inner plexiform and outer plexiform layers (Figure 10.4). A direct three-

Figure 10.3 Development of the human eye. (A) The retina develops as an outpocketing from the neural tube, called the optic vesicle. (B) The optic vesicle invaginates to form the optic cup. (C, D) The inner wall of the optic cup becomes the neural retina, while the outer wall becomes the pigment epithelium. (A–C from Hilfer and Yang, 1980; D courtesy of K. Tosney.)



(A) Section of retina



(B)

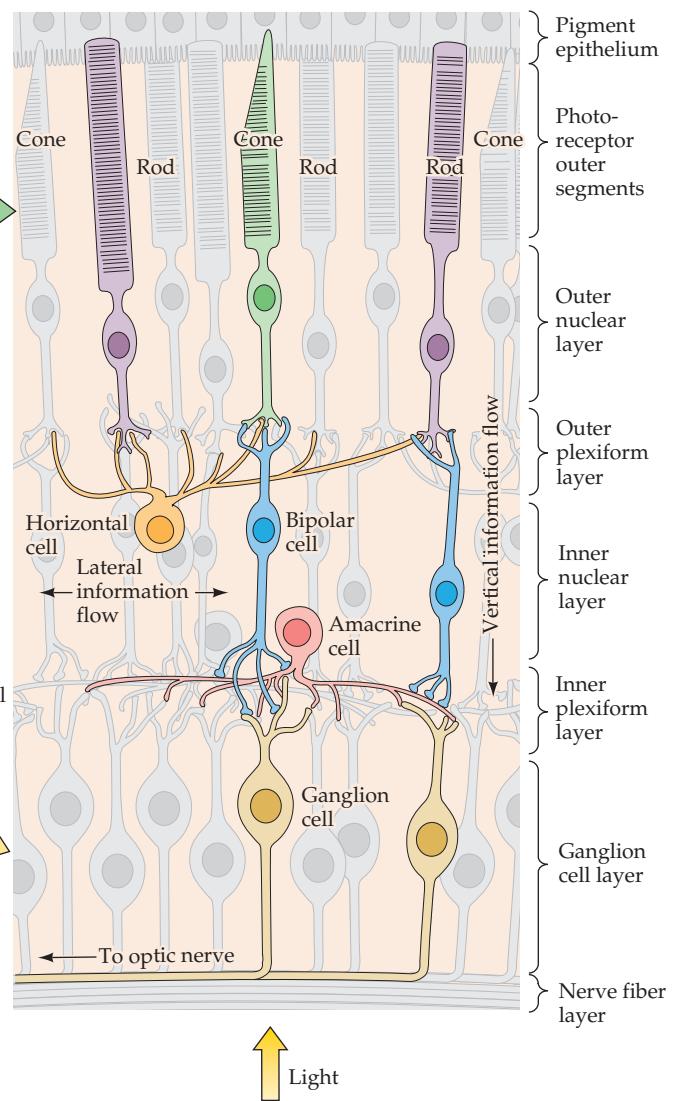


Figure 10.4 Structure of the retina. (A) Section of the retina showing overall arrangement of retinal layers. (B) Diagram of the basic circuitry of the retina. A three-neuron chain—photoreceptor, bipolar cell, and ganglion cell—provides the most direct route for transmitting visual information to the brain. Horizontal cells and amacrine cells mediate lateral interactions in the outer and inner plexiform layers, respectively. The terms *inner* and *outer* designate relative distances from the center of the eye (inner, near the center of the eye; outer, away from the center, or toward the pigment epithelium).

neuron chain—photoreceptor cell to bipolar cell to ganglion cell—is the major route of information flow from photoreceptors to the optic nerve.

There are two types of photoreceptors in the retina: **rods** and **cones**. Both types have an outer segment composed of membranous disks that contain light-sensitive photopigment and lies adjacent to the pigment epithelium, and an inner segment that contains the cell nucleus and gives rise to synaptic terminals that contact bipolar or horizontal cells (see also Figure 10.8). Absorption of light by the photopigment in the outer segment of the photoreceptors initiates a cascade of events that changes the membrane potential of the receptor, and therefore the amount of neurotransmitter released by the photoreceptor synapses onto the cells they contact. The synapses between photoreceptor terminals and bipolar cells (and horizontal cells) occur in the outer plexiform layer; more specifically, the cell bodies of photoreceptors make up the outer nuclear layer, whereas the cell bodies of bipolar cells lie in the inner nuclear layer. The short axonal processes of bipolar cells make synaptic contacts in turn on the dendritic processes of ganglion cells in the inner plexiform layer. The much larger axons of the ganglion cells form the **optic**

nerve and carry information about retinal stimulation to the rest of the central nervous system.

The two other types of neurons in the retina, **horizontal cells** and **amacrine cells**, have their cell bodies in the inner nuclear layer and have processes that are limited to the outer and inner plexiform layers respectively (see Figure 10.4). The processes of horizontal cells enable lateral interactions between photoreceptors and bipolar cells that maintain the visual system's sensitivity to luminance contrast over a wide range of light intensities. The processes of amacrine cells are postsynaptic to bipolar cell terminals and presynaptic to the dendrites of ganglion cells. Different subclasses of amacrine cells are thought to make distinct contributions to visual function. One class of amacrine cells, for example, plays an important role in transforming the sustained responses of bipolar cells to step changes in light intensity into transient onset or offset responses exhibited by some types of ganglion cells. Another type serves as an obligatory step in the pathway that transmits information from rod photoreceptors to retinal ganglion cells. The variety of amacrine cell subtypes illustrates the more general rule that although there are only five basic retinal cell types, there can be considerable diversity within a given cell type. This diversity is also a hallmark of retinal ganglion cells and the basis for pathways that convey different sorts of information to central targets in a parallel manner (see Chapter 11).

At first glance, the spatial arrangement of retinal layers seems counterintuitive, since light rays must pass through various non-light-sensitive elements of the retina as well as the retinal vasculature (which branches extensively on the inner surface of the retina—see Figure 11.1) before reaching the outer segments of the photoreceptors, where photons are absorbed (Figure 10.4). The reason for this curious feature of retinal organization lies in the special relationship that exists among the outer segments of the photoreceptors, the pigment epithelium, and the underlying choroid. Recall that the outer segments contain membranous disks that house the light-sensitive photopigment and other proteins involved in the transduction process. These disks are formed near the inner segment of the photoreceptor and move toward the tip of the outer segment, where they are shed. The pigment epithelium plays an essential role in removing the expended receptor disks; this is no small task, since all the disks in the outer segments are replaced every 12 days. In addition, the pigment epithelium contains the biochemical machinery that is required to regenerate photopigment molecules after they have been exposed to light. Finally, the capillaries in the choroid underlying the pigment epithelium are the primary source of nourishment for retinal photoreceptors. These functional considerations presumably explain why rods and cones are found in the outermost rather than the innermost layer of the retina. They also explain why disruptions in the normal relationships between the pigment epithelium and retinal photoreceptors such as those that occur in retinitis pigmentosa have severe consequences for vision (Box B).

Phototransduction

In most sensory systems, activation of a receptor by the appropriate stimulus causes the cell membrane to depolarize, ultimately stimulating an action potential and transmitter release onto the neurons it contacts. In the retina, however, photoreceptors do not exhibit action potentials; rather, light activation causes a graded change in membrane potential and a corresponding change in the rate of transmitter release onto postsynaptic neurons. Indeed,

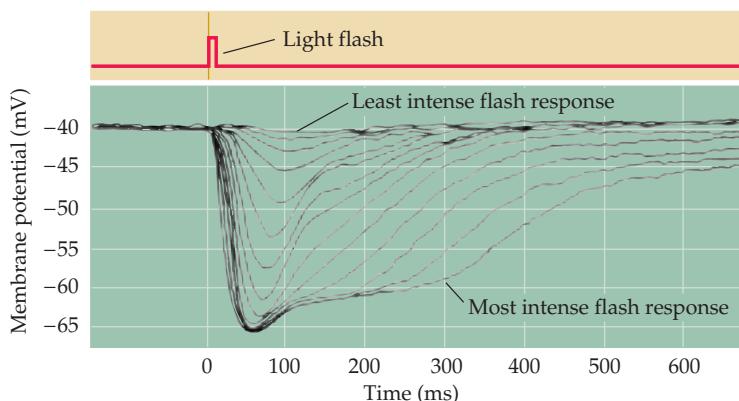


Figure 10.5 An intracellular recording from a single cone stimulated with different amounts of light (the cone has been taken from the turtle retina, which accounts for the relatively long time course of the response). Each trace represents the response to a brief flash that was varied in intensity. At the highest light levels, the response amplitude saturates (at about -65 mV). The hyperpolarizing response is characteristic of vertebrate photoreceptors; interestingly, some invertebrate photoreceptors depolarize in response to light. (After Schnapf and Baylor, 1987.)

much of the processing within the retina is mediated by graded potentials, largely because action potentials are not required to transmit information over the relatively short distances involved.

Perhaps even more surprising is that shining light on a photoreceptor, either a rod or a cone, leads to membrane *hyperpolarization* rather than depolarization (Figure 10.5). In the dark, the receptor is in a depolarized state, with a membrane potential of roughly -40 mV (including those portions of the cell that release transmitters). Progressive increases in the intensity of illumination cause the potential across the receptor membrane to become more negative, a response that saturates when the membrane potential reaches about -65 mV . Although the sign of the potential change may seem odd, the only logical requirement for subsequent visual processing is a consistent relationship between luminance changes and the rate of transmitter release from the photoreceptor terminals. As in other nerve cells, transmitter release from the synaptic terminals of the photoreceptor is dependent on voltage-sensitive Ca^{2+} channels in the terminal membrane. Thus, in the dark, when photoreceptors are relatively depolarized, the number of open Ca^{2+} channels in the synaptic terminal is high, and the rate of transmitter release is correspondingly great; in the light, when receptors are hyperpolarized, the number of open Ca^{2+} channels is reduced, and the rate of transmitter release is also reduced. The reason for this unusual arrangement compared to other sensory receptor cells is not known.

The relatively depolarized state of photoreceptors in the dark depends on the presence of ion channels in the outer segment membrane that permit Na^+ and Ca^{2+} ions to flow into the cell, thus reducing the degree of inside negativity (Figure 10.6). The probability of these channels in the outer segment being open or closed is regulated in turn by the levels of the nucleotide cyclic guanosine monophosphate (cGMP) (as in many other second messenger systems; see Chapter 7). In darkness, high levels of cGMP in the outer segment keep the channels open. In the light, however, cGMP levels drop and some of the channels close, leading to hyperpolarization of the outer segment membrane, and ultimately the reduction of transmitter release at the photoreceptor synapse.

The series of biochemical changes that ultimately leads to a reduction in cGMP levels begins when a photon is absorbed by the photopigment in the receptor disks. The photopigment contains a light-absorbing chromophore (**retinal**, an aldehyde of vitamin A) coupled to one of several possible proteins called **opsins** that tune the molecule's absorption of light to a particular region of the spectrum. Indeed, it is the different protein component of

Figure 10.6 Cyclic GMP-gated channels in the outer segment membrane are responsible for the light-induced changes in the electrical activity of photoreceptors (a rod is shown here, but the same scheme applies to cones). In the dark, cGMP levels in the outer segment are high; this molecule binds to the Na^+ -permeable channels in the membrane, keeping them open and allowing sodium (and other cations) to enter, thus depolarizing the cell. Exposure to light leads to a decrease in cGMP levels, a closing of the channels, and receptor hyperpolarization.

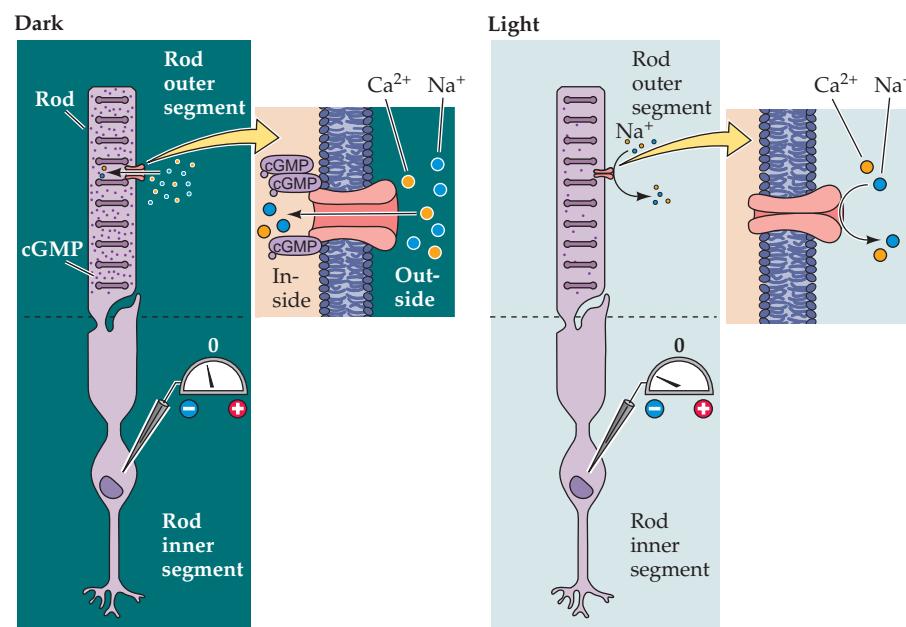
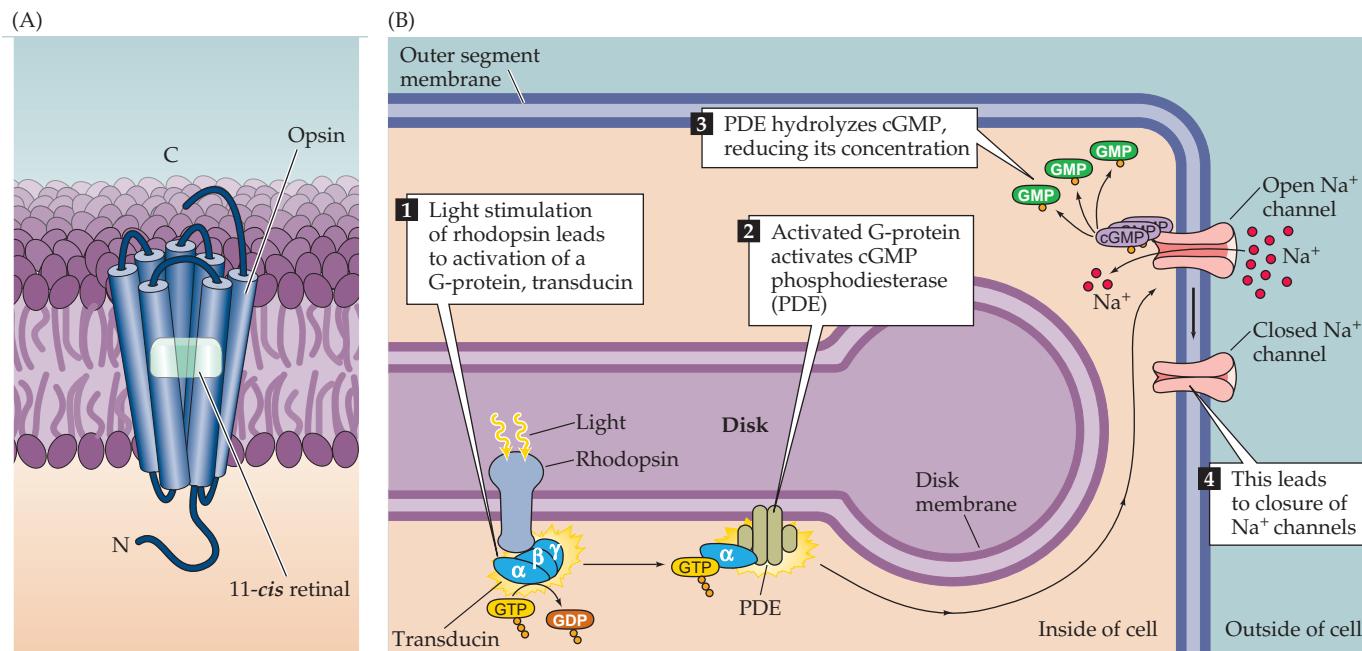


Figure 10.7 Details of phototransduction in rod photoreceptors. (A) The molecular structure of rhodopsin, the pigment in rods. (B) The second messenger cascade of phototransduction. Light stimulation of rhodopsin in the receptor disks leads to the activation of a G-protein (transducin), which in turn activates a phosphodiesterase (PDE). The phosphodiesterase hydrolyzes cGMP, reducing its concentration in the outer segment and leading to the closure of sodium channels in the outer segment membrane.

the photopigment in rods and cones that contributes to the functional specialization of these two receptor types. Most of what is known about the molecular events of phototransduction has been gleaned from experiments in rods, in which the photopigment is **rhodopsin** (Figure 10.7A). When the retinal moiety in the rhodopsin molecule absorbs a photon, its configuration changes from the 11-cis isomer to all-trans retinal; this change then triggers a series of alterations in the protein component of the molecule (Figure 10.7B). The changes lead, in turn, to the activation of an intracellular messenger called **transducin**, which activates a phosphodiesterase that hydrolyzes



Box B

Retinitis Pigmentosa

Retinitis pigmentosa (RP) refers to a heterogeneous group of hereditary eye disorders characterized by progressive vision loss due to a gradual degeneration of photoreceptors. An estimated 100,000 people in the United States have RP. In spite of the name, inflammation is not a prominent part of the disease process; instead the photoreceptor cells appear to die by apoptosis (determined by the presence of DNA fragmentation).

Classification of this group of disorders under one rubric is based on the clinical features commonly observed in these patients. The hallmarks of RP are night blindness, a reduction of peripheral vision, narrowing of the retinal vessels, and the migration of pigment from disrupted retinal pigment epithelium into the retina, forming clumps of various sizes, often next to retinal blood vessels (see figure).

Typically, patients first notice difficulty seeing at night due to the loss of rod photoreceptors; the remaining cone

photoreceptors then become the main-stay of visual function. Over many years, the cones also degenerate, leading to a progressive loss of vision. In most RP patients, visual field defects begin in the midperiphery, between 30° and 50° from the point of foveal fixation. The defective regions gradually enlarge, leaving islands of vision in the periphery and a constricted central field—a condition known as tunnel vision. When the visual field contracts to 20° or less and/or central vision is 20/200 or worse, the patient is categorized as legally blind.

Inheritance patterns indicate that RP can be transmitted in an X-linked (XLRP), autosomal dominant (ADRP), or recessive (ARRP) manner. In the United States, the percentage of these genetic types is estimated to be 9%, 16%, and 41%, respectively. When only one member of a pedigree has RP, the case is classified as “simplex,” which accounts for about a third of all cases.

Among the three genetic types of RP, ADRP is the mildest. These patients often retain good central vision until 60 years of age or older. In contrast, patients with the XLRP form of the disease are usually legally blind by 30 to 40 years of age. However, the severity and age of onset of the symptoms varies greatly among patients with the same type of RP, and even within the same family (when, presumably, all the affected members have the same genetic mutation).

To date, RP-inducing mutations of 30 genes have been identified. Many of these genes encode photoreceptor-specific proteins, several being associated with phototransduction in the rods. Among the latter are genes for rhodopsin, subunits of the cGMP phosphodiesterase, and the cGMP-gated

channel. Multiple mutations have been found in each of these cloned genes. For example, in the case of the rhodopsin gene, 90 different mutations have been identified among ADRP patients.

The heterogeneity of RP at all levels, from genetic mutations to clinical symptoms, has important implications for understanding the pathogenesis of the disease and designing therapies. Given the complex molecular etiology of RP, it is unlikely that a single cellular mechanism will explain the disease in all cases. Regardless of the specific mutation or causal sequence, the vision loss that is most critical to RP patients is due to the gradual degeneration of cones. In many cases, the protein that the RP-causing mutation affects is not even expressed in the cones; the prime example is rhodopsin—the rod-specific visual pigment. Therefore, the loss of cones may be an indirect result of a rod-specific mutation. In consequence, understanding and treating this disease presents a particularly difficult challenge.

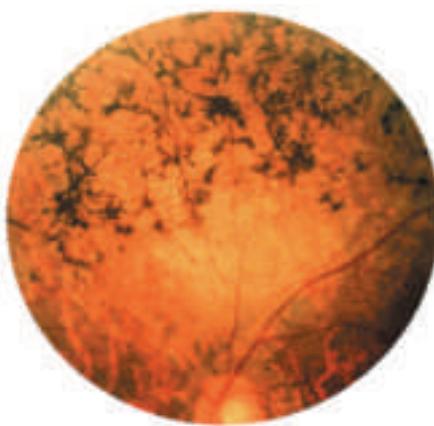
References

WELEBER, R. G. AND K. GREGORY-EVANS (2001) Retinitis pigmentosa and allied disorders. In *Retina*, 3rd Ed., Vol. 1: *Basic Science and Inherited Retinal Diseases*. S. J. Ryan (ed. in chief). St. Louis, MO: Mosby Year Book, pp. 362–460.

RATTNER, A., A. SUN AND J. NATHANS (1999) Molecular genetics of human retinal disease. *Annu. Rev. Genet.* 33: 89–131.

THE FOUNDATION FIGHTING BLINDNESS of Hunt Valley, MD, maintains a web site that provides updated information about many forms of retinal degeneration: www.blindness.org

RETNET provides updated information, including references to original articles, on genes and mutations associated with retinal diseases: www.sph.uth.tmc.edu/RetNet



Characteristic appearance of the retina in patients with retinitis pigmentosa. Note the dark clumps of pigment that are the hallmark of this disorder.

cGMP. All of these events take place within the disk membrane. The hydrolysis by phosphodiesterase at the disk membrane lowers the concentration of cGMP throughout the outer segment, and thus reduces the number of cGMP molecules that are available for binding to the channels in the surface of the outer segment membrane, leading to channel closure.

One of the important features of this complex biochemical cascade initiated by photon capture is that it provides enormous signal amplification. It has been estimated that a single light-activated rhodopsin molecule can activate 800 transducin molecules, roughly eight percent of the transducin molecules on the disk surface. Although each transducin molecule activates only one phosphodiesterase molecule, each of these is in turn capable of catalyzing the breakdown of as many as six cGMP molecules. As a result, the absorption of a single photon by a rhodopsin molecule results in the closure of approximately 200 ion channels, or about 2% of the number of channels in each rod that are open in the dark. This number of channel closures causes a net change in the membrane potential of about 1 mV.

Equally important is the fact that the magnitude of this amplification varies with the prevailing levels of illumination, a phenomenon known as **light adaptation**. At low levels of illumination, photoreceptors are the most sensitive to light. As levels of illumination increase, sensitivity decreases, preventing the receptors from saturating and thereby greatly extending the range of light intensities over which they operate. The concentration of Ca^{2+} in the outer segment appears to play a key role in the light-induced modulation of photoreceptor sensitivity. The cGMP-gated channels in the outer segment are permeable to both Na^+ and Ca^{2+} ; thus, light-induced closure of these channels leads to a net decrease in the internal Ca^{2+} concentration. This decrease triggers a number of changes in the phototransduction cascade, all of which tend to reduce the sensitivity of the receptor to light. For example, the decrease in Ca^{2+} increases the activity of quanlyate cyclase, the cGMP synthesizing enzyme, leading to an increase in cGMP levels. Likewise, the decrease in Ca^{2+} increases the affinity of the cGMP-gated channels for cGMP, reducing the impact of the light-induced reduction of cGMP levels. The regulatory effects of Ca^{2+} on the phototransduction cascade are only one part of the mechanism that adapts retinal sensitivity to background levels of illumination; another important contribution comes from neural interactions between horizontal cells and photoreceptor terminals (see below).

Once initiated, additional mechanisms limit the duration of this amplifying cascade and restore the various molecules to their inactivated states. The protein **arrestin**, for instance, blocks the ability of activated rhodopsin to activate transducin, and facilitates the breakdown of activated rhodopsin. The *all-trans* retinal then dissociates from the opsin, diffuses into the cytosol of the outer segment, is converted to *all-trans* retinol and is transported out of the outer segment and into the pigment epithelium, where appropriate enzymes ultimately convert it to *11-cis* retinal. After it is transported back into the outer segment, the *11-cis* retinal recombines with opsin in the receptor disks. The recycling of rhodopsin is critically important for maintaining the light sensitivity of photoreceptors. Even under intense levels of illumination, the rate of regeneration is sufficient to maintain a significant number of active photopigment molecules.

Functional Specialization of the Rod and Cone Systems

The two types of photoreceptors, rods and cones, are distinguished by shape (from which they derive their names), the type of photopigment they con-

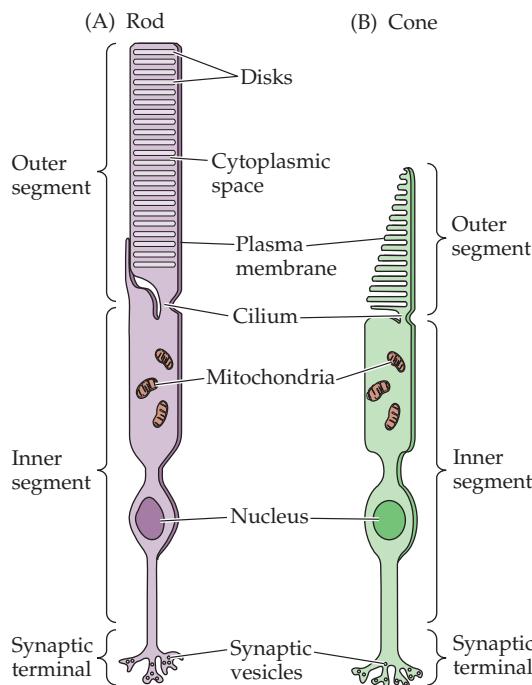


Figure 10.8 Structural differences between rods and cones. Although generally similar in structure, rods (A) and cones (B) differ in their size and shape, as well as in the arrangement of the membranous disks in their outer segments.

tain, distribution across the retina, and pattern of synaptic connections (Figure 10.8). These properties reflect the fact that the rod and cone systems (the receptors and their connections within the retina) are specialized for different aspects of vision. The rod system has very low spatial resolution but is extremely sensitive to light; it is therefore specialized for sensitivity at the expense of resolution. Conversely, the cone system has very high spatial resolution but is relatively insensitive to light; it is therefore specialized for acuity at the expense of sensitivity. The properties of the cone system also allow humans and many other animals to see color.

The range of illumination over which the rods and cones operate is shown in Figure 10.9. At the lowest levels of light, only the rods are activated. Such rod-mediated perception is called **scotopic vision**. The difficulty of making fine visual discriminations under very low light conditions where only the rod system is active is a common experience. The problem is primarily the poor resolution of the rod system (and, to a lesser degree, the fact that there is no perception of color in dim light because the cones are not involved to a significant degree). Although cones begin to contribute to visual perception at about the level of starlight, spatial discrimination at this light level is still very poor. As illumination increases, cones become more and more dominant in determining what is seen, and they are the major determinant of perception under relatively bright conditions such as normal indoor lighting or sunlight. The contributions of rods to vision drops out nearly entirely in so-called **photopic vision** because their response to light saturates—that is, the membrane potential of individual rods no longer varies as a function of illumination because all of the membrane channels are closed (see Figure 10.5). **Mesopic vision** occurs in levels of light at which both rods and cones contribute—at twilight, for example. From these considerations it should be clear that most of what we think of as normal “seeing” is mediated by the cone system, and that loss of cone function is devastating, as occurs in

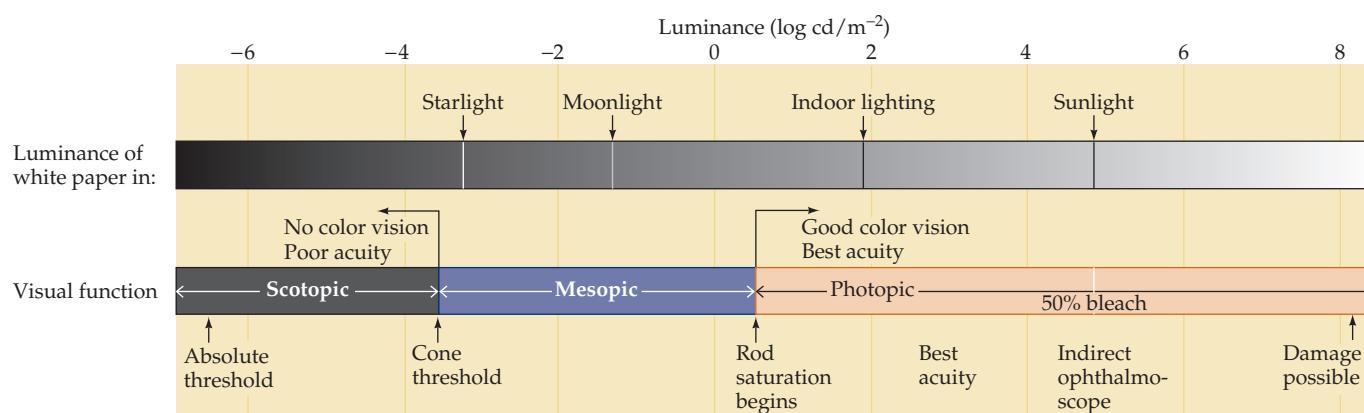


Figure 10.9 The range of luminance values over which the visual system operates. At the lowest levels of illumination, only rods are activated. Cones begin to contribute to perception at about the level of starlight and are the only receptors that function under relatively bright conditions.

elderly individuals suffering from macular degeneration (Box C). People who have lost cone function are legally blind, whereas those who have lost rod function only experience difficulty seeing at low levels of illumination (night blindness; see Box B).

Differences in the transduction mechanisms utilized by the two receptor types is a major factor in the ability of rods and cones to respond to different ranges of light intensity. For example, rods produce a reliable response to a single photon of light, whereas more than 100 photons are required to produce a comparable response in a cone. It is not, however, that cones fail to effectively capture photons. Rather, the change in current produced by single photon capture in cones is comparatively small and difficult to distinguish from noise. Another difference is that the response of an individual cone does not saturate at high levels of steady illumination, as does the rod response. Although both rods and cones adapt to operate over a range of luminance values, the adaptation mechanisms of the cones are more effective. This difference in adaptation is apparent in the time course of the response of rods and cones to light flashes. The response of a cone, even to a bright light flash that produces the maximum change in photoreceptor current, recovers in about 200 milliseconds, more than four times faster than rod recovery.

The arrangement of the circuits that transmit rod and cone information to retinal ganglion cells also contributes to the different characteristics of scotopic and photopic vision. In most parts of the retina, rod and cone signals converge on the same ganglion cells; i.e., individual ganglion cells respond to both rod and cone inputs, depending on the level of illumination. The early stages of the pathways that link rods and cones to ganglion cells, however, are largely independent. For example, the pathway from rods to ganglion cells involves a distinct class of bipolar cell (called rod bipolar) that, unlike cone bipolar cells, does not contact retinal ganglion cells. Instead, rod bipolar cells synapse with the dendritic processes of a specific class of amacrine cell that makes gap junctions and chemical synapses with the terminals of cone bipolars; these processes, in turn, make synaptic contacts on the dendrites of ganglion cells in the inner plexiform layer. As a consequence, the circuits linking the rods and cones to retinal ganglion cells differ dramatically in their degree of convergence. Each rod bipolar cell is contacted by a number of rods, and many rod bipolar cells contact a given amacrine cell. In contrast, the cone system is much less convergent. Thus, each retinal ganglion cell that dominates central vision (called midget gan-

Box C

Macular Degeneration

An estimated six million people in the United States suffer from a condition known as **age-related macular degeneration (AMD)**, which causes a progressive loss of central vision. Since central vision is critical for sight, diseases that affect the macula (see Figure 11.1) severely limit the ability to perform visual tasks. Indeed, AMD is the most common cause of vision loss in people over age 55, and its incidence is rising with the increasing percentage of elderly individuals in the population.

The underlying problem, which remains poorly understood, is degeneration of the photoreceptors. Usually, patients first notice a blurring of central vision when performing tasks such as reading. Images may also appear distorted. A graph paper-like chart known as the Amsler grid is used as a simple test for early signs of AMD. By focusing on a marked spot in the middle of the grid, the patient can assess whether the parallel and perpendicular lines on the grid appear blurred or distorted. Blurred central vision often progresses to having blind spots within central vision, and in most cases both eyes are eventually involved.

Although the risk of developing AMD clearly increases with age, the causes of the disease are not known. Various studies have implicated hereditary factors, cardiovascular disease, environmental factors such as smoking and light exposure, and nutritional causes. Indeed, it may be that all these contribute to the risk of developing AMD.

In descriptive terms, macular degeneration is broadly divided into two types. In the *exudative-neovascular form*, or "wet" AMD, which accounts for 10% of all cases, abnormal blood vessel growth occurs under the macula. These blood vessels leak fluid and blood into the retina and cause damage to the photore-

ceptors. Wet AMD tends to progress rapidly and can cause severe damage; rapid loss of central vision may occur over just a few months. The treatment for this form of the disease is laser therapy. By transferring thermal energy, the laser beam destroys the leaky blood vessels under the macula, thus slowing the rate of vision loss. A disadvantage of this approach is that the high thermal energy delivered by the beam also destroys nearby healthy tissue. An improvement in the laser treatment of AMD involves a light-activated drug to target abnormal blood vessels. Once the drug is administered, relatively low energy laser pulses aimed at the abnormal blood vessels are delivered to stimulate the drug, which in turn destroys the abnormal blood vessels with minimal damage to the surrounding tissue.

The remaining 90% of AMD cases are the nonexudative, or "dry" form. In these patients there is a gradual disappearance of the retinal pigment epithelium, resulting in circumscribed areas of atrophy. Since photoreceptor loss follows the disappearance of the pigment epithelium, the affected retinal areas have little or no visual function. Vision loss from dry AMD occurs more gradually, typically over the course of many years. These patients usually retain some central vision, although the loss can be severe enough to compromise performance of tasks that require seeing details. Unfortunately, at the present time there is no treatment for dry AMD. A radical and quite fascinating new approach that offers some promise entails surgically repositioning the retina away from the abnormal area.

Occasionally, macular degeneration occurs in much younger individuals. Many of these cases are caused by various mutations, each with its own clinical manifestations and genetic cause. The

most common form of juvenile macular degeneration is known as *Stargardt disease*, which is inherited as an autosomal recessive. Patients are usually diagnosed before they reach the age of 20. Although the progression of vision loss is variable, most of these patients are legally blind by age 50. Mutations that cause Stargardt disease have been identified in the *ABCR* gene, which codes for a protein that transports retinoids across the photoreceptor membrane. Thus, the visual cycle of photopigment regeneration may be disrupted in this form of macular degeneration, presumably by dysfunctional proteins encoded by the abnormal gene. Interestingly, the *ABCR* gene is expressed only in rods, suggesting that the cones may have their own visual cycle enzymes.

References

- FINE, S. L., J. W. BERGER, M. G. MAGUIRE AND A. C. HO (2000) Drug therapy: Age-related macular degeneration. *NEJM* 342: 483–492.
- SARKS, S. H. AND J. P. SARKS (2001) Age-related macular degeneration—atrophic form. In *Retina*, 3rd Ed., Vol. 2: *Medical Retina*. S. J. Ryan (ed.-in-chief). St. Louis, MO: Mosby Year Book, pp. 1071–1102.
- ELMAN, M. J. AND S. L. FINE (2001) Exudative age-related macular degeneration. In *Retina*, 3rd Ed., Vol. 2: *Medical Retina*. S. J. Ryan (ed.-in-chief). St. Louis, MO: Mosby Year Book, pp. 1103–114.
- DEUTMAN, A. F. (2001) Macular dystrophies. In *Retina*, 3rd Ed., Vol. 2: *Medical Retina*. S. J. Ryan (ed.-in-chief). St. Louis, MO: Mosby Year Book, pp. 1186–1240.
- THE FOUNDATION FIGHTING BLINDNESS of Hunt Valley, MD, maintains a web site that provides updated information about many forms of retinal degeneration: www.blindness.org
- RETNET provides updated information, including references to original articles, on genes and mutations associated with retinal diseases: www.sph.uth.tmc.edu/RetNet

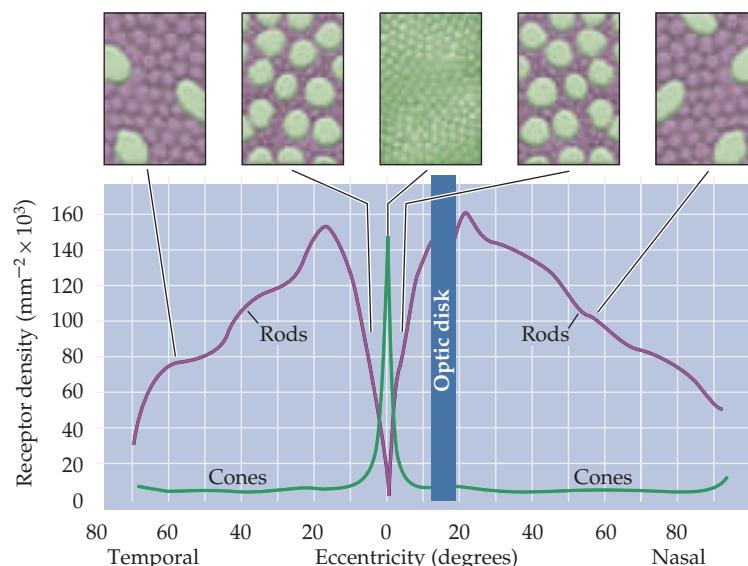
glion cells) receives input from only one cone bipolar cell, which, in turn, is contacted by a single cone. Convergence makes the rod system a better detector of light, because small signals from many rods are pooled to generate a large response in the bipolar cell. At the same time, convergence reduces the spatial resolution of the rod system, since the source of a signal in a rod bipolar cell or retinal ganglion cell could have come from anywhere within a relatively large area of the retinal surface. The one-to-one relationship of cones to bipolar and ganglion cells is, of course, just what is required to maximize acuity.

Anatomical Distribution of Rods and Cones

The distribution of rods and cones across the surface of the retina also has important consequences for vision (Figure 10.10). Despite the fact that perception in typical daytime light levels is dominated by cone-mediated vision, the total number of rods in the human retina (about 90 million) far exceeds the number of cones (roughly 4.5 million). As a result, the density of rods is much greater than cones throughout most of the retina. However, this relationship changes dramatically in the **fovea**, a highly specialized region of the central retina that measures about 1.2 millimeters in diameter (Figure 10.11). In the fovea, cone density increases almost 200-fold, reaching, at its center, the highest receptor packing density anywhere in the retina. This high density is achieved by decreasing the diameter of the cone outer segments such that foveal cones resemble rods in their appearance. The increased density of cones in the fovea is accompanied by a sharp decline in the density of rods. In fact, the central 300 μm of the fovea, called the **foveola**, is totally rod-free.

The extremely high density of cone receptors in the fovea, and the one-to-one relationship with bipolar cells and retinal ganglion cells (see earlier), endows this component of the cone system with the capacity to mediate high visual acuity. As cone density declines with eccentricity and the degree of convergence onto retinal ganglion cells increases, acuity is markedly reduced. Just 6° eccentric to the line of sight, acuity is reduced by 75%, a fact

Figure 10.10 Distribution of rods and cones in the human retina. Graph illustrates that cones are present at a low density throughout the retina, with a sharp peak in the center of the fovea. Conversely, rods are present at high density throughout most of the retina, with a sharp decline in the fovea. Boxes at top illustrate the appearance of face on sections through the outer segments of the photoreceptors at different eccentricities. The increased density of cones in the fovea is accompanied by a striking reduction in the diameter of their outer segments.



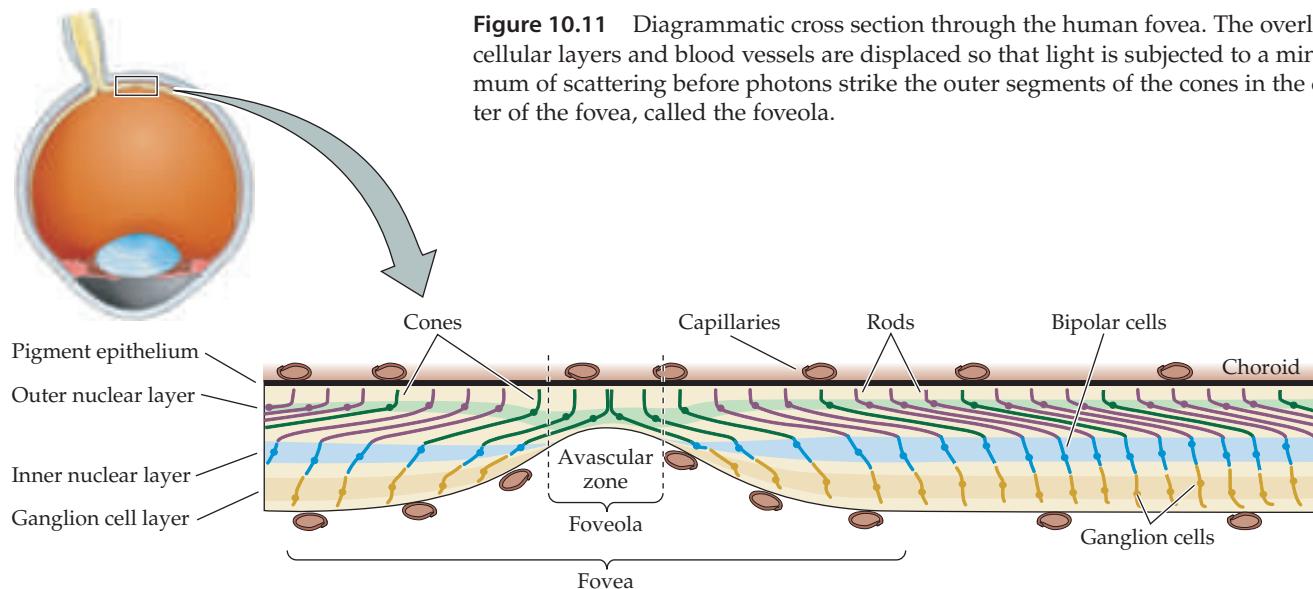


Figure 10.11 Diagrammatic cross section through the human fovea. The overlying cellular layers and blood vessels are displaced so that light is subjected to a minimum of scattering before photons strike the outer segments of the cones in the center of the fovea, called the foveola.

that can be readily appreciated by trying to read the words on any line of this page beyond the word being fixated on. The restriction of highest acuity vision to such a small region of the retina is the main reason humans spend so much time moving their eyes (and heads) around—in effect directing the foveas of the two eyes to objects of interest (see Chapter 19). It is also the reason why disorders that affect the functioning of the fovea have such devastating effects on sight (see Box C). Conversely, the exclusion of rods from the fovea, and their presence in high density away from the fovea, explain why the threshold for detecting a light stimulus is lower outside the region of central vision. It is easier to see a dim object (such as a faint star) by looking slightly away from it, so that the stimulus falls on the region of the retina that is richest in rods (see Figure 10.10).

Another anatomical feature of the fovea (which literally means “pit”) that contributes to the superior acuity of the cone system is that the layers of cell bodies and processes that overlie the photoreceptors in other areas of the retina are displaced around the fovea, and especially the foveola (see Figure 10.11). As a result, photons are subjected to a minimum of scattering before they strike the photoreceptors. Finally, another potential source of optical distortion that lies in the light path to the receptors—the retinal blood vessels—are diverted away from the foveola. This central region of the fovea is therefore dependent on the underlying choroid and pigment epithelium for oxygenation and metabolic sustenance.

Cones and Color Vision

A special property of the cone system is color vision. Perceiving color allows humans (and many other animals) to discriminate objects on the basis of the distribution of the wavelengths of light that they reflect to the eye. While differences in luminance (i.e., overall light intensity) are often sufficient to distinguish objects, color adds another perceptual dimension that is especially useful when differences in luminance are subtle or nonexistent. Color obviously gives us a quite different way of perceiving and describing the world we live in.

Unlike rods, which contain a single photopigment, there are three types of cones that differ in the photopigment they contain. Each of these photopigments has a different sensitivity to light of different wavelengths, and for this reason are referred to as “blue,” “green,” and “red” or, more appropriately, short (S), medium (M), and long (L) wavelength cones—terms that more or less describe their spectral sensitivities (Figure 10.12). This nomenclature implies that individual cones provide color information for the wavelength of light that excites them best. In fact, individual cones, like rods, are entirely color blind in that their response is simply a reflection of the number of photons they capture, regardless of the wavelength of the photon (or, more properly, its vibrational energy). It is impossible, therefore, to determine whether the change in the membrane potential of a particular cone has arisen from exposure to many photons at wavelengths to which the receptor is relatively insensitive, or fewer photons at wavelengths to which it is most sensitive. This ambiguity can only be resolved by *comparing* the activity in different classes of cones. Based on the responses of individual ganglion cells, and cells at higher levels in the visual pathway (see Chapter 11), comparisons of this type are clearly involved in how the visual system extracts color information from spectral stimuli. Despite these insights, a full understanding of the neural mechanisms that underlie color perception has been elusive (Box D).

Much additional information about color vision has come from studies of individuals with abnormal color detecting abilities. Color vision deficiencies result either from the inherited failure to make one or more of the cone pigments or from an alteration in the absorption spectra of cone pigments (or, rarely, from lesions in the central stations that process color information; see Chapter 11). Under normal conditions, most people can match any color in a test stimulus by adjusting the intensity of three superimposed light sources generating long, medium, and short wavelengths. The fact that only three such sources are needed to match (nearly) all the perceived colors is strong

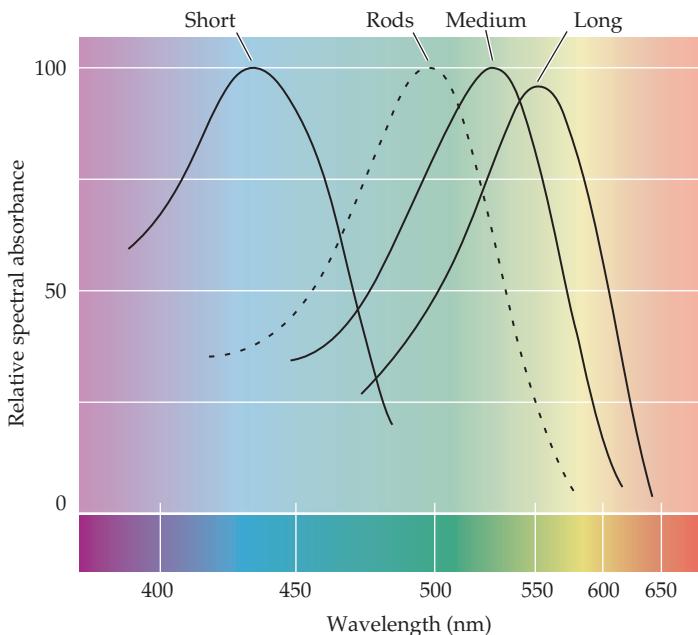


Figure 10.12 Color vision. The light absorption spectra of the four photopigments in the normal human retina. (Recall that light is defined as electromagnetic radiation having wavelengths between ~400 and 700 nm.) The solid curves indicate the three kinds of cone opsins; the dashed curve shows rod rhodopsin for comparison. Absorbance is defined as the log value of the intensity of incident light divided by intensity of transmitted light.

Box D

The Importance of Context in Color Perception

Seeing color logically demands that retinal responses to different wavelengths in some way be *compared*. The discovery of the three human cone types and their different absorption spectra is correctly regarded, therefore, as the basis for human color vision. Nevertheless, how these human cone types and the higher-order neurons they contact (see Chapter 11) produce the sensations of color is still unclear. Indeed, this issue has been debated by some of the greatest minds in science (Hering, Helmholtz, Maxwell, Schroedinger, and Mach, to name only a few) since Thomas Young first proposed that humans must have three different receptive “particles”—i.e., the three cone types.

A fundamental problem has been that, although the relative activities of three cone types can more or less explain the colors perceived in color-matching experiments performed in the laboratory, the perception of color is strongly influenced by context. For example, a patch returning the exact same spectrum of wavelengths to the eye can appear quite different depending on its surround, a phenomenon called *color contrast* (Figure A). Moreover, test patches returning different spectra to the eye can appear to be the same color, an effect called *color constancy* (Figure B). Although these phenomena were well known in the nineteenth century, they were not accorded a central place in color vision theory until Edwin Land’s work in the 1950s. In his most famous demonstration, Land (who among other achievements founded the Polaroid company and became a billionaire) used a collage of colored papers that have been referred to as “the Land Mondrians” because of their similarity to the work of the Dutch artist Piet Mondrian.

Using a telemetric photometer and three adjustable illuminators generating short, middle, and long wavelength light, Land showed that two patches that in

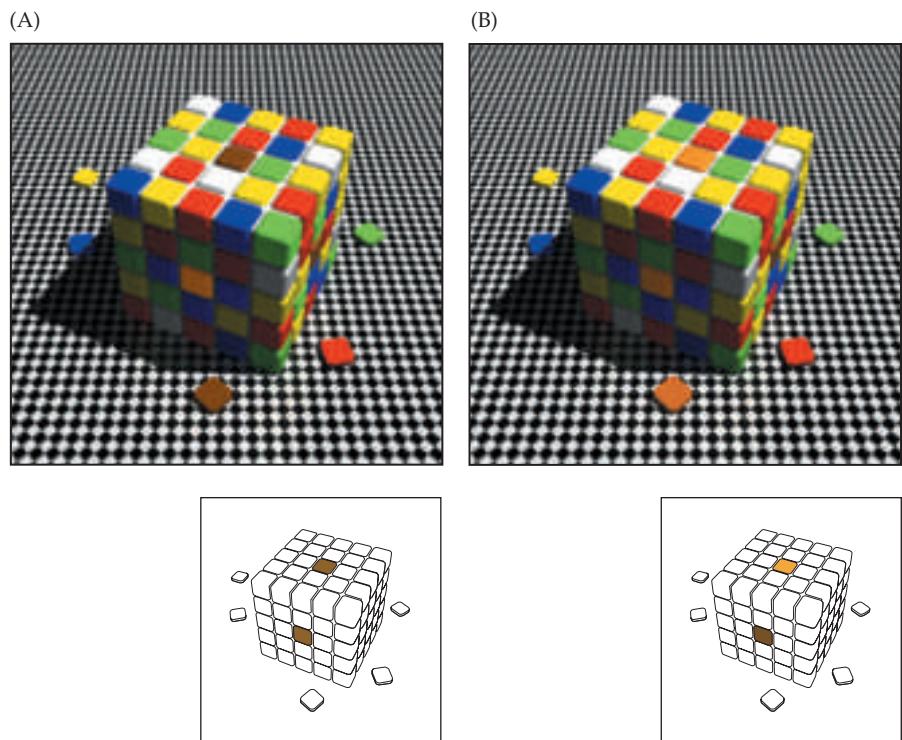
white light appeared quite different in color (e.g., green and brown) continued to look their respective colors even when the three illuminators were adjusted so that the light being returned from the “green” surfaces produced exactly the same readings on the three telephotometers as had previously come from the “brown” surface—a striking demonstration of color constancy.

The phenomena of color contrast and color constancy have led to a heated modern debate about how color percepts are generated that now spans several decades. For Land, the answer lay in a series of ratiometric equations that could integrate the spectral returns of different regions over the entire scene. It was rec-

ognized even before Land’s death in 1991, however, that his so-called Retinex theory did not work in all circumstances and was in any event a description rather than an explanation. An alternative explanation of these contextual aspects of color vision is that color, like brightness, is generated empirically according to what spectral stimuli have typically signified in past experience (see Box E).

References

- LAND, E. (1986) Recent advances in Retinex theory. *Vis. Res.* 26: 7–21.
PURVES, D. AND R. B. LOTTO (2003) *Why We See What We Do: An Empirical Theory of Vision*, Chapters 5 and 6. Sunderland MA: Sinauer Associates, pp. 89–138.

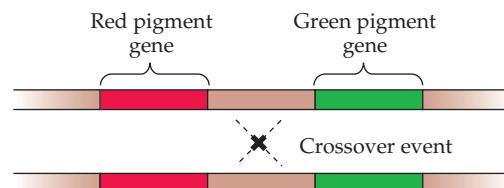


The genesis of contrast and constancy effects by exactly the same context. The two panels demonstrate the effects on apparent color when two *similarly* reflective target surfaces (A) or two *differently* reflective target surfaces (B) are presented in the *same* context in which all the information provided is consistent with illumination that differs only in intensity. The appearances of the relevant target surfaces in a neutral context are shown in the insets below. (From Purves and Lotto, 2003)

confirmation of the fact that color sensation is based on the relative levels of activity in three sets of cones with different absorption spectra. That color vision is **trichromatic** was first recognized by Thomas Young at the beginning of the nineteenth century (thus, people with normal color vision are called *trichromats*). For about 5–6% of the male population in the United States and a much smaller percentage of the female population, however, color vision is more limited. Only two bandwidths of light are needed to match all the colors that these individuals can perceive; the third color category is simply not seen. Such **dichromacy**, or “color blindness” as it is commonly called, is inherited as a recessive, sex-linked characteristic and exists in two forms: *protanopia*, in which all color matches can be achieved by using only green and blue light, and *deuteranopia*, in which all matches can be achieved by using only blue and red light. In another major class of color deficiencies, all three light sources (i.e., short, medium, and long wavelengths) are needed to make all possible color matches, but the matches are made using values that are significantly different from those used by most individuals. Some of these *anomalous trichromats* require more red than normal to match other colors (protanomalous trichromats); others require more green than normal (deutanomalous trichromats).

Jeremy Nathans and his colleagues at Johns Hopkins University have provided a deeper understanding of these color vision deficiencies by identifying and sequencing the genes that encode the three human cone pigments (Figure 10.13). The genes that encode the red and green pigments show a high degree of sequence homology and lie adjacent to each other on the X chromosome, thus explaining the prevalence of color blindness in males. In contrast, the blue-sensitive pigment gene is found on chromosome 7 and is quite different in its amino acid sequence. These facts suggest that the red and green pigment genes evolved relatively recently, perhaps as a result of the duplication of a single ancestral gene; they also explain why most color vision abnormalities involve the red and green cone pigments.

Human dichromats lack one of the three cone pigments, either because the corresponding gene is missing or because it exists as a hybrid of the red and green pigment genes (see Figure 10.13). For example, some dichromats lack the green pigment gene altogether, while others have a hybrid gene that is thought to produce a red-like pigment in the “green” cones. Anomalous trichromats also possess hybrid genes, but these genes elaborate pigments



Different crossover events can lead to:

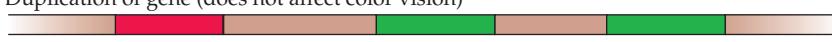
- (1) Hybrid gene 
 - (2) Loss of gene 
 - (3) Duplication of gene (does not affect color vision) 
- Patterns in color-blind men

Figure 10.13 Many deficiencies of color vision are the result of genetic alterations in the red or green cone pigments due to the crossing over of chromosomes during meiosis. This recombination can lead to the loss of a gene, the duplication of a gene, or the formation of a hybrid with characteristics distinct from those of normal genes.

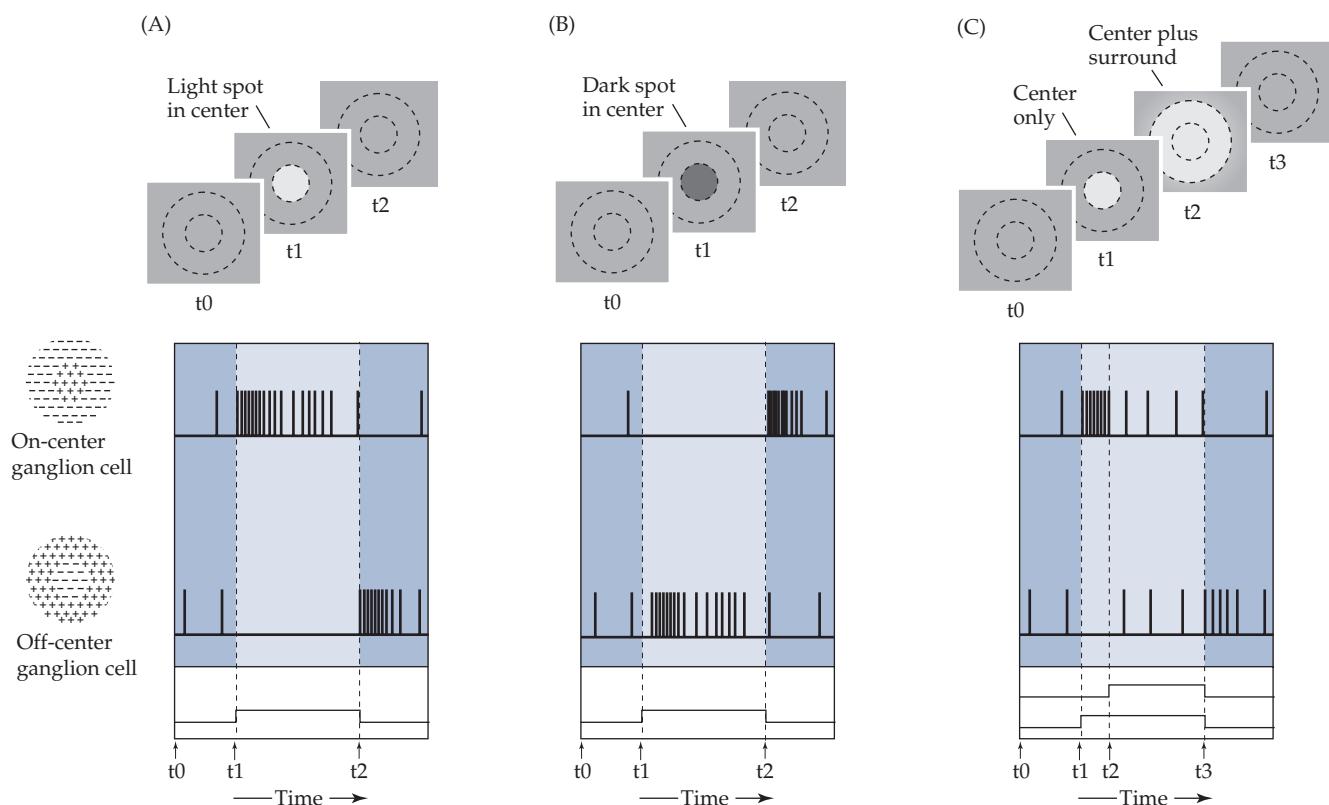
whose spectral properties lie between those of the normal red and green pigments. Thus, although most anomalous trichromats have distinct sets of medium and long-wavelength cones, there is more overlap in their absorption spectra than in normal trichromats, and thus less difference in how the two sets of cones respond to a given wavelength (with resulting anomalies in color perception).

Retinal Circuits for Detecting Luminance Change

Despite the esthetically pleasing nature of color vision, most of the information in visual scenes consists of spatial variations in light intensity (a black and white movie, for example, has most of the information a color version has, although it is deficient in some respects and usually is less fun to watch). How the spatial patterns of light and dark that fall on the photoreceptors are deciphered by central targets has been a vexing problem (Box E). To understand what is accomplished by the complex neural circuits within the retina during this process, it is useful to start by considering the responses of individual retinal ganglion cells to small spots of light. Stephen Kuffler, working at Johns Hopkins University in the 1950s, pioneered this approach by characterizing the responses of single ganglion cells in the cat retina. He found that each ganglion cell responds to stimulation of a small circular patch of the retina, which defines the cell's receptive field (see Chapter 8 for discussion of receptive fields). Based on these responses, Kuffler distinguished two classes of ganglion cells, "on"-center and "off"-center (Figure 10.14).

Turning on a spot of light in the receptive field center of an **on-center ganglion cell** produces a burst of action potentials. The same stimulus applied to the receptive field center of an **off-center ganglion cell** reduces the rate of

Figure 10.14 The responses of on-center and off-center retinal ganglion cells to stimulation of different regions of their receptive fields. Upper panels indicate the time sequence of stimulus changes. (A) Effects of light spot in the receptive field center. (B) Effects of dark spot in the receptive field center. (C) Effects of light spot in the center followed by the addition of light in the surround.



Box E

The Perception of Light Intensity

Understanding the link between retinal stimulation and what we see (perception) is arguably the central problem in vision, and the relation of luminance (a physical measurement of light intensity) and brightness (the sensation elicited by light intensity) is probably the simplest place to consider this challenge.

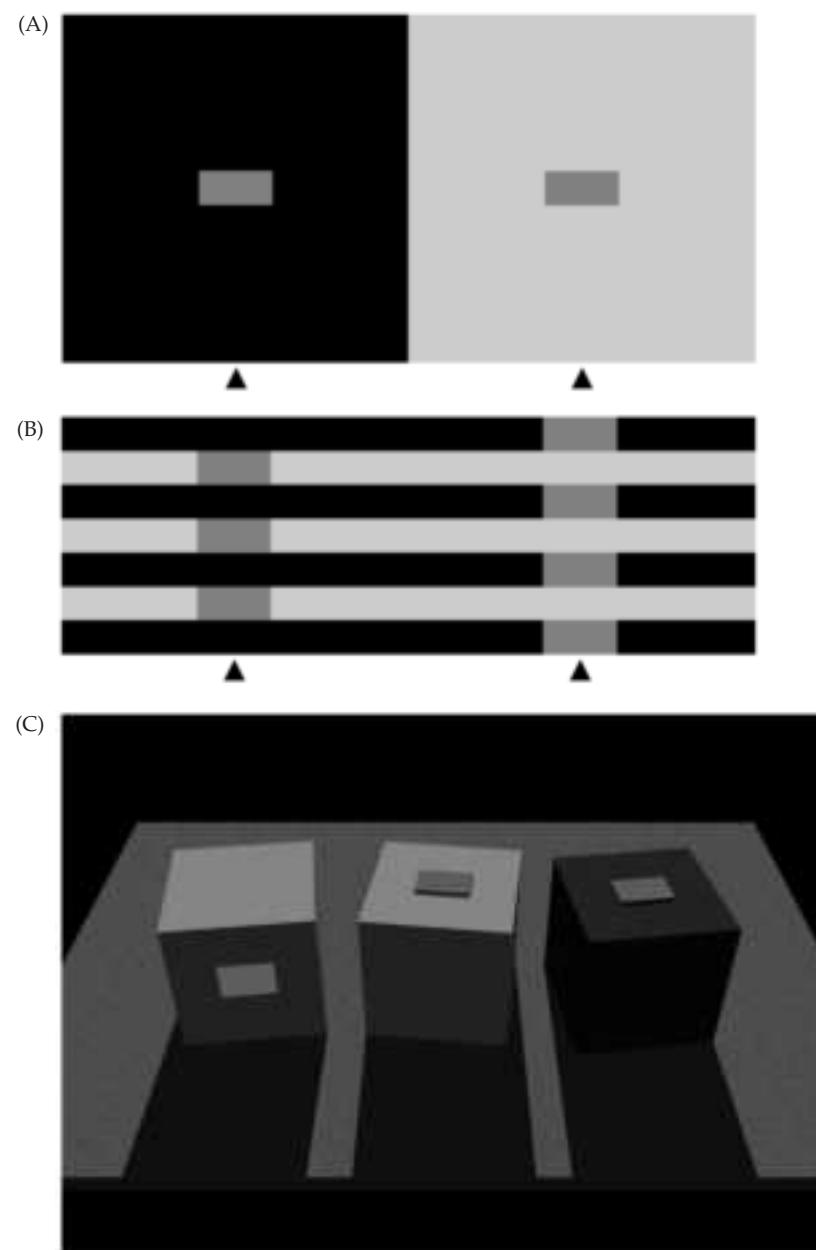
As indicated in the text, how we see the brightness differences (i.e., contrast) between adjacent territories with distinct luminances depends in the first instance on the relative firing rate of retinal ganglion cells, modified by lateral interactions. However, there is a problem with the assumption that the central nervous system simply “reads out” these relative rates of ganglion cell activity to sense brightness. The difficulty, as in perceiving color, is that the brightness of a given target is markedly affected by its context in ways that are difficult or impossible to explain in terms of the retinal output as such. The accompanying figures, which illustrate two simultaneous brightness contrast illusions, help make this point. In Figure A, two photometrically identical (equiluminant) gray squares appear differently bright as a function of the background in which they are presented.

A conventional interpretation of this phenomenon is that the receptive field properties illustrated in Figures 10.14 through 10.17 cause ganglion cells to fire differently depending on whether the surround of the equiluminant target is dark or light. The demonstration in Figure B, however, undermines this explanation, since in this case the target surrounded by more dark area actually looks *darker* than the same target surrounded by more light area.

An alternative interpretation of luminance perception that can account for these puzzling phenomena is that brightness percepts are generated on a statistical basis as a means of contending with the inherent ambiguity of luminance (i.e., the fact that a given value of lumi-

nance can be generated by many different combinations of illumination and surface reflectance properties). Since to be successful an observer has to respond

to the real-world sources of luminance and not to light intensity as such, this ambiguity of the retinal stimulus presents a quandary. A plausible solution to



(A) Standard illusion of simultaneous brightness contrast. (B) Another illusion of simultaneous brightness contrast that is difficult to explain in conventional terms. (C) Cartoons of some possible sources of the standard simultaneous brightness contrast illusion in (A). (Courtesy of R. Beau Lotto and Dale Purves.)

the inherent uncertainty of the relationship between luminance values and their actual sources would be to generate the sensation of brightness elicited by a given luminance (e.g., in the brightness of the identical test patches in the figure) on the basis of what the luminance of the test patches had typically turned out to be in the past experience of human observers. To get the gist of this explanation consider Figure C, which illustrates the point that the two equiluminant target patches in Figure A could have been generated by two differently painted surfaces in different illuminants, as in a comparison of the target patches on the left and middle cubes, or two similarly

reflecting surfaces in similar amounts of light, as in a comparison of the target patches on the middle and right cubes. An expedient—and perhaps the only—way the visual system can cope with this ambiguity is to generate the perception of the stimulus in Figure A (and in Figure B) empirically, i.e., based on what the target patches typically turned out to signify in the past. Since the equiluminant targets will have arisen from a variety of possible sources, it makes sense to have the brightness elicited by the patches determined statistically by the relative frequency of occurrence of that luminance in the particular context in which it is presented. The advantage of seeing

luminance according to the relative probabilities of the possible sources of the stimulus is that percepts generated in this way give the observer the best chance of making appropriate behavioral responses to profoundly ambiguous stimuli.

References

- ADELSON, E. H. (1999) Light perception and lightness illusions. In *The Cognitive Neurosciences*, 2nd Ed. M. Gazzaniga (ed.). Cambridge, MA: MIT Press, pp. 339–351.
- PURVES, D. AND R. B. LOTTO (2003) *Why We See What We Do: An Empirical Theory of Vision*, Chapters 3 and 4. Sunderland MA: Sinauer Associates, pp. 41–87.

discharge, and when the spot of light is turned off, the cell responds with a burst of action potentials (Figure 10.14A). Complementary patterns of activity are found for each cell type when a dark spot is placed in the receptive field center (Figure 10.14B). Thus, on-center cells increase their discharge rate to luminance *increments* in the receptive field center, whereas off-center cells increase their discharge rate to luminance *decrements* in the receptive field center.

On- and off-center ganglion cells are present in roughly equal numbers. The receptive fields have overlapping distributions, so that every point on the retinal surface (that is, every part of visual space) is analyzed by several on-center and several off-center ganglion cells. A rationale for having these two distinct types of retinal ganglion cells was suggested by Peter Schiller and his colleagues at the Massachusetts Institute of Technology, who examined the effects of pharmacologically inactivating on-center ganglion cells on a monkey's ability to detect a variety of visual stimuli. After silencing on-center ganglion cells, the animals showed a deficit in their ability to detect stimuli that were brighter than the background; however, they could still see objects that were darker than the background.

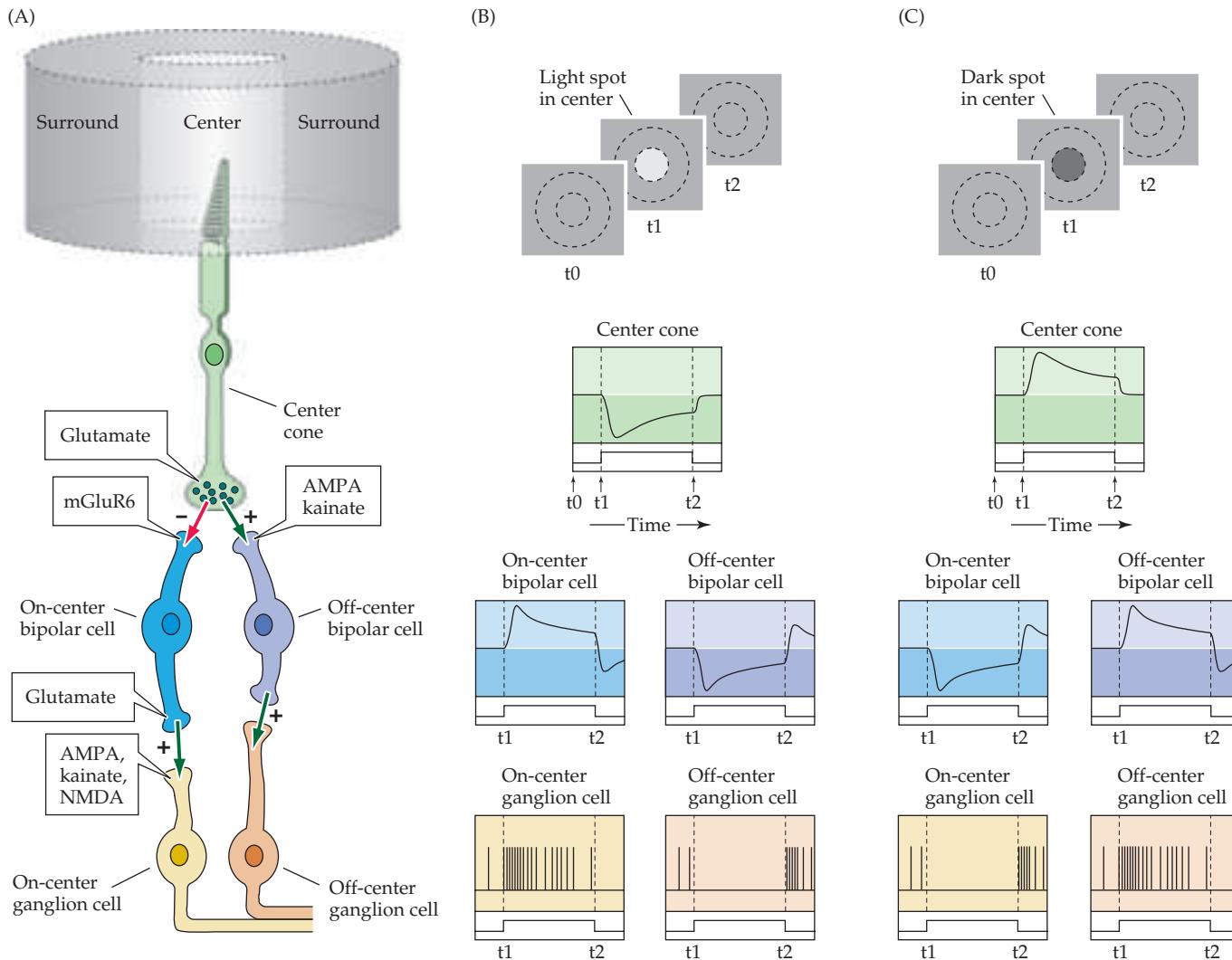
These observations imply that information about increases or decreases in luminance is carried separately to the brain by the axons of these two different types of retinal ganglion cells. Having separate luminance "channels" means that changes in light intensity, whether increases or decreases, are always conveyed to the brain by an increased number of action potentials. Because ganglion cells rapidly adapt to changes in luminance, their "resting" discharge rate in constant illumination is relatively low. Although an increase in discharge rate above resting level serves as a reliable signal, a decrease in firing rate from an initially low rate of discharge might not. Thus, having luminance changes signaled by two classes of adaptable cells provides unambiguous information about both luminance increments and decrements.

The functional differences between these two ganglion cell types can be understood in terms of both their anatomy and their physiological proper-

ties and relationships (Figure 10.15). On- and off-center ganglion cells have dendrites that arborize in separate strata of the inner plexiform layer, forming synapses selectively with the terminals of on- and off-center bipolar cells that respond to luminance increases and decreases, respectively (Figure 10.15A). As mentioned previously, the principal difference between ganglion cells and bipolar cells lies in the nature of their electrical response. Like most other cells in the retina, bipolar cells have graded potentials rather than action potentials. Graded depolarization of bipolar cells leads to an increase in transmitter release (glutamate) at their synapses and consequent depolarization of the on-center ganglion cells that they contact via AMPA, kainate, and NMDA receptors.

The selective response of on- and off-center bipolar cells to light increments and decrements is explained by the fact that they express different types of glutamate receptors (Figure 10.15A). Off-center bipolar cells have ionotropic receptors (AMPA and kainate) that cause the cells to depolarize in response to glutamate released from photoreceptor terminals. In contrast, on-center bipolar cells express a G-protein-coupled metabotropic glutamate receptor (mGluR6). When bound to glutamate, these receptors activate an intracellular cascade that closes cGMP-gated Na^+ channels, reducing inward

Figure 10.15 Circuitry responsible for generating receptive field center responses of retinal ganglion cells. (A) Functional anatomy of cone inputs to the center of a ganglion cell receptive field. A plus indicates a sign-conserving synapse; a minus represents a sign-inverting synapse. (B) Responses of various cell types to the presentation of a light spot in the center of the ganglion cell receptive field. (C) Responses of various cell types to the presentation of a dark spot in the center of the ganglion cell receptive field.



current and hyperpolarizing the cell. Thus, glutamate has opposite effects on these two classes of cells, depolarizing off-center bipolar cells and hyperpolarizing on-center cells. Photoreceptor synapses with off-center bipolar cells are called sign-conserving, since the sign of the change in membrane potential of the bipolar cell (depolarization or hyperpolarization) is the same as that in the photoreceptor (Figure 10.15B,C). Photoreceptor synapses with on-center bipolar cells are called sign-inverting because the change in the membrane potential of the bipolar cell is the opposite of that in the photoreceptor.

In order to understand the response of on- and off-center bipolar cells to changes in light intensity, recall that photoreceptors hyperpolarize in response to light increments, decreasing their release of neurotransmitter (Figure 10.15B). Under these conditions, on-center bipolar cells contacted by the photoreceptors are freed from the hyperpolarizing influence of the photoreceptor's transmitter, and they depolarize. In contrast, for off-center cells, the reduction in glutamate represents the withdrawal of a depolarizing influence, and these cells hyperpolarize. Decrements in light intensity naturally have the opposite effect on these two classes of bipolar cells, hyperpolarizing on-center cells and depolarizing off-center ones (Figure 10.15C).

Kuffler's work also called attention to the fact that retinal ganglion cells do not act as simple photodetectors. Indeed, most ganglion cells are relatively poor at signaling differences in the level of diffuse illumination. Instead, they are sensitive to *differences* between the level of illumination that falls on the receptive field center and the level of illumination that falls on the surround—that is, to **luminance contrast**. The center of a ganglion cell receptive field is surrounded by a concentric region that, when stimulated, antagonizes the response to stimulation of the receptive field center (see Figure 10.14C). For example, as a spot of light is moved from the center of the receptive field of an on-center cell toward its periphery, the response of the cell to the spot of light decreases (Figure 10.16). When the spot falls completely outside the center (that is, in the surround), the response of the cell falls below its resting level; the cell is effectively inhibited until the distance from the center is so great that the spot no longer falls on the receptive field at all, in which case the cell returns to its resting level of firing. Off-center

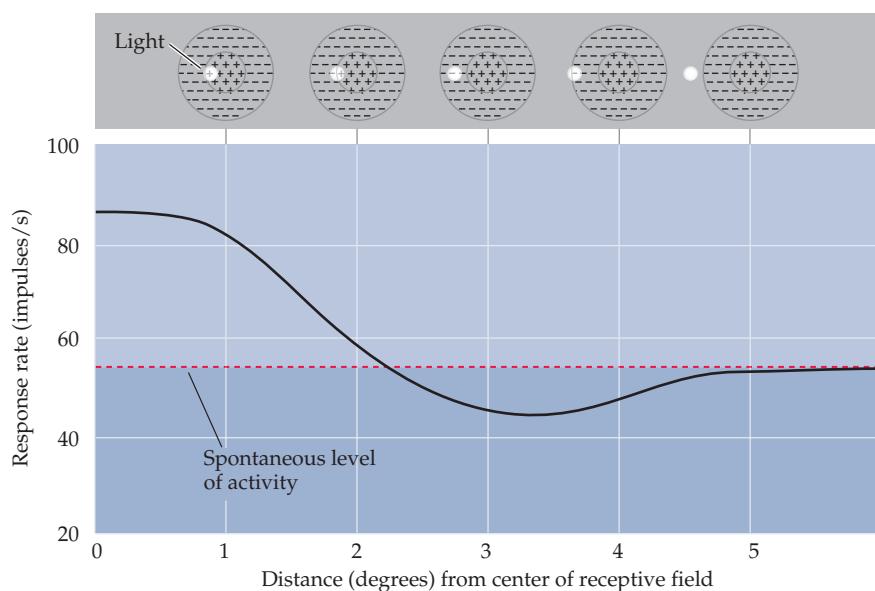
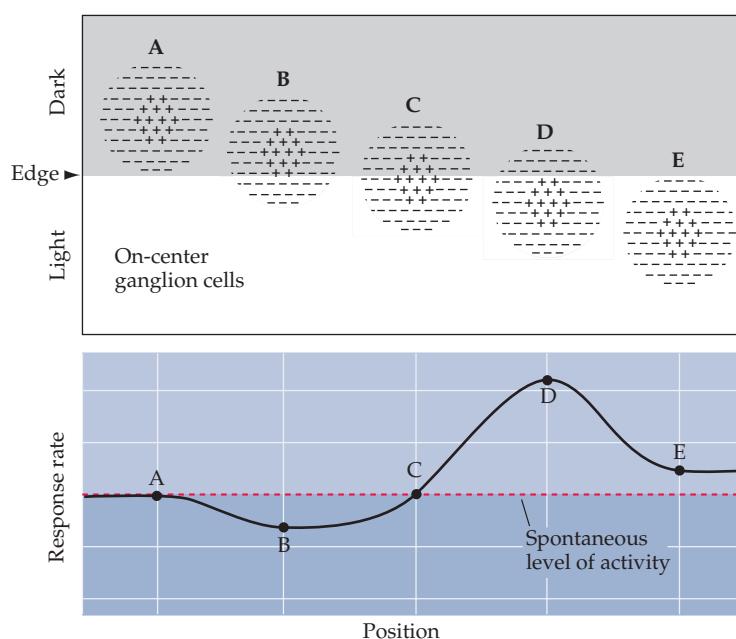


Figure 10.16 Rate of discharge of an on-center ganglion cell to a spot of light as a function of the distance of the spot from the receptive field center. Zero on the x axis corresponds to the center; at a distance of 5°, the spot falls outside the receptive field.

Figure 10.17 Responses of a hypothetical population of on-center ganglion cells whose receptive fields (A–E) are distributed across a light-dark edge. Those cells whose activity is most affected have receptive fields that lie along the light-dark edge.



cells exhibit a similar surround antagonism. Stimulation of the surround by light opposes the decrease in firing rate that occurs when the center is stimulated alone, and reduces the response to light decrements in the center (compare Figures 10.14A and 10.14C).

Because of their antagonistic surrounds, ganglion cells respond much more vigorously to small spots of light confined to their receptive field centers than to large spots, or to uniform illumination of the visual field (see Figure 10.14C).

To appreciate how center-surround antagonism makes the ganglion cell sensitive to luminance contrast, consider the activity levels in a hypothetical population of on-center ganglion cells whose receptive fields are distributed across a retinal image of a light-dark edge (Figure 10.17). The neurons whose firing rates are most affected by this stimulus—either increased (neuron D) or decreased (neuron B)—are those with receptive fields that lie along the light-dark border; those with receptive fields completely illuminated (or completely darkened) are less affected (neurons A and E). Thus, the information supplied by the retina to central visual stations for further processing does not give equal weight to all regions of the visual scene; rather, it emphasizes the regions where there are differences in luminance.

Contribution of Retinal Circuits to Light Adaptation

In addition to making ganglion cells especially sensitive to light-dark borders in the visual scene, center-surround mechanisms make a significant contribution to the process of **light adaptation**. As illustrated for an on-center cell in Figure 10.18, the response rate of a ganglion cell to a small spot of light turned on in its receptive field center varies as a function of the spot's intensity. In fact, response rate is proportional to the spot's intensity over a range of about one log unit. However, the intensity of spot illumination required to evoke a given discharge rate is dependent on the background level of illumination. Increases in background level of illumination are accompanied by adaptive shifts in the cell's operating range such that

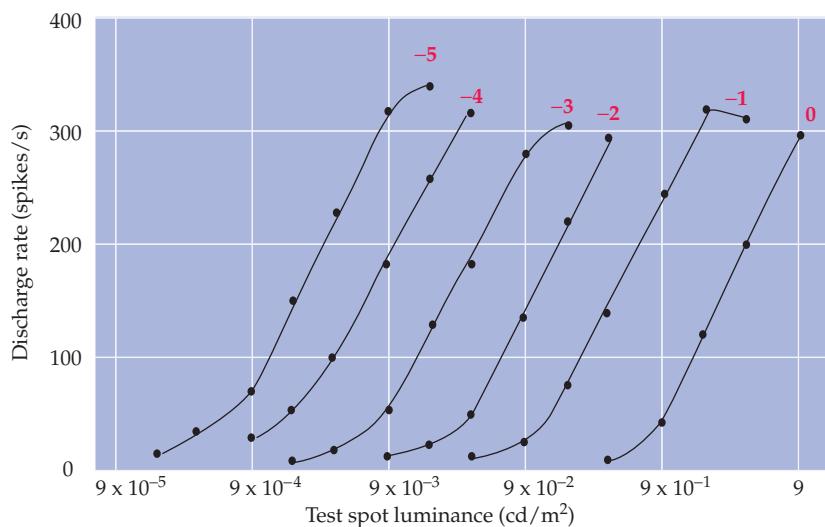


Figure 10.18 A series of curves illustrating the discharge rate of a single on-center ganglion cell to the onset of a small test spot of light in the center of its receptive field. Each curve represents the discharge rate evoked by spots of varying intensity at a constant background level of illumination, which is given by the red numbers at the top of each curve (the highest background level is 0, the lowest -5). The response rate is proportional to stimulus intensity over a range of 1 log unit, but the operating range shifts to the right as the background level of illumination increases.

greater stimulus intensities are required to achieve the same discharge rate. Thus, firing rate is not an absolute measure of light intensity, but rather signals the difference from background level of illumination.

Because the range of light intensities over which we can see is enormous compared to the narrow range of ganglion cell discharge rates (see Figure 10.9), adaptational mechanisms are essential. By scaling the ganglion cell's response to ambient levels of illumination, the entire dynamic range of a neuron's firing rate is used to encode information about intensity differences over the range of luminance values that are relevant for a given visual scene. Due to the antagonistic center-surround organization of retinal ganglion cells, the signal sent to the brain from the retina downplays the background level of illumination (see Figure 10.14). This arrangement presumably explains why the relative brightness of objects remains much the same over a wide range of lighting conditions. In bright sunlight, for example, the print on this page reflects considerably more light to the eye than it does in room light. In fact, the *print* reflects more light in sunlight than the *paper* reflects in room light; yet it continues to look black and the page white, indoors or out.

Like the mechanism responsible for generating the on- and off-center response, the antagonistic surround of ganglion cells is a product of interactions that occur at the early stages of retinal processing (Figure 10.19). Much of the antagonism is thought to arise via lateral connections established by horizontal cells and receptor terminals. Horizontal cells receive synaptic inputs from photoreceptor terminals and are linked via gap junctions with a vast network of other horizontal cells distributed over a wide area of the retinal surface. As a result, the activity in horizontal cells reflects levels of illumination over a broad area of the retina. Although the details of their actions are not entirely clear, horizontal cells are thought to exert their influence via the release of neurotransmitter directly onto photoreceptor terminals, regulating the amount of transmitter that the photoreceptors release onto bipolar cell dendrites.

Glutamate release from photoreceptor terminals has a depolarizing effect on horizontal cells (sign-conserving synapse), while the transmitter released from horizontal cells (GABA) has a hyperpolarizing influence on photoreceptor terminals (sign-inverting synapse) (Figure 10.19A). As a result, the net effect of inputs from the horizontal cell network is to oppose changes in the

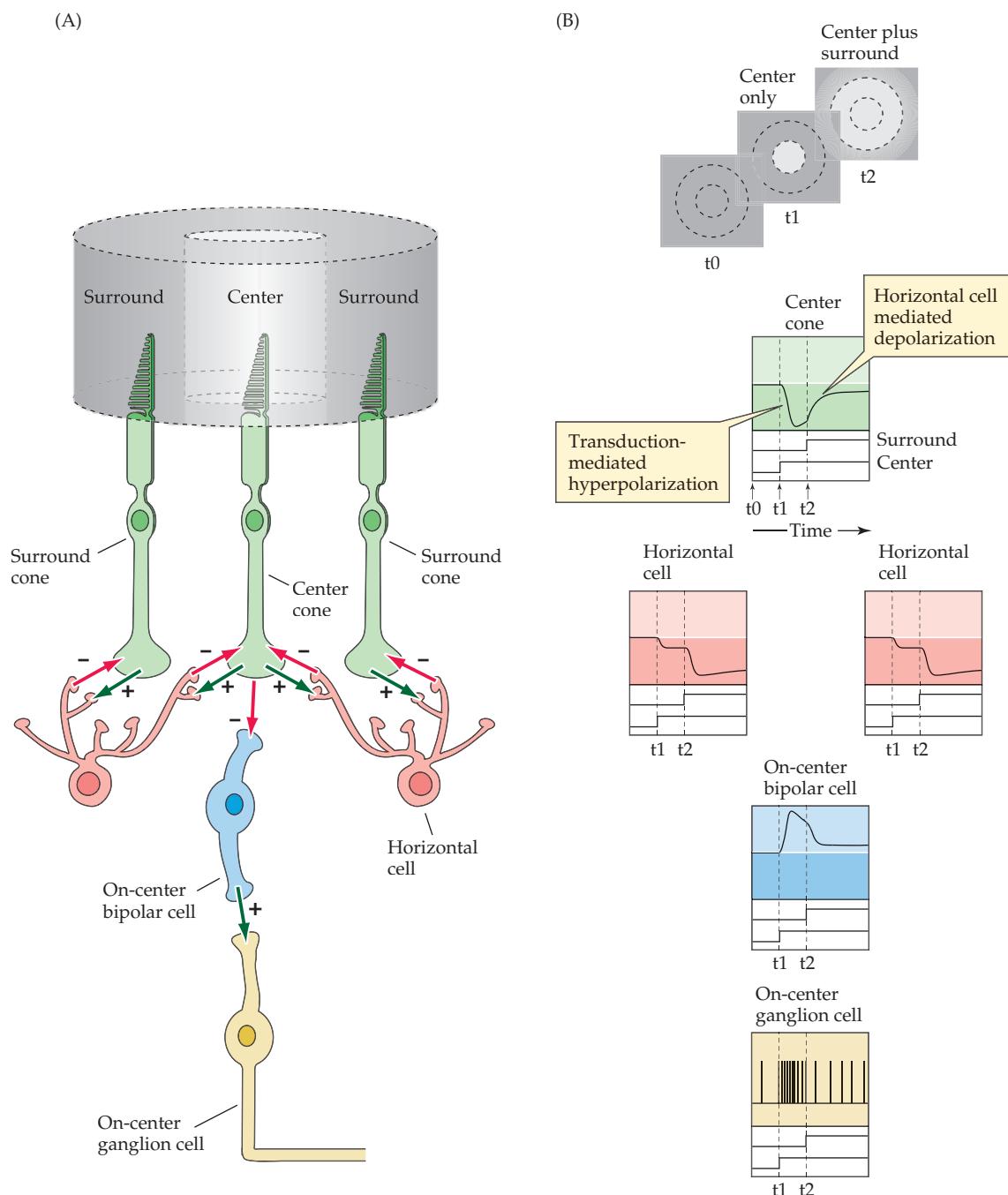


Figure 10.19 Circuitry responsible for generating the receptive field surround of an on-center retinal ganglion cell. (A) Functional anatomy of horizontal cell inputs responsible for surround antagonism. A plus indicates a sign-conserving synapse; a minus represents a sign-inverting synapse. (B) Responses of various cell types to the presentation of a light spot in the center of the receptive field (t1) followed by the addition of light stimulation in the surround (t2). Light stimulation of the surround leads to hyperpolarization of the horizontal cells and a decrease in the release of inhibitory transmitter (GABA) onto the photoreceptor terminals. The net effect is to depolarize the center cone terminal, offsetting much of the hyperpolarization induced by the transduction cascade in the center cone's outer segment.

membrane potential of the photoreceptor that are induced by phototransduction events in the outer segment. How these events lead to surround suppression in an on-center ganglion cell is illustrated in Figure 10.19. A small spot of light centered on a photoreceptor supplying input to the center of the ganglion cell's receptive field produces a strong hyperpolarizing response in the photoreceptor. Under these conditions, changes in the membrane potential of the horizontal cells that synapse with the photoreceptor terminal are relatively small, and the response of the photoreceptor to light is largely determined by its phototransduction cascade (Figure 10.19B). With the addition of light to the surround, however, the impact of the horizontal network becomes significantly greater; the light-induced reduction in the release of glutamate from the photoreceptors in the surround leads to a strong hyperpolarization of the horizontal cells whose processes converge on the terminal of the photoreceptor in the receptive field center. The reduction in GABA release from the horizontal cells has a depolarizing effect on the membrane potential of the central photoreceptor, reducing the light-evoked response and ultimately reducing the firing rate of the on-center ganglion cell.

Thus, even at the earliest stages in visual processing, neural signals do not represent the absolute numbers of photons that are captured by receptors, but rather the relative intensity of stimulation—how much the current level of stimulation differs from ambient levels. While it may seem that the actions of horizontal cells decrease the sensitivity of the retina, they play a critical role in allowing the full range of the photoreceptor's electrical response (about 30 mV) to be applied to the limited range of stimulus intensities that are present at any given moment. The network mechanisms of adaptation described here function in conjunction with cellular mechanisms in the receptor outer segments that regulate the sensitivity of the phototransduction cascade at different light levels. Together, they allow retinal circuits to convey the most salient aspects of luminance changes to the central stages of the visual system described in the following chapter.

Summary

The light that falls on photoreceptors is transformed by retinal circuitry into a pattern of action potentials that ganglion cell axons convey to the visual centers in the brain. This process begins with phototransduction, a biochemical cascade that ultimately regulates the opening and closing of ion channels in the membrane of the photoreceptor's outer segment, and thereby the amount of neurotransmitter the photoreceptor releases. Two systems of photoreceptors—rods and cones—allow the visual system to meet the conflicting demands of sensitivity and acuity, respectively. Retinal ganglion cells operate quite differently from the photoreceptor cells. The center-surround arrangement of ganglion cell receptive fields makes these neurons particularly sensitive to luminance contrast and relatively insensitive to the overall level of illumination. It also allows the retina to adapt, such that it can respond effectively over the enormous range of illuminant intensities in the world. The underlying organization is generated by the synaptic interactions between photoreceptors, horizontal cells, and bipolar cells in the outer plexiform layer. As a result, the signal sent to the visual centers in the brain is already highly processed when it leaves the retina, emphasizing those aspects of the visual scene that convey the most information.

Additional Reading

Reviews

ARSHAVSKY, V. Y., T. D. LAMB AND E. N. PUGH JR. (2002) G proteins and phototransduction. *Annu. Rev. Physiol.* 64: 153–187.

BURNS, M. E. AND D. A. BAYLOR (2001) Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annu. Rev. Neurosci.* 24: 779–805.

NATHANS, J. (1987) Molecular biology of visual pigments. *Annu. Rev. Neurosci.* 10: 163–194.

SCHNAPF, J. L. AND D. A. BAYLOR (1987) How photoreceptor cells respond to light. *Sci. Amer.* 256 (April): 40–47.

STERLING, P. (1990) Retina. In *The Synaptic Organization of the Brain*, G. M. Shepherd (ed.). New York: Oxford University Press, pp. 170–213.

STRYER, L. (1986) Cyclic GMP cascade of vision. *Annu. Rev. Neurosci.* 9: 87–119.

Important Original Papers

BAYLOR, D. A., M. G. F. FUORTES AND P. M. O'BRYAN (1971) Receptive fields of cones in the retina of the turtle. *J. Physiol. (Lond.)* 214: 265–294.

DOWLING, J. E. AND F. S. WERBLIN (1969) Organization of the retina of the mud puppy, *Necturus maculosus*. I. Synaptic structure. *J. Neurophysiol.* 32: 315–338.

ENROTH-CUGELL, C. AND R. M. SHAPLEY (1973) Adaptation and dynamics of cat retinal ganglion cells. *J. Physiol.* 233: 271–309.

FASENKO, E. E., S. S. KOLESNIKOV AND A. L. LYUBARSKY (1985) Induction by cyclic GMP of cationic conductance in plasma membrane of retinal rod outer segment. *Nature* 313: 310–313.

KUFFLER, S. W. (1953) Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.* 16: 37–68.

NATHANS, J., D. THOMAS AND D. S. HOGNESS (1986) Molecular genetics of human color vision: The genes encoding blue, green and red pigments. *Science* 232: 193–202.

NATHANS, J., T. P. PIANTANIDA, R. EDDY, T. B. SHOWS AND D. S. HOGNESS (1986) Molecular genetics of inherited variation in human color vision. *Science* 232: 203–210.

SCHILLER, P. H., J. H. SANDELL AND J. H. R. MAUNSELL (1986) Functions of the “on” and “off” channels of the visual system. *Nature* 322: 824–825.

WERBLIN, F. S. AND J. E. DOWLING (1969) Organization of the retina of the mud puppy, *Necturus maculosus*. II. Intracellular recording. *J. Neurophysiol.* 32: 339–354.

Books

BARLOW, H. B. AND J. D. MOLLON (1982) *The Senses*. London: Cambridge University Press.

DOWLING, J. E. (1987) *The Retina: An Approachable Part of the Brain*. Cambridge, MA: Belknap Press.

FAIN, G. L. (2003) *Sensory Transduction*. Sunderland, MA: Sinauer Associates.

HART, W. M. J. (ed.) (1992) *Adler's Physiology of the Eye: Clinical Application*, 9th Ed. St. Louis, MO: Mosby Year Book.

HELMHOLTZ, H. L. F. VON (1924) *Helmholtz's Treatise on Physiological Optics*, Vol. I–III. Transl. from the Third German Edition by J. P. C. Southall. Menasha, WI: George Banta Publishing Company.

HOGAN, M. J., J. A. ALVARADO AND J. E. WEDDELL (1971) *Histology of the Human Eye: An Atlas and Textbook*. Philadelphia: Saunders.

HUBEL, D. H. (1988) *Eye, Brain, and Vision*, Scientific American Library Series. New York: W. H. Freeman.

HURVICH, L. (1981) *Color Vision*. Sunderland, MA: Sinauer Associates, pp. 180–194.

OGLE, K. N. (1964) *Researches in Binocular Vision*. Hafner: New York.

OYSTER, C. (1999) *The Human Eye: Structure and Function*. Sunderland, MA: Sinauer Associates.

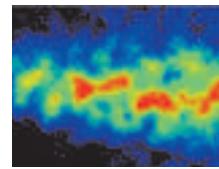
POLYAK, S. (1957) *The Vertebrate Visual System*. Chicago: The University of Chicago Press.

RODIECK, R. W. (1973) *The Vertebrate Retina*. San Francisco: W. H. Freeman.

RODIECK, R. W. (1998) *First Steps in Seeing*. Sunderland, MA: Sinauer Associates.

WANDELL, B. A. (1995) *Foundations of Vision*. Sunderland, MA: Sinauer Associates.

Chapter 11



Central Visual Pathways

Overview

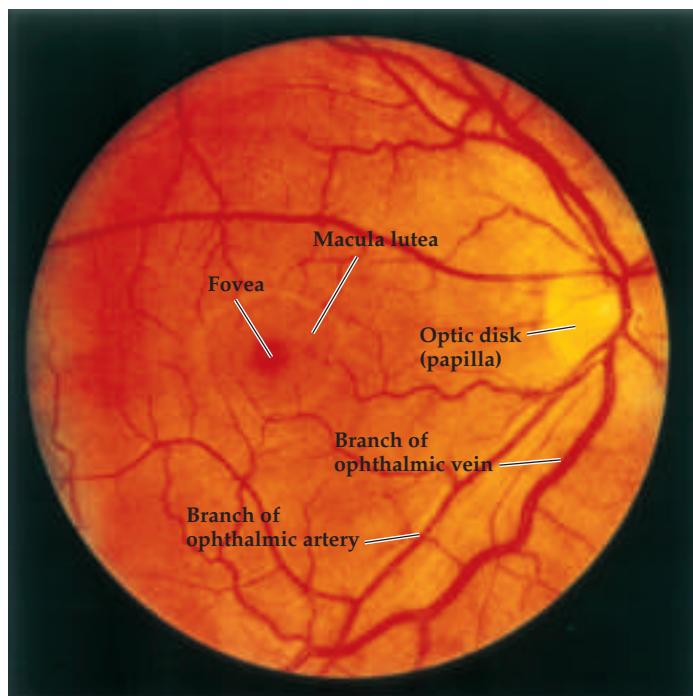
Information supplied by the retina initiates interactions between multiple subdivisions of the brain that eventually lead to conscious perception of the visual scene, at the same time stimulating more conventional reflexes such as adjusting the size of the pupil, directing the eyes to targets of interest, and regulating homeostatic behaviors that are tied to the day/night cycle. The pathways and structures that mediate this broad range of functions are necessarily diverse. Of these, the primary visual pathway from the retina to the dorsal lateral geniculate nucleus in the thalamus and on to the primary visual cortex is the most important and certainly the most thoroughly studied component of the visual system. Different classes of neurons within this pathway encode the varieties of visual information—luminance, spectral differences, orientation, and motion—that we ultimately see. The parallel processing of different categories of visual information continues in cortical pathways that extend beyond primary visual cortex, supplying a variety of visual areas in the occipital, parietal, and temporal lobes. Visual areas in the temporal lobe are primarily involved in object recognition, whereas those in the parietal lobe are concerned with motion. Normal vision depends on the integration of information in all these cortical areas. The processes underlying visual perception are not understood and remain one of the central challenges of modern neuroscience.

Central Projections of Retinal Ganglion Cells

Ganglion cell axons exit the retina through a circular region in its nasal part called the **optic disk** (or optic papilla), where they bundle together to form the **optic nerve**. This region of the retina contains no photoreceptors and, because it is insensitive to light, produces the perceptual phenomenon known as the **blind spot** (Box A). The optic disk is easily identified as a whitish circular area when the retina is examined with an ophthalmoscope; it also is recognized as the site from which the ophthalmic artery and veins enter (or leave) the eye (Figure 11.1). In addition to being a conspicuous retinal landmark, the appearance of the optic disk is a useful gauge of intracranial pressure. The subarachnoid space surrounding the optic nerve is continuous with that of the brain; as a result, increases in intracranial pressure—a sign of serious neurological problems such as a space-occupying lesion—can be detected as *papilledema*, a swelling of the optic disk.

Axons in the optic nerve run a straight course to the **optic chiasm** at the base of the diencephalon. In humans, about 60% of these fibers cross in the chiasm, while the other 40% continue toward the thalamus and midbrain targets on the same side. Once past the chiasm, the ganglion cell axons on each

Figure 11.1 The retinal surface of the left eye, viewed with an ophthalmoscope. The optic disk is the region where the ganglion cell axons leave the retina to form the optic nerve; it is also characterized by the entrance and exit, respectively, of the ophthalmic arteries and veins that supply the retina. The macula lutea can be seen as a distinct area at the center of the optical axis (the optic disk lies nasally); the macula is the region of the retina that has the highest visual acuity. The fovea is a depression or pit about 1.5 mm in diameter that lies at the center of the macula (see Chapter 10).



side form the **optic tract**. Thus, the optic tract, unlike the optic nerve, contains fibers from *both* eyes. The partial crossing (or decussation) of ganglion cell axons at the optic chiasm allows information from corresponding points on the two retinas to be processed by approximately the same cortical site in each hemisphere, an important issue that is considered in the next section.

The ganglion cell axons in the optic tract reach a number of structures in the diencephalon and midbrain (Figure 11.2). The major target in the diencephalon is the **dorsal lateral geniculate nucleus** of the thalamus. Neurons in the lateral geniculate nucleus, like their counterparts in the thalamic relays of other sensory systems, send their axons to the cerebral cortex via the internal capsule. These axons pass through a portion of the internal capsule called the **optic radiation** and terminate in the **primary visual cortex**, or **striate cortex** (also referred to as **Brodmann's area 17** or **V1**), which lies largely along and within the calcarine fissure in the occipital lobe. The **retinogeniculostriate pathway**, or **primary visual pathway**, conveys information that is essential for most of what is thought of as seeing. Thus, damage anywhere along this route results in serious visual impairment.

A second major target of the ganglion cell axons is a collection of neurons that lies between the thalamus and the midbrain in a region known as the **pretectum**. Although small in size compared to the lateral geniculate nucleus, the pretectum is particularly important as the coordinating center for the **pupillary light reflex** (i.e., the reduction in the diameter of the pupil that occurs when sufficient light falls on the retina) (Figure 11.3). The initial component of the pupillary light reflex pathway is a bilateral projection from the retina to the pretectum. Pretectal neurons, in turn, project to the **Edinger-Westphal nucleus**, a small group of nerve cells that lies close to the nucleus of the oculomotor nerve (cranial nerve III) in the midbrain. The Edinger-Westphal nucleus contains the preganglionic parasympathetic neurons that send their axons via the oculomotor nerve to terminate on neurons in the ciliary

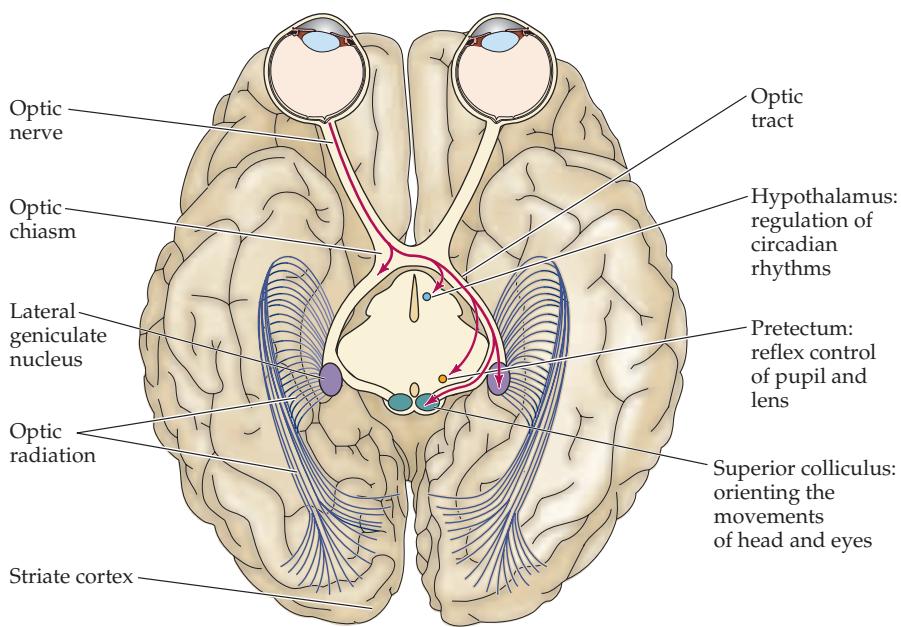


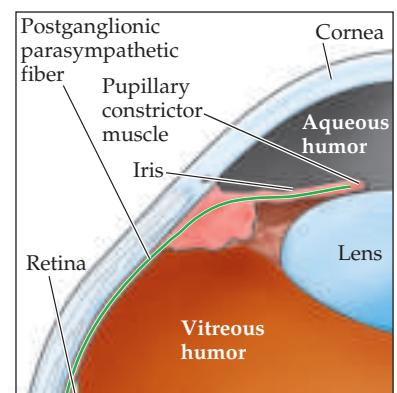
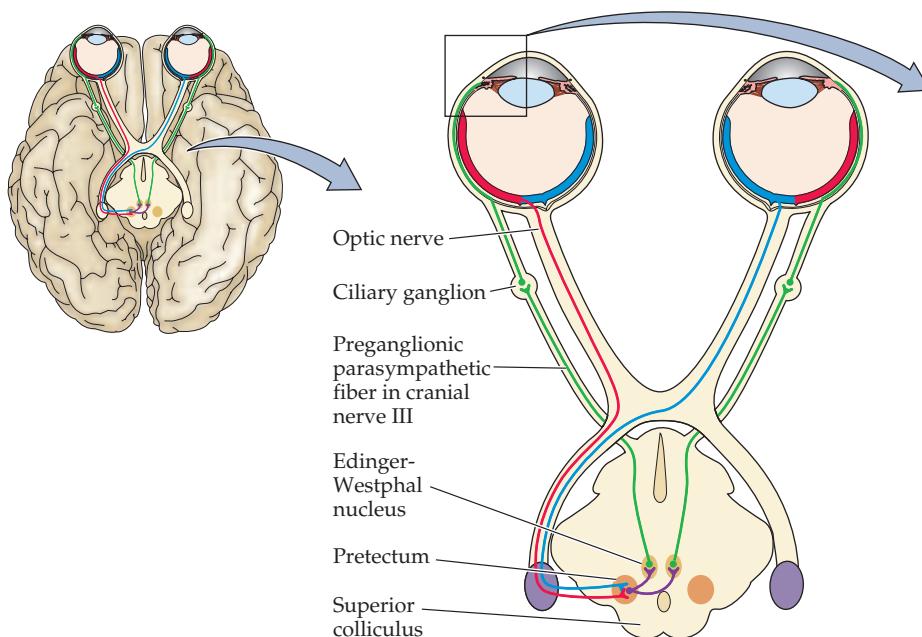
Figure 11.2 Central projections of retinal ganglion cells. Ganglion cell axons terminate in the lateral geniculate nucleus of the thalamus, the superior colliculus, the pretectum, and the hypothalamus. For clarity, only the crossing axons of the right eye are shown (view is looking up at the inferior surface of the brain).



ganglion (see Chapter 19). Neurons in the ciliary ganglion innervate the constrictor muscle in the iris, which decreases the diameter of the pupil when activated. Shining light in the eye thus leads to an increase in the activity of pretectal neurons, which stimulates the Edinger-Westphal neurons and the ciliary ganglion neurons they innervate, thus constricting the pupil.

In addition to its normal role in regulating the amount of light that enters the eye, the pupillary reflex provides an important diagnostic tool that allows the physician to test the integrity of the visual sensory apparatus, the motor outflow to the pupillary muscles, and the central pathways that medi-

Figure 11.3 The circuitry responsible for the pupillary light reflex. This pathway includes bilateral projections from the retina to the pretectum and projections from the pretectum to the Edinger-Westphal nucleus. Neurons in the Edinger-Westphal nucleus terminate in the ciliary ganglion, and neurons in the ciliary ganglion innervate the pupillary constrictor muscles. Notice that the afferent axons activate both Edinger-Westphal nuclei via the neurons in the pretectum.



Box A

The Blind Spot

It is logical to suppose that a visual field defect (called a *scotoma*) arising from damage to the retina or central visual pathways would be obvious to the individual suffering from such pathology. When the deficit involves a peripheral region of the visual field, however, a scotoma often goes unnoticed until a car accident or some other mishap all too dramatically reveals the sensory loss. In fact, all of us have a physiological scotoma of which we are quite unaware, the so-called "blind spot." The blind spot is the substantial gap in each monocular visual field that corresponds to the location of the optic disk, the receptor-free region of the retina where the optic nerve leaves the eye (see Figure 11.1).

To find the "blind spot" of the right eye, close the left eye and fixate on the X shown in the figure here, holding the book about 30–40 centimeters away. Now take a pencil in your right hand and, without breaking fixation, move the tip slowly toward the X from the right side of the page. At some point, the tip of the pencil (indeed the whole end of the pencil) will disappear; mark this point and continue to move the pencil to the left until it reappears; then make another mark. The borders of the blind spot along the vertical axis can be determined in the same way by moving the pencil

up and down so that its path falls between the two horizontal marks. To prove that information from the region of visual space bounded by the marks is really not perceived, put a penny inside the demarcated area. When you fixate the X with both eyes and then close the left eye, the penny will disappear, a seemingly magical event that amazed the French royal court when it was first reported by the natural philosopher Edmé Mariotte in 1668.

How can we be unaware of such a large defect in the visual field (typically about 5°–8°)? The optic disk is located in the nasal retina of each eye. With both eyes open, information about the corresponding region of visual space is, of course, available from the temporal retina of the other eye. But this fact does not explain why the blind spot remains undetected with one eye closed. When the world is viewed monocularly, the visual system appears to "fill-in" the missing part of the scene based on the information supplied by the regions surrounding the optic disk. To observe this phenomenon, notice what happens when a pencil or some other object lies across the optic disk representation. Remarkably, the pencil looks complete! Although electrophysiological recordings have shown that neurons in the visual

cortex whose receptive fields lie in the optic disk representation can be activated by stimulating the regions that surround the optic disk of the contralateral eye, suggesting that "filling-in" the blind spot is based on cortical mechanisms that integrate information from different points in the visual field, the mechanism of this striking phenomenon is not clear. Herman von Helmholtz pointed out in the nineteenth century that it may just be that this part of the visual world is ignored, the pencil being completed across the blind spot because the rest of the scene simply "collapses" around it.

References

- FIORANI, M., M. G. P. ROSA, R. GATTASS AND C. E. ROCHA-MIRANDA (1992) Dynamic surrounds of receptive fields in striate cortex: A physiological basis for perceptual completion? *Proc. Natl. Acad. Sci. USA* 89: 8547–8551.
- GILBERT, C. D. (1992) Horizontal integration and cortical dynamics. *Neuron* 9: 1–13.
- RAMACHANDRAN, V. S. AND T. L. GREGORY (1991) Perceptual filling in of artificially induced scotomas in human vision. *Nature* 350: 699–702.
- VON HELMHOLTZ, H. (1968). *Helmholtz's Treatise on Physiological Optics*, Vols. I–III (Translated from the Third German Ed. published in 1910). J. P. C. Southall (ed.). New York: Dover Publications. See pp. 204ff in Vol. III.



ate the reflex. Under normal conditions, the pupils of both eyes respond identically, regardless of which eye is stimulated; that is, light in one eye produces constriction of both the stimulated eye (the direct response) and the unstimulated eye (the consensual response; see Figure 11.3). Comparing the response in the two eyes is often helpful in localizing a lesion. For example, a direct response in the left eye without a consensual response in the right eye suggests a problem with the visceral motor outflow to the right eye, possibly as a result of damage to the oculomotor nerve or Edinger-Westphal nucleus in the brainstem. Failure to elicit a response (either direct or indirect) to stimulation of the left eye if both eyes respond normally to stimulation of the right eye suggests damage to the sensory input from the left eye, possibly to the left retina or optic nerve.

There are several other important targets of retinal ganglion cell axons. One is the **suprachiasmatic nucleus** of the hypothalamus, a small group of neurons at the base of the diencephalon (see Box A in Chapter 20). The **retino-hypothalamic pathway** is the route by which variation in light levels influences the broad spectrum of visceral functions that are entrained to the day/night cycle (see Chapters 20 and 27). Another target is the **superior colliculus**, a prominent structure visible on the dorsal surface of the midbrain (see Figure 1.14). The superior colliculus coordinates head and eye movements to visual (as well as other) targets; its functions are considered in Chapter 19.

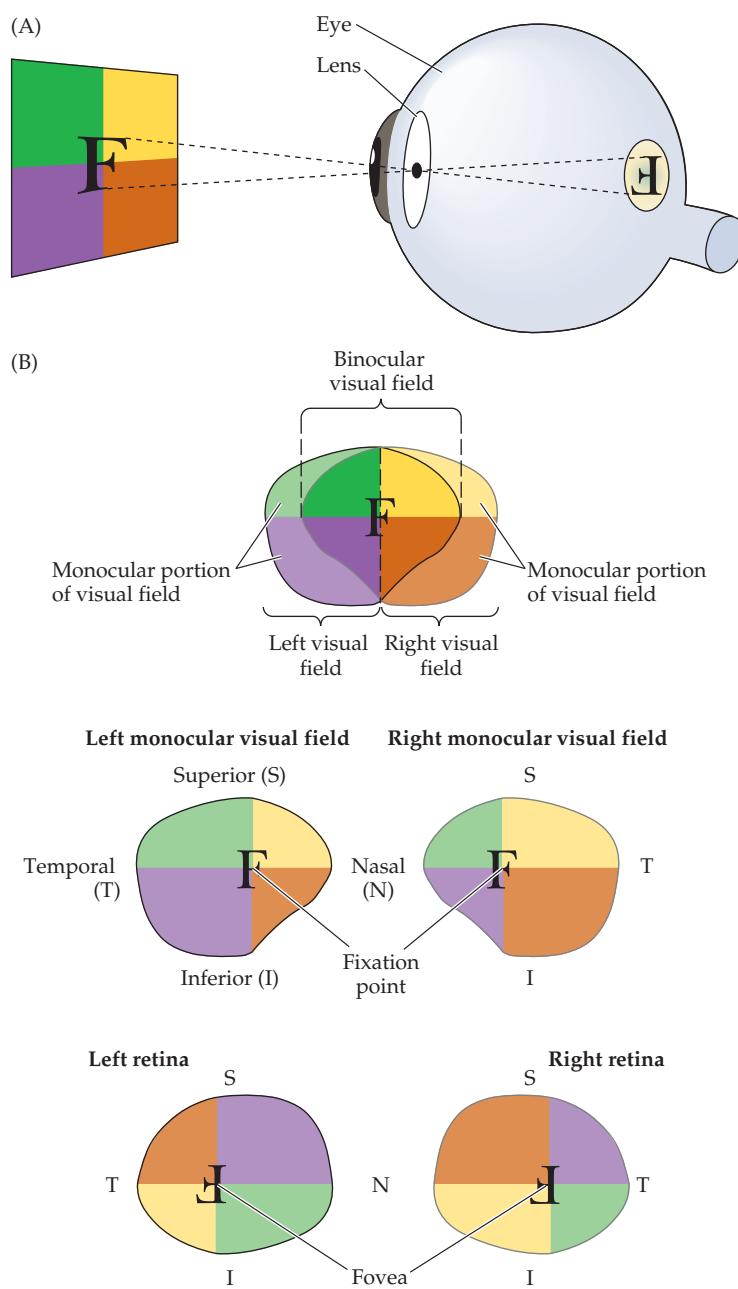
The type of visual information required to perform the functions of these different retinal targets is quite different. Reading the text on this page, for example, requires a high-resolution sampling of the retinal image, whereas regulating circadian rhythms and adjusting the pupil accordingly require only a measure of overall changes in light levels, and little or no information about the features of the image. It should come as no surprise, then, that there is a diversity of ganglion cell types that provide information appropriate to the functions of these different targets.

Projections to the lateral geniculate nucleus (which are described in more detail later) arise from at least three broad classes of ganglion cells, whose tuning properties are appropriate for mediating the richness of visual perception (high acuity, color, motion). In contrast, projections to the hypothalamus and the pretectum arise from ganglion cells that lack these properties and are highly suited for detecting luminance flux. The retinal specializations responsible for constructing these distinct classes of retinal ganglion cells are only beginning to be identified; they include not only differences in ganglion cell synaptic connections, but in the locus of the phototransduction event itself. Unlike the majority of ganglion cells, which depend on rods and cones for their sensitivity to light, the ganglion cells that project to the hypothalamus and pretectum express their own light-sensitive photopigment (*melanopsin*) and are capable of modulating their response to changes in light levels in the absence of signals from rods and cones. The presence of light sensitivity within this class of ganglion cells presumably explains why normal circadian rhythms are maintained in animals that have completely lost form vision due to degeneration of rod and cone photoreceptors.

The Retinotopic Representation of the Visual Field

The spatial relationships among the ganglion cells in the retina are maintained in most of their central targets as orderly representations or “maps” of visual space. Most of these structures receive information from both eyes, requiring that these inputs be integrated to form a coherent map of individ-

Figure 11.4 Projection of the visual fields onto the left and right retinas. (A) Projection of an image onto the surface of the retina. The passage of light rays through the pupil of the eye results in images that are inverted and left-right reversed on the retinal surface. (B) Retinal quadrants and their relation to the organization of monocular and binocular visual fields, as viewed from the back surface of the eyes. Vertical and horizontal lines drawn through the center of the fovea define retinal quadrants (bottom). Comparable lines drawn through the point of fixation define visual field quadrants (center). Color coding illustrates corresponding retinal and visual field quadrants. The overlap of the two monocular visual fields is shown at the top.



ual points in space. As a general rule, information from the left half of the visual world, whether it originates from the left or right eye, is represented in the right half of the brain, and vice versa.

Understanding the neural basis for the appropriate arrangement of inputs from the two eyes requires considering how images are projected onto the two retinas, and the central destination of the ganglion cells located in different parts of the retina. Each eye sees a part of visual space that defines its **visual field** (Figure 11.4A). For descriptive purposes, each retina and its corresponding visual field are divided into quadrants. In this scheme, the surface of the retina is subdivided by vertical and horizontal lines that intersect at the center of the fovea (Figure 11.4B). The vertical line divides the retina into **nasal** and **temporal divisions** and the horizontal line divides the retina

into **superior** and **inferior divisions**. Corresponding vertical and horizontal lines in visual space (also called meridians) intersect at the **point of fixation** (the point in visual space that falls on the fovea) and define the quadrants of the visual field. The crossing of light rays diverging from different points on an object at the pupil causes the images of objects in the visual field to be inverted and left-right reversed on the retinal surface. As a result, objects in the temporal part of the visual field are seen by the nasal part of the retina, and objects in the superior part of the visual field are seen by the inferior part of the retina. (It may help in understanding Figure 11.4B to imagine that you are looking at the back surfaces of the retinas, with the corresponding visual fields projected onto them.)

With both eyes open, the two foveas are normally aligned on a single target in visual space, causing the visual fields of both eyes to overlap extensively (see Figure 11.4B and Figure 11.5). This **binocular field** of view consists of two symmetrical visual hemifields (left and right). The left binocular hemifield includes the nasal visual field of the right eye and the temporal visual field of the left eye; the right hemifield includes the temporal visual field of

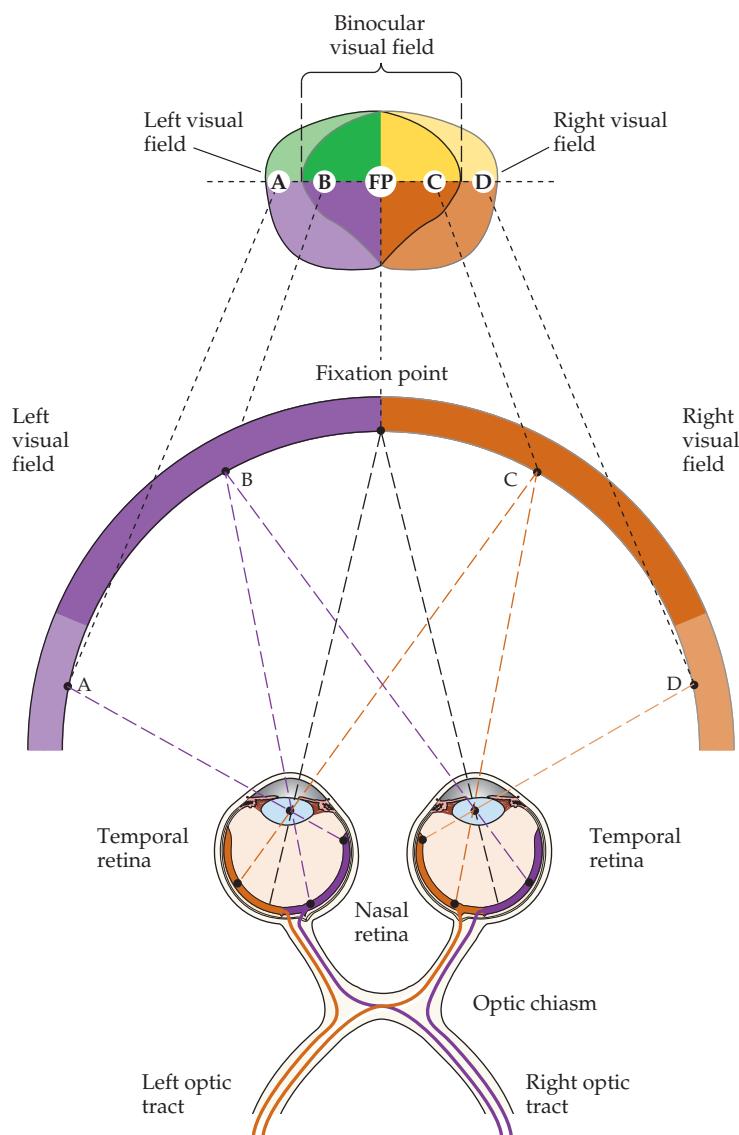


Figure 11.5 Projection of the binocular field of view onto the two retinas and its relation to the crossing of fibers in the optic chiasm. Points in the binocular portion of the left visual field (B) fall on the nasal retina of the left eye and the temporal retina of the right eye. Points in the binocular portion of the right visual field (C) fall on the nasal retina of the right eye and the temporal retina of the left eye. Points that lie in the monocular portions of the left and right visual fields (A and D) fall on the left and right nasal retinas, respectively. The axons of ganglion cells in the nasal retina cross in the optic chiasm, whereas those from the temporal retina do not. As a result, information from the left visual field is carried in the right optic tract, and information from the right visual field is carried in the left optic tract.

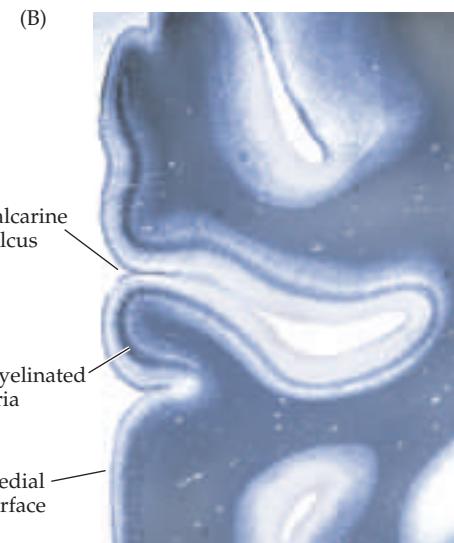
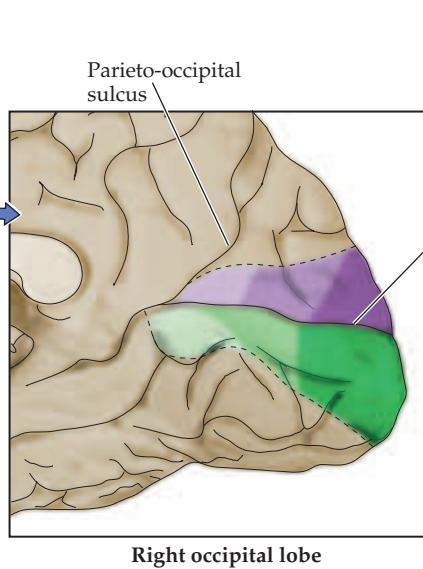
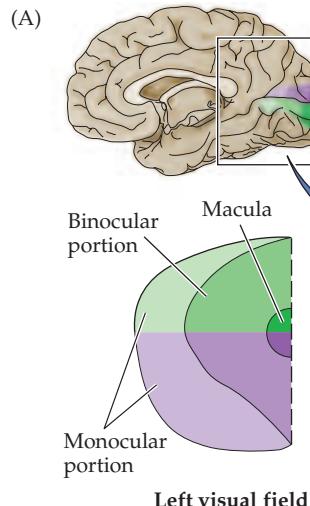
the right eye and the nasal visual field of the left eye. The temporal visual fields are more extensive than the nasal visual fields, reflecting the size of the nasal and temporal retinas respectively. As a result, vision in the periphery of the field of view is strictly monocular, mediated by the most medial portion of the nasal retina. Most of the rest of the field of view can be seen by both eyes; i.e., individual points in visual space lie in the nasal visual field of one eye and the temporal visual field of the other. It is worth noting, however, that the shape of the face and nose impact the extent of this region of binocular vision. In particular, the inferior nasal visual fields are less extensive than the superior nasal fields, and consequently the binocular field of view is smaller in the lower visual field than in the upper (see Figure 11.4B).

Ganglion cells that lie in the nasal division of each retina give rise to axons that cross in the chiasm, while those that lie in the temporal retina give rise to axons that remain on the same side (see Figure 11.5). The boundary (or line of decussation) between contralaterally and ipsilaterally projecting ganglion cells runs through the center of the fovea and defines the border between the nasal and temporal hemiretinas. Images of objects in the left visual hemifield (such as point B in Figure 11.5) fall on the nasal retina of the left eye and the temporal retina of the right eye, and the axons from ganglion cells in these regions of the two retinas project through the right optic tract. Objects in the right visual hemifield (such as point C in Figure 11.5) fall on the nasal retina of the right eye and the temporal retina of the left eye; the axons from ganglion cells in these regions project through the left optic tract. As mentioned previously, objects in the monocular portions of the visual hemifields (points A and D in Figure 11.5) are seen only by the most peripheral nasal retina of each eye; the axons of ganglion cells in these regions (like the rest of the nasal retina) run in the contralateral optic tract. Thus, unlike the optic nerve, the optic tract contains the axons of ganglion cells that originate in both eyes and represent the contralateral field of view.

Optic tract axons terminate in an orderly fashion within their target structures thus generating well ordered maps of the contralateral hemifield. For the primary visual pathway, the map of the contralateral hemifield that is established in the lateral geniculate nucleus is maintained in the projections of the lateral geniculate nucleus to the striate cortex (Figure 11.6). Thus the



Figure 11.6 Visuotopic organization of the striate cortex in the right occipital lobe, as seen in mid-sagittal view. (A) The primary visual cortex occupies a large part of the occipital lobe. The area of central vision (the fovea) is represented over a disproportionately large part of the caudal portion of the lobe, whereas peripheral vision is represented more anteriorly. The upper visual field is represented below the calcarine sulcus, the lower field above the calcarine sulcus. (B) Photomicrograph of a coronal section of the human striate cortex, showing the characteristic myelinated band, or stria, that gives this region of the cortex its name. The calcarine sulcus on the medial surface of the occipital lobe is indicated. (B courtesy of T. Andrews and D. Purves.)



fovea is represented in the posterior part of the striate cortex, whereas the more peripheral regions of the retina are represented in progressively more anterior parts of the striate cortex. The upper visual field is mapped below the calcarine sulcus, and the lower visual field above it. As in the somatic sensory system, the amount of cortical area devoted to each unit area of the sensory surface is not uniform, but reflects the density of receptors and sensory axons that supply the peripheral region. Like the representation of the hand region in the somatic sensory cortex, the representation of the macula is therefore disproportionately large, occupying most of the caudal pole of the occipital lobe.

Visual Field Deficits

A variety of retinal or more central pathologies that involve the primary visual pathway can cause visual field deficits that are limited to particular regions of visual space. Because the spatial relationships in the retinas are maintained in central visual structures, a careful analysis of the visual fields can often indicate the site of neurological damage. Relatively large visual field deficits are called **anopsias** and smaller ones are called **scotomas** (see Box A). The former term is combined with various prefixes to indicate the specific region of the visual field from which sight has been lost (Figures 11.7 and 11.8).

Damage to the retina or one of the optic nerves before it reaches the chiasm results in a loss of vision that is limited to the eye of origin. In contrast, damage in the region of the optic chiasm—or more centrally—results in specific types of deficits that involve the visual fields of both eyes (Figure 11.8). Damage to structures that are central to the optic chiasm, including the optic tract, lateral geniculate nucleus, optic radiation, and visual cortex, results in deficits that are limited to the contralateral visual hemifield. For example, interruption of the optic tract on the right results in a loss of sight in the left visual field (that is, blindness in the temporal visual field of the left eye and the nasal visual field of the right eye). Because such damage affects corresponding parts of the visual field in each eye, there is a complete loss of vision in the affected region of the binocular visual field, and the deficit is referred to as a **homonymous hemianopsia** (in this case, a left homonymous hemianopsia).

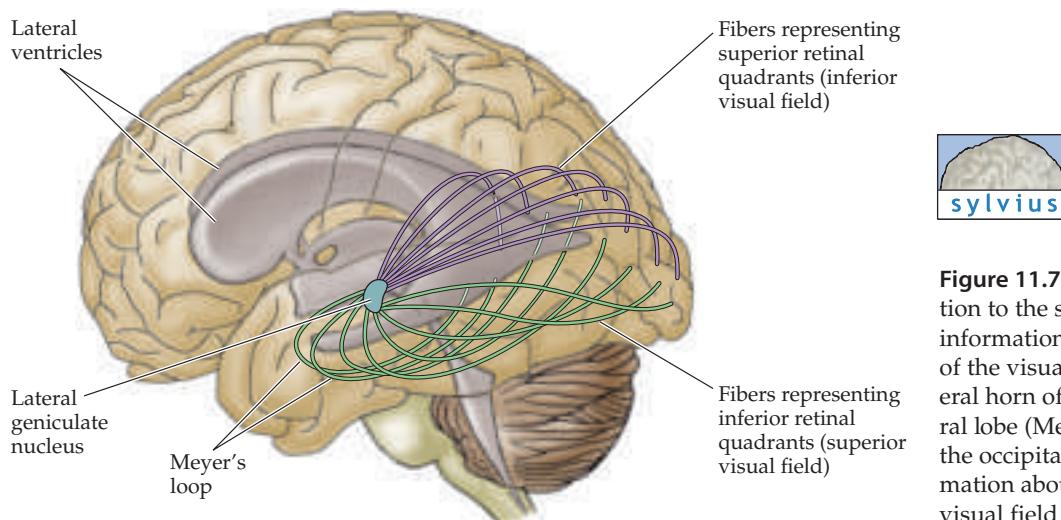


Figure 11.7 Course of the optic radiation to the striate cortex. Axons carrying information about the superior portion of the visual field sweep around the lateral horn of the ventricle in the temporal lobe (Meyer’s loop) before reaching the occipital lobe. Those carrying information about the inferior portion of the visual field travel in the parietal lobe.

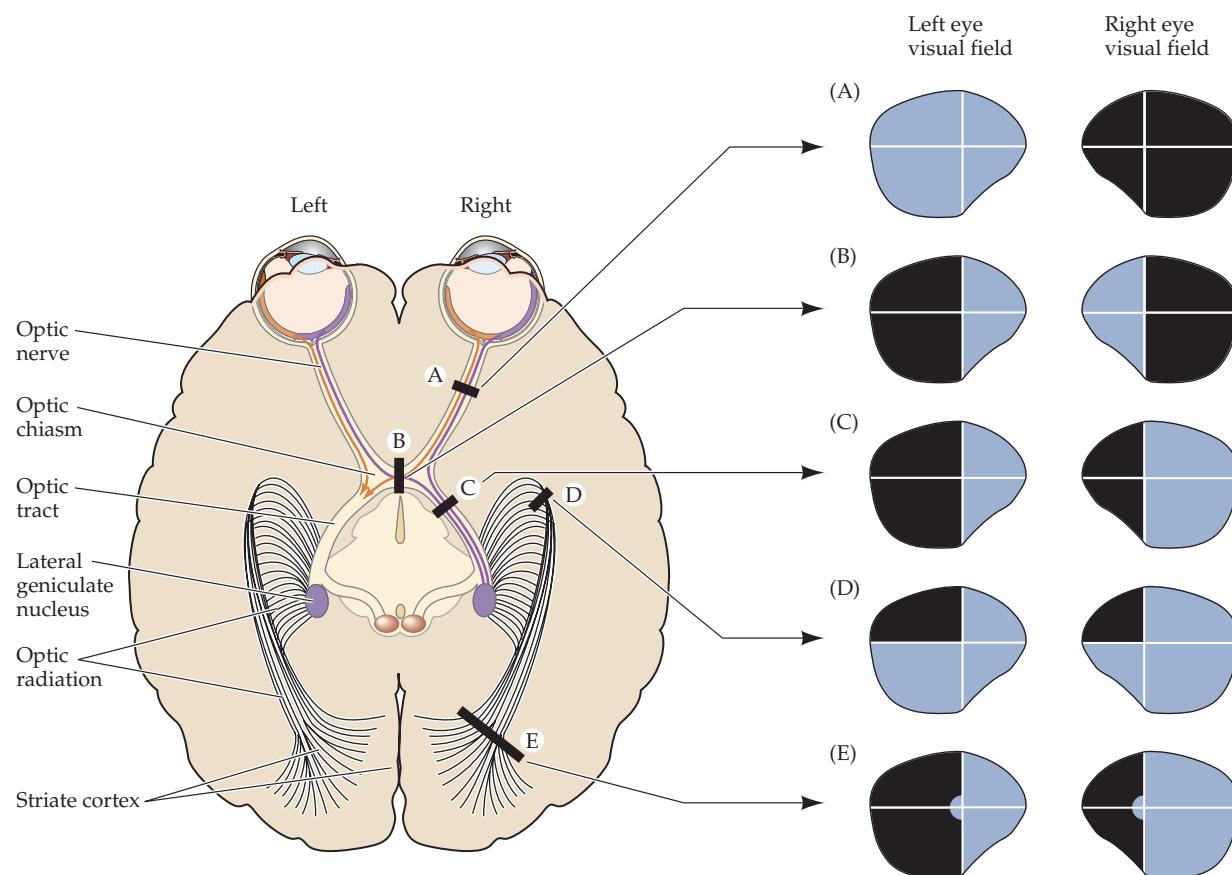


Figure 11.8 Visual field deficits resulting from damage at different points along the primary visual pathway. The diagram on the left illustrates the basic organization of the primary visual pathway and indicates the location of various lesions. The right panels illustrate the visual field deficits associated with each lesion. (A) Loss of vision in right eye. (B) Bitemporal (heteronomous) hemianopsia. (C) Left homonymous hemianopsia. (D) Left superior quadrantanopsia. (E) Left homonymous hemianopsia with macular sparing.



In contrast, damage to the optic chiasm results in visual field deficits that involve noncorresponding parts of the visual field of each eye. For example, damage to the middle portion of the optic chiasm (which is often the result of pituitary tumors) can affect the fibers that are crossing from the nasal retina of each eye, leaving the uncrossed fibers from the temporal retinas intact. The resulting loss of vision is confined to the temporal visual field of each eye and is known as **bitemporal hemianopsia**. It is also called **heteronomous hemianopsia** to emphasize that the parts of the visual field that are lost in each eye do not overlap. Individuals with this condition are able to see in both left and right visual fields, provided both eyes are open. However, all information from the most peripheral parts of visual fields (which are seen only by the nasal retinas) is lost.

Damage to central visual structures is rarely complete. As a result, the deficits associated with damage to the chiasm, optic tract, optic radiation, or visual cortex are typically more limited than those shown in Figure 11.8. This is especially true for damage along the optic radiation, which fans out under the temporal and parietal lobes in its course from the lateral geniculate nucleus to the striate cortex. Some of the optic radiation axons run out into the temporal lobe on their route to the striate cortex, a branch called **Meyer's loop** (see Figure 11.7). Meyer's loop carries information from the superior portion of the contralateral visual field. More medial parts of the optic radiation, which pass under the cortex of the parietal lobe, carry information from the inferior portion of the contralateral visual field. Damage to parts of the temporal lobe with involvement of Meyer's loop can thus result in a superior

homonymous quadrantanopsia; damage to the optic radiation underlying the parietal cortex results in an inferior homonymous quadrantanopsia.

Injury to central visual structures can also lead to a phenomenon called *macular sparing*, i.e., the loss of vision throughout wide areas of the visual field, with the exception of foveal vision. Macular sparing is commonly found with damage to the cortex, but can be a feature of damage anywhere along the length of the visual pathway. Although several explanations for macular sparing have been offered, including overlap in the pattern of crossed and uncrossed ganglion cells supplying central vision, the basis for this selective preservation is not clear.

The Functional Organization of the Striate Cortex

Much in the same way that Stephen Kuffler explored the response properties of individual retinal ganglion cells (see Chapter 10), David Hubel and Torsten Wiesel used microelectrode recordings to examine the properties of neurons in more central visual structures.

The responses of neurons in the lateral geniculate nucleus were found to be remarkably similar to those in the retina, with a center-surround receptive field organization and selectivity for luminance increases or decreases. However, the small spots of light that were so effective at stimulating neurons in the retina and lateral geniculate nucleus were largely ineffective in visual cortex. Instead, most cortical neurons in cats and monkeys responded vigorously to light-dark bars or edges, and only if the bars were presented at a particular range of orientations within the cell's receptive field (Figure 11.9). The responses of cortical neurons are thus tuned to the orientation of edges, much like cone receptors are tuned to the wavelength of light; the peak in the tuning curve (the orientation to which a cell is most responsive) is referred to as the neuron's preferred orientation. By sampling the responses of a large number of single cells, Hubel and Weisel demonstrated that all edge orientations were roughly equally represented in visual cortex. As a

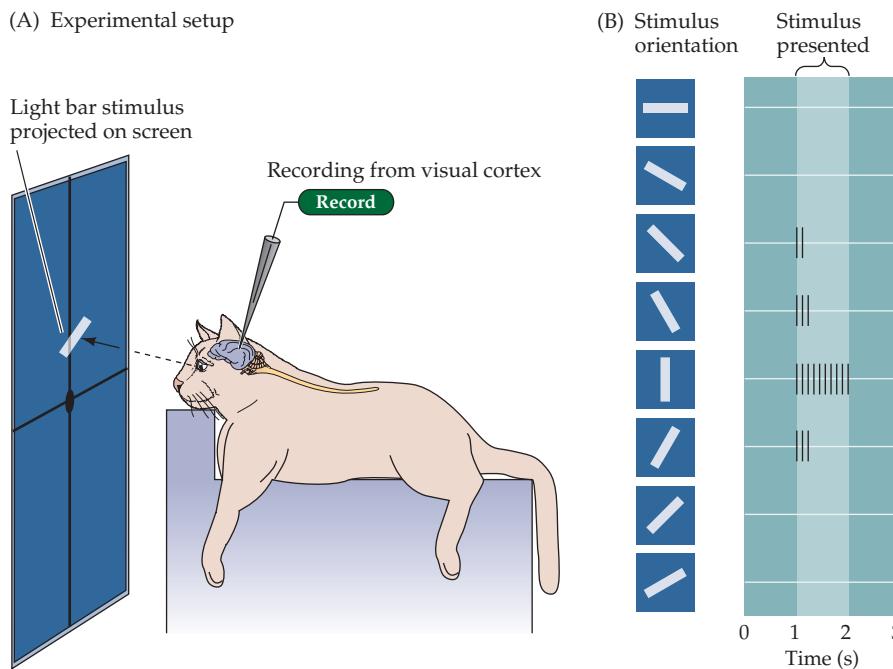


Figure 11.9 Neurons in the primary visual cortex respond selectively to oriented edges. (A) An anesthetized animal is fitted with contact lenses to focus the eyes on a screen, where images can be projected; an extracellular electrode records the neuronal responses. (B) Neurons in the primary visual cortex typically respond vigorously to a bar of light oriented at a particular angle and weakly—or not at all—to other orientations.

result, a given orientation in a visual scene appears to be “encoded” in the activity of a distinct population of **orientation-selective neurons**.

Hubel and Wiesel also found that there are subtly different subtypes within a class of neurons that preferred the same orientation. For example, the receptive fields of some cortical cells, which they called **simple cells**, were composed of spatially separate “on” and “off” response zones, as if the “on” and “off” centers of lateral geniculate cells that supplied these neurons were arrayed in separate parallel bands. Other neurons, referred to as **complex cells**, exhibited mixed “on” and “off” responses throughout their receptive field, as if they received their inputs from a number of simple cells. Further analysis uncovered cortical neurons sensitive to the *length* of the bar of light that was moved across their receptive field, decreasing their rate of response when the bar exceeded a certain length. Still other cells responded selectively to the *direction* in which an edge moved across their receptive field. Although the mechanisms responsible for generating these selective responses are still not well understood, there is little doubt that the specificity of the receptive field properties of neurons in the striate cortex (and beyond) plays an important role in determining the basic attributes of visual scenes.

Another feature that distinguishes the responses of neurons in the striate cortex from those at earlier stages in the primary visual pathway is **binocularity**. Although the lateral geniculate nucleus receives inputs from both eyes, the axons terminate in separate layers, so that individual geniculate

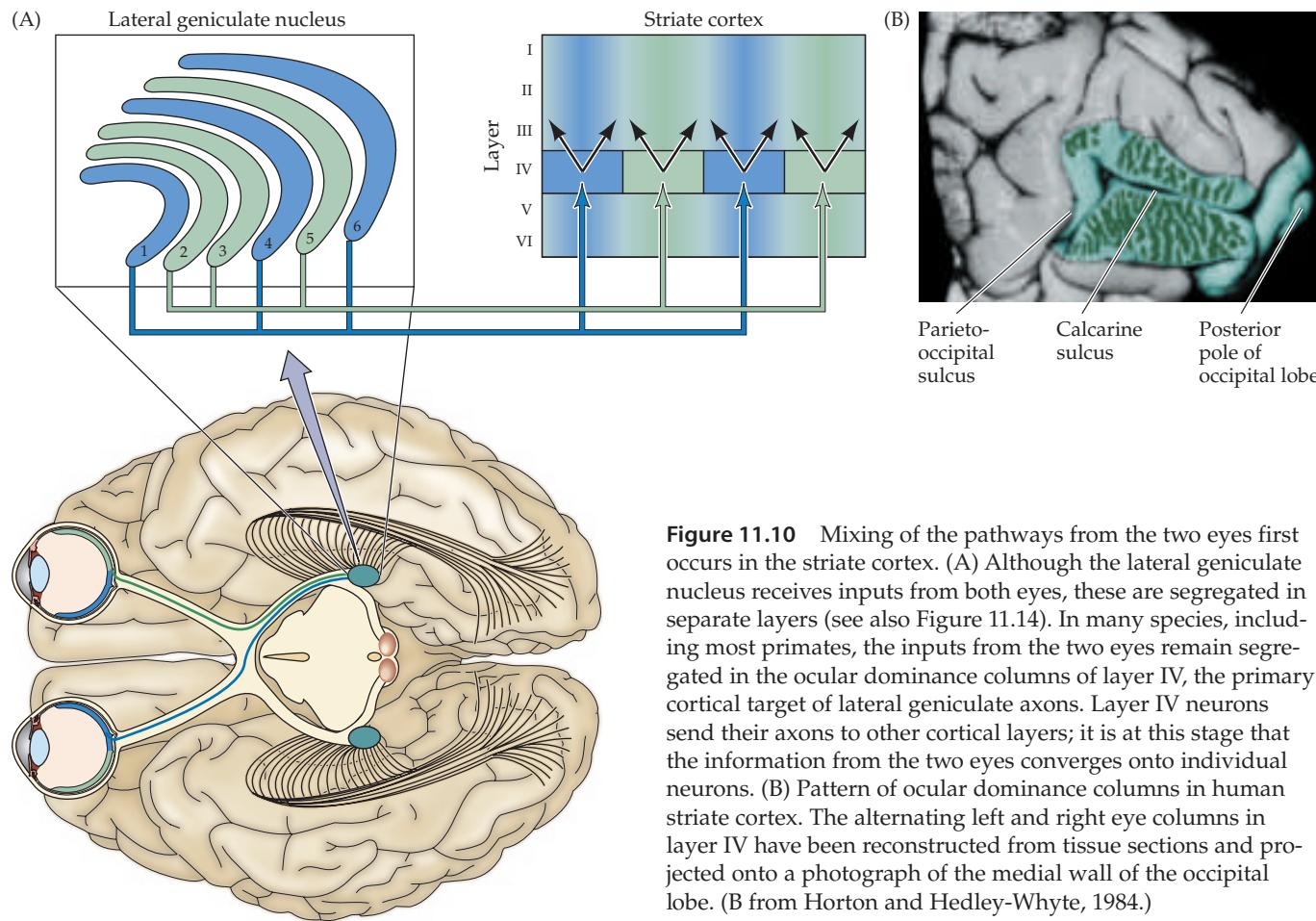


Figure 11.10 Mixing of the pathways from the two eyes first occurs in the striate cortex. (A) Although the lateral geniculate nucleus receives inputs from both eyes, these are segregated in separate layers (see also Figure 11.14). In many species, including most primates, the inputs from the two eyes remain segregated in the ocular dominance columns of layer IV, the primary cortical target of lateral geniculate axons. Layer IV neurons send their axons to other cortical layers; it is at this stage that the information from the two eyes converges onto individual neurons. (B) Pattern of ocular dominance columns in human striate cortex. The alternating left and right eye columns in layer IV have been reconstructed from tissue sections and projected onto a photograph of the medial wall of the occipital lobe. (B from Horton and Hedley-Whyte, 1984.)

neurons are monocular, driven by either the left or right eye but not by both (Figure 11.10; see also Figure 11.14). In some species, including most (but not all) primates, inputs from the left and right eyes remain segregated to some degree even beyond the geniculate because the axons of geniculate neurons terminate in alternating eye-specific columns within cortical layer IV—the so-called **ocular dominance columns** (see the next section). Beyond this point, the signals from the two eyes are combined at the cellular level. Thus, most cortical neurons have binocular receptive fields, and these fields are almost identical, having the same size, shape, preferred orientation, and roughly the same position in the visual field of each eye.

Bringing together the inputs from the two eyes at the level of the striate cortex provides a basis for **stereopsis**, the special sensation of depth that arises from viewing nearby objects with two eyes instead of one. Because the two eyes look at the world from slightly different angles, objects that lie in front of or behind the plane of fixation project to noncorresponding points on the two retinas. To convince yourself of this fact, hold your hand at arm's length and fixate on the tip of one finger. Maintain fixation on the finger as you hold a pencil in your other hand about half as far away. At this distance, the image of the pencil falls on noncorresponding points on the two retinas and will therefore be perceived as two separate pencils (a phenomenon called double vision, or *diplopia*). If the pencil is now moved toward the finger (the point of fixation), the two images of the pencil fuse and a single pencil is seen in front of the finger. Thus, for a small distance on either side of the plane of fixation, where the disparity between the two views of the world remains modest, a single image is perceived; the disparity between the two eye views of objects nearer or farther than the point of fixation is interpreted as depth (Figure 11.11).

Although the neurophysiological basis of stereopsis is not understood, some neurons in the striate cortex and in other visual cortical areas have receptive field properties that make them good candidates for extracting information about binocular disparity. Unlike many binocular cells whose monocular receptive fields sample the same region of visual space, these neurons have monocular fields that are slightly displaced (or perhaps differ in their internal organization) so that the cell is maximally activated by stimuli that fall on noncorresponding parts of the retinas. Some of these neurons (so-called **far cells**) discharge to disparities beyond the plane of fixation, while others (**near cells**) respond to disparities in front of the plane of fixation. The pattern of activity in these different classes of neurons seems likely to contribute to sensations of stereoscopic depth (Box B).

Interestingly, the preservation of the binocular responses of cortical neurons is contingent on the normal activity from the two eyes during early postnatal life. Anything that creates an imbalance in the activity of the two eyes—for example, the clouding of one lens or the abnormal alignment of the eyes during infancy (strabismus)—can permanently reduce the effectiveness of one eye in driving cortical neurons, and thus impair the ability to use binocular information as a cue for depth. Early detection and correction of visual problems is therefore essential for normal visual function in maturity (see Chapter 23).

The Columnar Organization of the Striate Cortex

The variety of response properties exhibited by cortical neurons raises the question of how neurons with different receptive fields are arranged within striate cortex. For the most part, the responses of neurons are qualitatively

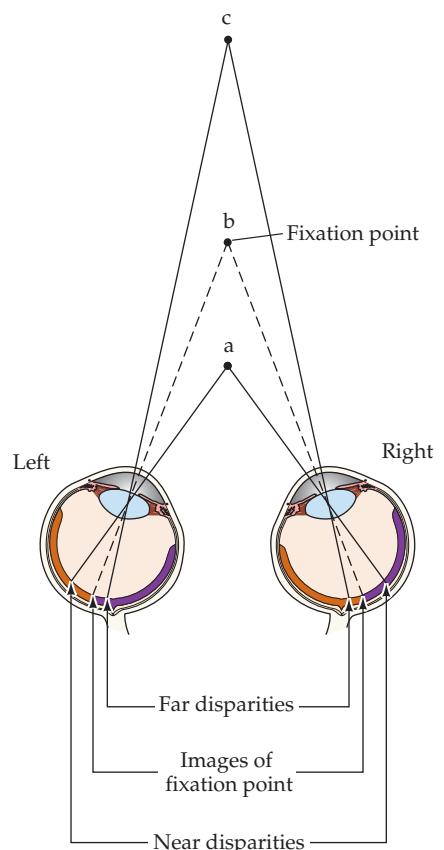


Figure 11.11 Binocular disparities are generally thought to be the basis of stereopsis. When the eyes are fixated on point b, points that lie beyond the plane of fixation (point c) or in front of the point of fixation (point a) project to noncorresponding points on the two retinas. When these disparities are small, the images are fused and the disparity is interpreted by the brain as small differences in depth. When the disparities are greater, double vision occurs (although this normal phenomenon is generally unnoticed).

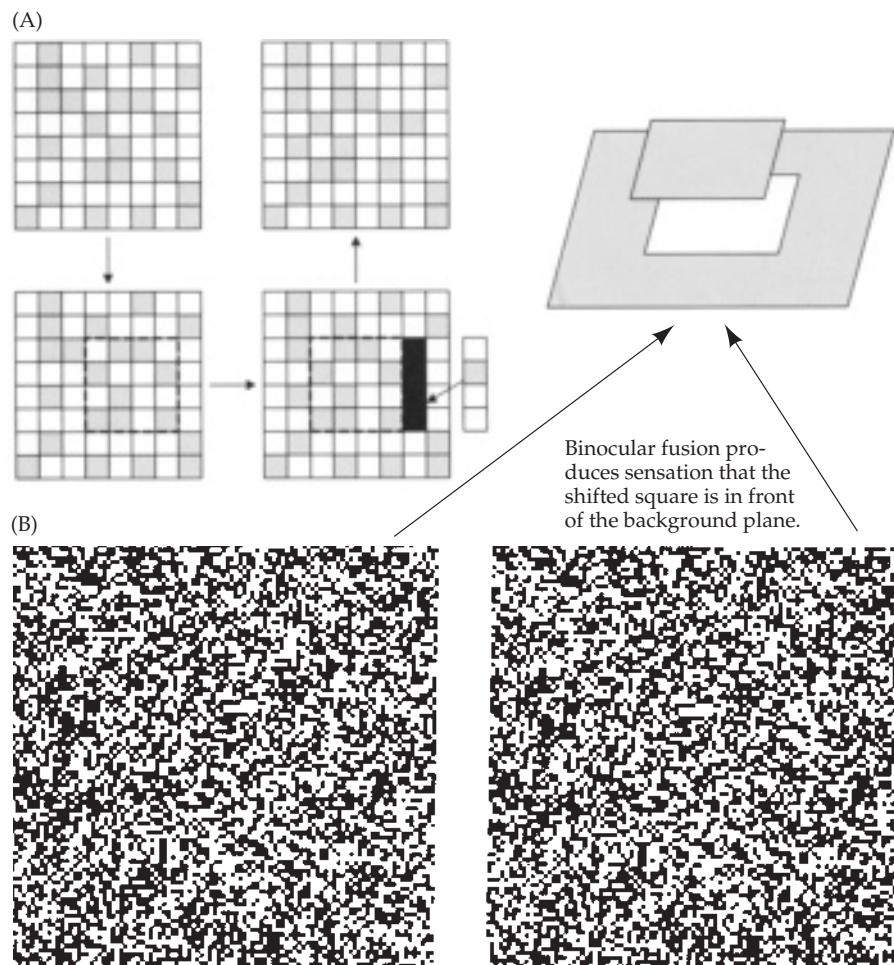
Box B

Random Dot Stereograms and Related Amusements

An important advance in studies of stereopsis was made in 1959 when Bela Julesz, then working at the Bell Laboratories in Murray Hill, New Jersey, discovered an ingenious way of showing that stereoscopy depends on matching information seen by the two eyes without any prior recognition of what object(s) such matching might generate. Julesz, a Hungarian whose background was in engineering and physics, was working on the problem of how to “break” camouflage. He surmised that the brain’s ability to fuse the slightly different views of the two eyes to bring out new information would be an aid in overcoming military camouflage. Julesz also realized that, if his hypothesis was correct, a hidden figure in a random pattern presented to the two eyes should emerge when a portion of the otherwise identical pattern was shifted horizontally in the view of one eye or the other. A horizontal shift in one direction would cause the hidden object to appear in front of the plane of the background, whereas a shift in the other direction would cause the hidden object to appear in back of the plane. Such a figure, called a random dot stereogram, and the method of its creation are shown in Figures A and B. The two images can be easily fused in a stereoscope (like the

familiar Viewmaster® toy) but can also be fused simply by allowing the eyes to diverge. Most people find it easiest to do this by imagining that they are looking “through” the figure; after some seconds, during which the brain tries to make sense of what it is presented with, the two images merge and the hidden figure appears (in this case, a square that occupies the middle portion of the figure). The random dot stereogram has been widely used in stereoscopic research for about 40 years, although how such stimuli elicit depth remains very much a matter of dispute.

An impressive—and extraordinarily popular—derivative of the random dot stereogram is the autostereogram (Figure C). The possibility of autostereograms was first discerned by the nineteenth-century British physicist David Brewster. While staring at a Victorian wallpaper with an iterated but offset pattern, he noticed that when the patterns were fused, he perceived two different planes. The plethora of autostereograms that can be seen today in posters, books, and newspapers are close cousins of the random dot stereogram in that computers are used to shift patterns of iterated



Random dot stereograms and autostereograms. (A) to construct a random dot stereogram, a random dot pattern is created to be observed by one eye. The stimulus for the other eye is created by copying the first image, displacing a particular region horizontally, and then filling in the gap with a random sample of dots. (B) When the right and left images are viewed simultaneously but independently by the two eyes (by using a stereoscope or fusing the images by converging or diverging the eyes), the shifted region (a square) appears to be in a different plane from the other dots. (A after Wandell, 1995.)

information with respect to each other. The result is that different planes emerge from what appears to be a meaningless array of visual information (or, depending on the taste of the creator, an apparently "normal" scene in which the iterated and displaced information is hidden). Some autostereograms are designed to reveal the hidden figure when the eyes diverge, and others when they converge. (Looking at a plane more distant than the plane of the surface causes divergence; looking at a plane in front of the picture causes the eyes to converge; see Figure 11.11.)

The elevation of the autostereogram to a popular art form should probably be attributed to Chris W. Tyler, a student of Julesz's and a visual psychophysicist, who was among the first to create commercial autostereograms. Numerous graphic artists—preeminently in Japan, where the popularity of the autostereogram has been enormous—have gener-

ated many of such images. As with the random dot stereogram, the task in viewing the autostereogram is not clear to the observer. Nonetheless, the hidden figure emerges, often after minutes of effort in which the brain automatically tries to make sense of the occult information.

References

- JULESZ, B. (1971) *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.
 JULESZ, B. (1995) *Dialogues on Perception*. Cambridge, MA: MIT Press.
 N. E. THING ENTERPRISES (1993) *Magic Eye: A New Way of Looking at the World*. Kansas City: Andrews and McMeel.

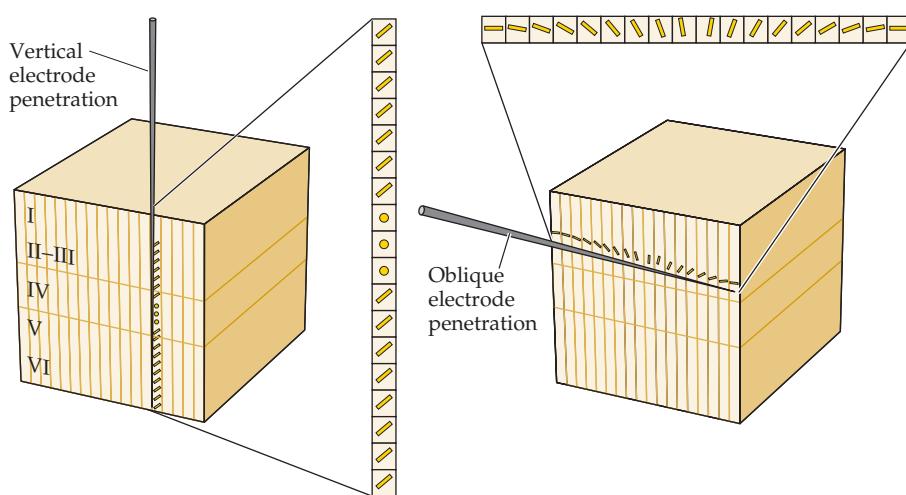
(C)



(C) An autostereogram. The hidden figure (three geometrical forms) emerges by diverging the eyes in this case. (C courtesy of Jun Oi.)

similar at any one point in primary visual cortex, but tend to shift smoothly across its surface. With respect to orientation, for example, all the neurons encountered in an electrode penetration perpendicular to the surface at a particular point will very likely have the same orientation preference, forming a "column" of cells with similar response properties. Adjacent columns, however, usually have slightly different orientation preferences; the sequence of orientation preferences encountered along a tangential electrode penetration gradually shifts as the electrode advances (Figure 11.12). Thus, orientation preference is mapped in the cortex, much like receptive field

Figure 11.12 Columnar organization of orientation selectivity in the monkey striate cortex. Vertical electrode penetrations encounter neurons with the same preferred orientations, whereas oblique penetrations show a systematic change in orientation across the cortical surface. The circles denote the lack of orientation-selective cells in layer IV.



location (Box C). Unlike the map of visual space, however, the map of orientation preference is iterated many times, such that the same orientation preference is repeated at approximately 1-mm intervals across the striate cortex. This iteration presumably ensures that there are neurons for each region of visual space that represent the full range of orientation values. The orderly progression of orientation preference (as well as other properties that are mapped in this systematic way) is accommodated within the orderly map of visual space by the fact that the mapping is relatively coarse. Each small region of visual space is represented by a set of neurons whose receptive fields cover the full range of orientation preferences, the set being distributed over several millimeters of the cortical surface

The columnar organization of the striate cortex is equally apparent in the binocular responses of cortical neurons. Although most neurons in the striate cortex respond to stimulation of both eyes, the relative strength of the inputs from the two eyes varies from neuron to neuron. At the extremes of this continuum are neurons that respond almost exclusively to the left or right eye; in the middle are those that respond equally well to both eyes. As in the case of orientation preference, vertical electrode penetrations tend to encounter neurons with similar ocular preference (or **ocular dominance**, as it is usually called), whereas tangential penetrations show gradual shifts in ocular dominance. And, like the arrangement of orientation preference, a movement of about a millimeter across the surface is required to sample the full complement of ocular dominance values (Figure 11.13). These shifts in ocular dominance result from the ocular segregation of the inputs from lateral geniculate nucleus within cortical layer IV (see Figure 11.10).

Although the modular arrangement of the visual cortex was first recognized on the basis of these orientation and ocular dominance columns, further work has shown that other stimulus features such as color, direction of motion, and spatial frequency also tend to be distributed in iterated patterns that are systematically related to each other (for example, orientation columns tend to intersect ocular dominance columns at right angles). In short, the striate cortex is composed of repeating units, or modules, that contain all the neuronal machinery necessary to analyze a small region of visual space for a variety of different stimulus attributes. As described in Box D in Chapter 8, a number of other cortical regions show a similar columnar arrangement of their processing circuitry.

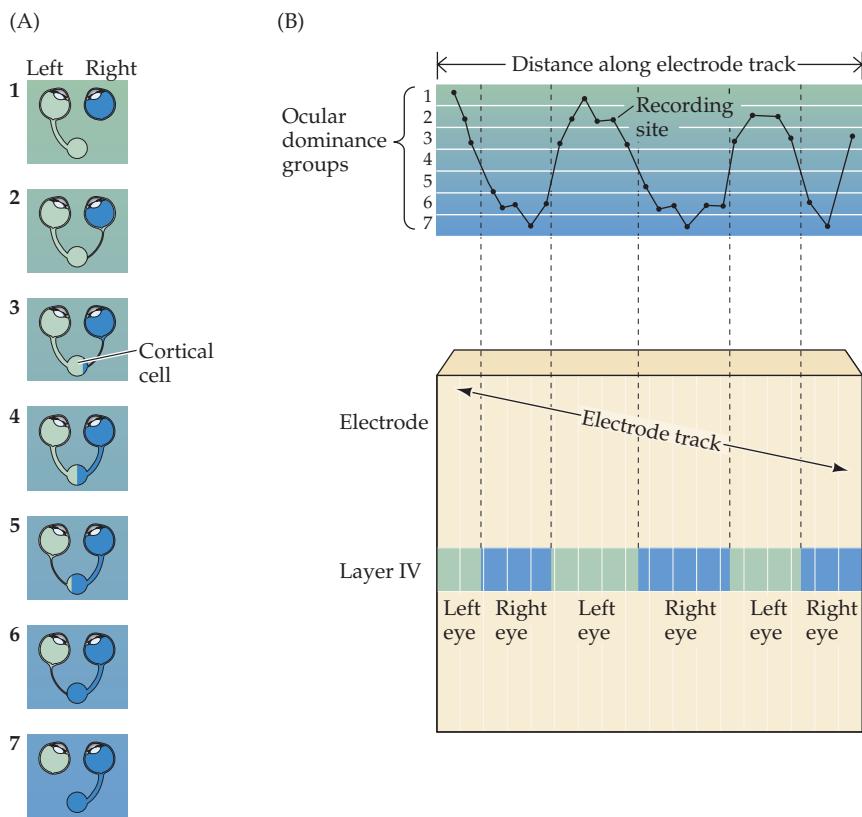


Figure 11.13 Columnar organization of ocular dominance. (A) Cortical neurons in all layers vary in the strength of their response to the inputs from the two eyes, from complete domination by one eye to equal influence of the two eyes. (B) Tangential electrode penetration across the superficial cortical layers reveals a gradual shift in the ocular dominance of the recorded neurons from one eye to the other. In contrast, all neurons encountered in a vertical electrode penetration (other than those neurons that lie in layer IV) tend to have the same ocular dominance.

Division of Labor within the Primary Visual Pathway

In addition to being specific for input from one eye or the other, the layers in the lateral geniculate are also distinguished on the basis of cell size: Two ventral layers are composed of large neurons and are referred to as the **magnocellular layers**, while more dorsal layers are composed of small neurons and are referred to as the **parvocellular layers**. The magno- and parvocellular layers receive inputs from distinct populations of ganglion cells that exhibit corresponding differences in cell size. M ganglion cells that terminate in the magnocellular layers have larger cell bodies, more extensive dendritic fields, and larger-diameter axons than the P ganglion cells that terminate in the parvocellular layers (Figure 11.14A). Moreover, the axons of relay cells in the magno- and parvocellular layers of the lateral geniculate nucleus terminate on distinct populations of neurons located in separate strata within layer 4 of striate cortex. Thus the retinogeniculate pathway is composed of parallel **magnocellular and parvocellular streams** that convey distinct types of information to the initial stages of cortical processing.

The response properties of the M and P ganglion cells provide important clues about the contributions of the magno- and parvocellular streams to visual perception. M ganglion cells have larger receptive fields than P cells, and their axons have faster conduction velocities. M and P ganglion cells also differ in ways that are not so obviously related to their morphology. M cells respond transiently to the presentation of visual stimuli, while P cells respond in a sustained fashion. Moreover, P ganglion cells can transmit information about color, whereas M cells cannot. P cells convey color information because their receptive field centers and surrounds are driven by different classes of cones (i.e., cones responding with greatest sensitivity to

Box C

Optical Imaging of Functional Domains in the Visual Cortex

The recent availability of optical imaging techniques has made it possible to visualize how response properties, such as the selectivity for edge orientation or ocular dominance, are mapped across the cortical surface. These methods generally rely on intrinsic signals (changes in the amount of light reflected from the cortical surface) that correlate with levels of neural activity. Such signals are thought to arise at least in part from local changes in the ratio of oxyhemoglobin and deoxyhemoglobin that accompany such activity, more active areas having a higher deoxyhemoglobin/oxyhemoglobin ratio (see also Box A in Chapter 1). This change can be detected when the cortical surface is illuminated with red light (605–700 nm). Under these conditions, active cortical regions absorb more light than less active ones. With the use of a sensitive video camera, and averaging over a number of trials (the changes are small, 1 or 2 parts per thousand), it is possible to visualize these differences and use them to map cortical patterns of activity (Figure A).

This approach has now been successfully applied to both striate and extrastri-

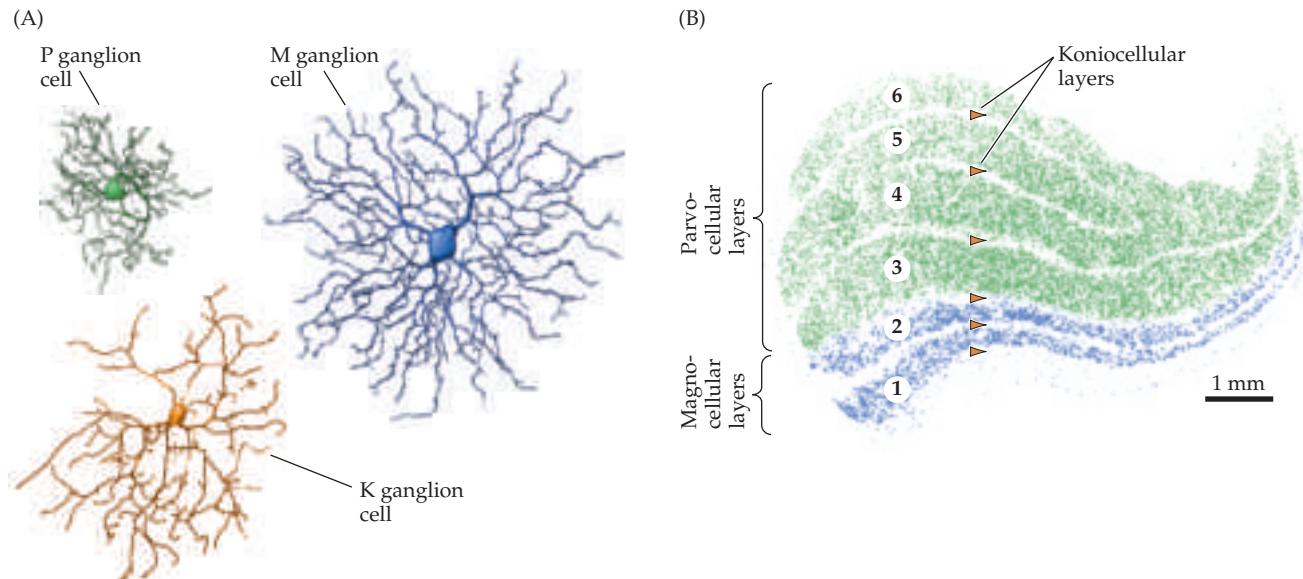
ate areas in both experimental animals and human patients undergoing neurosurgery. The results emphasize that maps of stimulus features are a general principle of cortical organization. For example, orientation preference is mapped in a continuous fashion such that adjacent positions on the cortical surface tend to have only slightly shifted orientation preferences. However, there are points where continuity breaks down. Around these points, orientation preference is represented in a radial pattern resembling a pinwheel, covering the whole 180° of possible orientation values (Figure B).

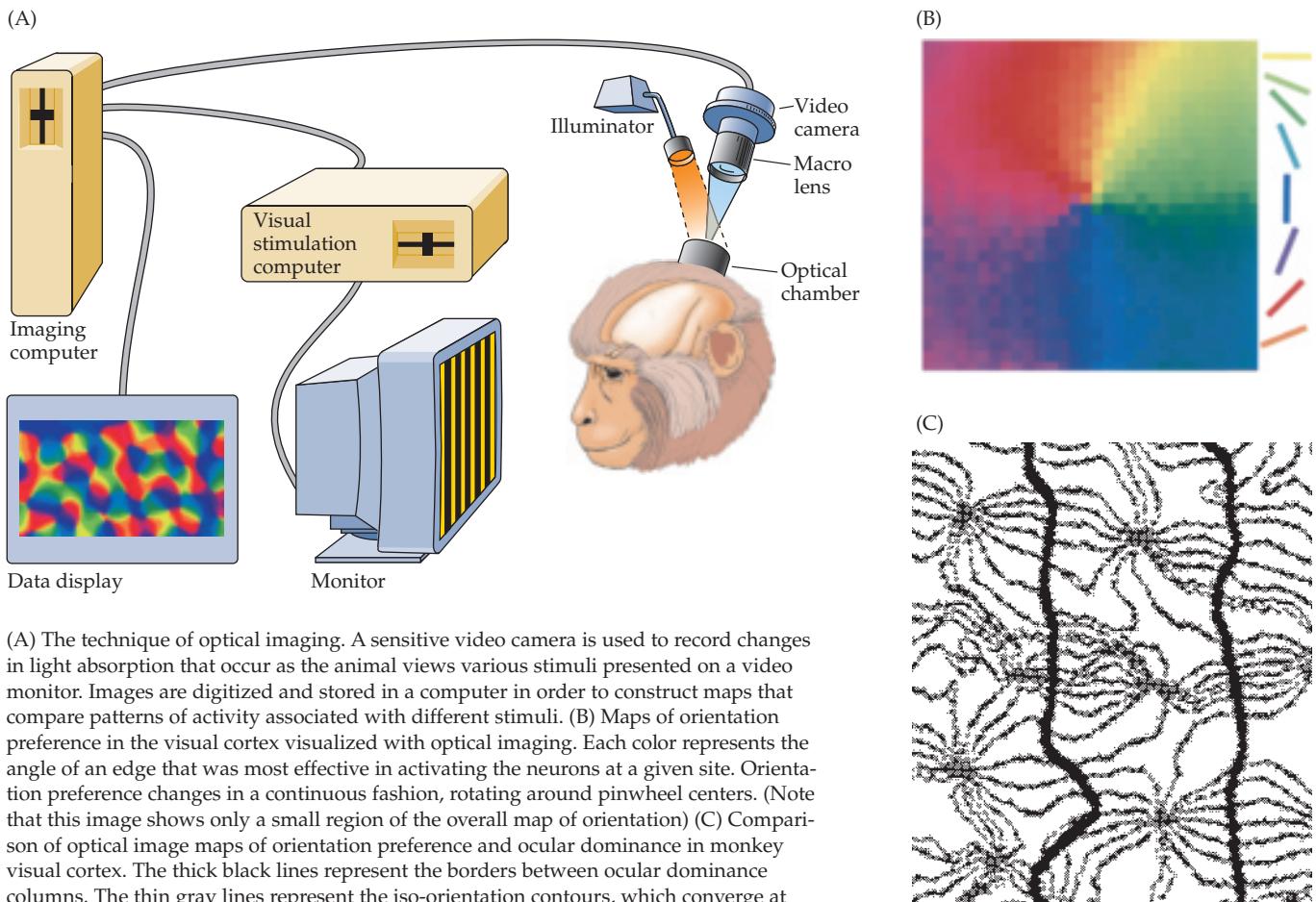
This powerful technique can also be used to determine how maps for different stimulus properties are arranged relative to one another, and to detect additional maps such as that for direction of motion. A comparison of ocular dominance bands and orientation preference maps, for example, shows that pinwheel centers are generally located in the center of ocular dominance bands, and that the iso-orientation contours that emanate from the pinwheel centers run orthogonal to the borders of ocular dominance bands (Figure C). An orderly relation-

ship between maps of orientation selectivity and direction selectivity has also been demonstrated. These systematic relationships between the functional maps that coexist within primary visual cortex are thought to ensure that all combinations of stimulus features (orientation, direction, ocular dominance, and spatial frequency) are analyzed for all regions of visual space.

References

- BLASDEL, G. G. AND G. SALAMA (1986) Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* 321: 579–585.
- BONHOEFFER, T. AND A. GRINVALD (1993) The layout of iso-orientation domains in area 18 of the cat visual cortex: Optical imaging reveals a pinwheel-like organization. *J. Neurosci* 13: 4157–4180.
- BONHOEFFER, T. AND A. GRINVALD (1996) Optical imaging based on intrinsic signals: The methodology. In *Brain Mapping: The Methods*, A. Toga (ed.). New York: Academic Press.
- OBERMAYER, K. AND G. G. BLASDEL (1993) Geometry of orientation and ocular dominance columns in monkey striate cortex. *J. Neurosci*. 13: 4114–4129.
- WELIKY, M., W. H. BOSKING AND D. FITZPATRICK (1996) A systematic map of direction preference in primary visual cortex. *Nature* 379: 725–728.





(A) The technique of optical imaging. A sensitive video camera is used to record changes in light absorption that occur as the animal views various stimuli presented on a video monitor. Images are digitized and stored in a computer in order to construct maps that compare patterns of activity associated with different stimuli. (B) Maps of orientation preference in the visual cortex visualized with optical imaging. Each color represents the angle of an edge that was most effective in activating the neurons at a given site. Orientation preference changes in a continuous fashion, rotating around pinwheel centers. (Note that this image shows only a small region of the overall map of orientation) (C) Comparison of optical image maps of orientation preference and ocular dominance in monkey visual cortex. The thick black lines represent the borders between ocular dominance columns. The thin gray lines represent the iso-orientation contours, which converge at orientation pinwheel centers (arrow). Iso-orientation contour lines generally intersect the borders of ocular dominance bands at right angles. (B from Bonhoeffer and Grinvald, 1993; C from Obermeyer and Blasdel, 1993.)

short-, medium-, or long-wavelength light). For example, some P ganglion cells have centers that receive inputs from long-wavelength (“red”) sensitive cones and surrounds that receive inputs from medium-wavelength (“green”) cones. Others have centers that receive inputs from “green cones” and surrounds from “red cones” (see Chapter 10). As a result, P cells are sensitive to differences in the wavelengths of light striking their receptive field center

► Figure 11.14 Magno- and parvocellular streams. (A) Tracings of M and P ganglion cells as seen in flat mounts of the retina after staining by the Golgi method. M cells have large-diameter cell bodies and large dendritic fields. They supply the magnocellular layers of the lateral geniculate nucleus. P cells have smaller cell bodies and dendritic fields. They supply the parvocellular layers of the lateral geniculate nucleus. (B) Photomicrograph of the human lateral geniculate nucleus showing the magnocellular and parvocellular layers. (A after Watanabe and Rodieck, 1989; B courtesy of T. Andrews and D. Purves.)

and surround. Although M ganglion cells also receive inputs from cones, there is no difference in the type of cone input to the receptive field center and surround; the center and surround of each M cell receptive field is driven by all cone types. The absence of cone specificity to center-surround antagonism makes M cells largely insensitive to differences in the wavelengths of light that strike their receptive field centers and surrounds, and they are thus unable to transmit color information to their central targets.

The contribution of the magno- and parvocellular streams to visual perception has been tested experimentally by examining the visual capabilities of monkeys after selectively damaging either the magno- or parvocellular layers of the lateral geniculate nucleus. Damage to the magnocellular layers has little effect on visual acuity or color vision, but sharply reduces the ability to perceive rapidly changing stimuli. In contrast, damage to the parvocellular layers has no effect on motion perception but severely impairs visual acuity and color perception. These observations suggest that the visual information conveyed by the parvocellular stream is particularly important for high spatial resolution vision—the detailed analysis of the shape, size, and color of objects. The magnocellular stream, on the other hand, appears critical for tasks that require high temporal resolution, such as evaluating the location, speed and direction of a rapidly moving object.

In addition to the magno- and parvocellular streams, a third distinct anatomical pathway—the **koniocellular, or K-cell pathway**—has been identified within the lateral geniculate nucleus. Neurons contributing to the K-cell pathway reside in the interlaminar zones that separate lateral geniculate layers; these neurons receive inputs from fine-caliber retinal axons and project in a patchy fashion to the superficial layers (layers II and III) of striate cortex. Although the contribution of the K-cell pathway to perception is not understood, it appears that some aspects of color vision, especially information derived from short-wavelength-sensitive cones, may be transmitted via the K-cell rather than the P-cell pathway. Why short-wavelength-sensitive cone signals should be processed differently from middle- and long-wavelength information is not clear, but the distinction may reflect the earlier evolutionary origin of the K-cell pathway (see Chapter 10).

The Functional Organization of Extrastriate Visual Areas

Anatomical and electrophysiological studies in monkeys have led to the discovery of a multitude of areas in the occipital, parietal, and temporal lobes that are involved in processing visual information (Figure 11.15). Each of these areas contains a map of visual space, and each is largely dependent on the primary visual cortex for its activation. The response properties of the neurons in some of these regions suggest that they are specialized for different aspects of the visual scene. For example, the **middle temporal area (MT)** contains neurons that respond selectively to the direction of a moving edge without regard to its color. In contrast, neurons in another cortical area called **V4** respond selectively to the color of a visual stimulus without regard to its direction of movement. These physiological findings are supported by behavioral evidence; thus, damage to area MT leads to a specific impairment in a monkey's ability to perceive the direction of motion in a stimulus pattern, while other aspects of visual perception remain intact.

Recent functional imaging studies have indicated a similar arrangement of visual areas within human extrastriate cortex. Using retinotopically restricted stimuli, it has been possible to localize at least 10 separate repre-

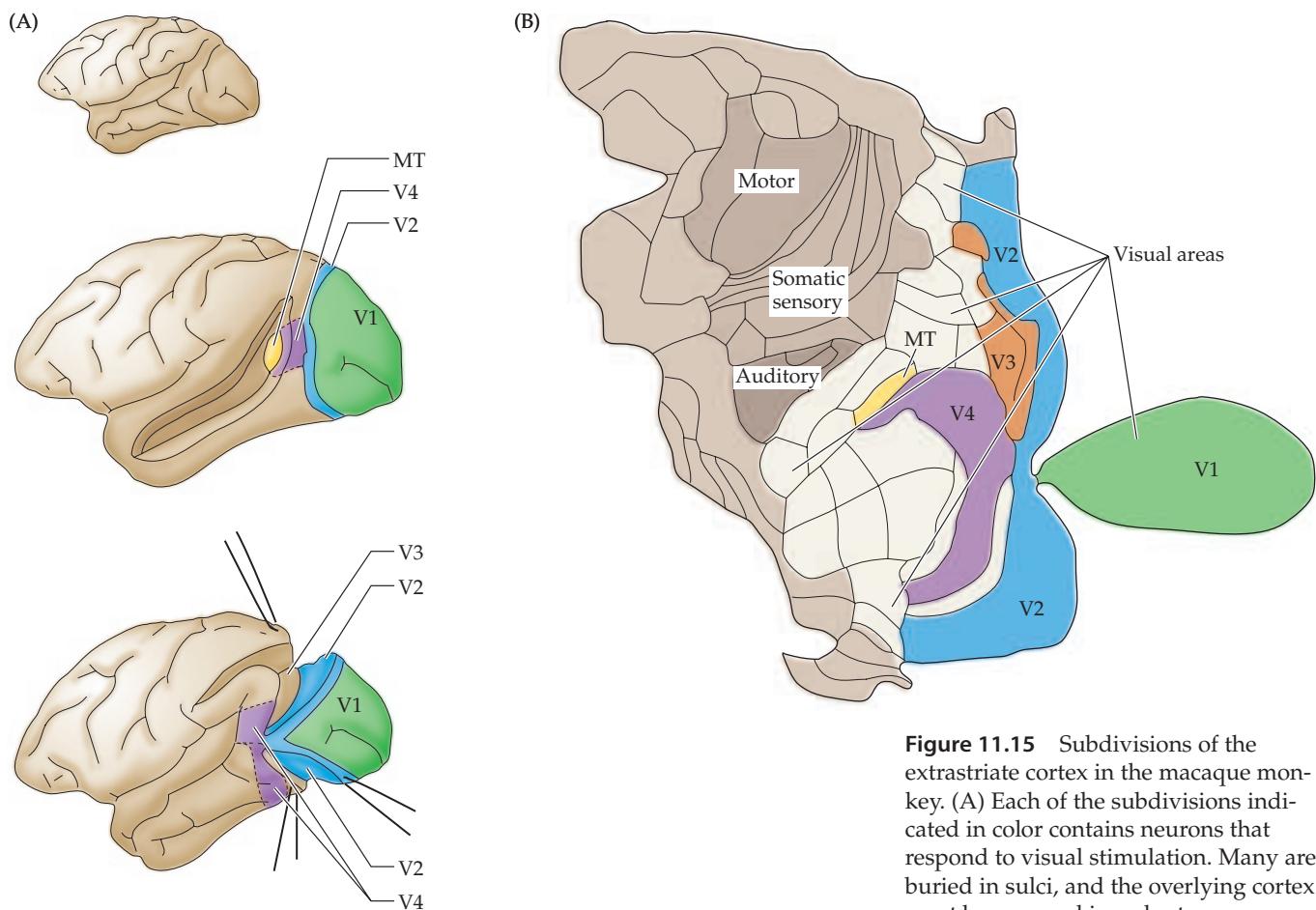


Figure 11.15 Subdivisions of the extrastriate cortex in the macaque monkey. (A) Each of the subdivisions indicated in color contains neurons that respond to visual stimulation. Many are buried in sulci, and the overlying cortex must be removed in order to expose them. Some of the more extensively studied extrastriate areas are specifically identified (V2, V3, V4, and MT). V1 is the primary visual cortex; MT is the middle temporal area. (B) The arrangement of extrastriate and other areas of neocortex in a flattened view of the monkey neocortex. There are at least 25 areas that are predominantly or exclusively visual in function, plus 7 other areas suspected to play a role in visual processing. (A after Maunsell and Newson, 1987; B after Felleman and Van Essen, 1991.)

sentations of the visual field (Figure 11.16). One of these areas exhibits a large motion-selective signal, suggesting that it is the homologue of the motion-selective middle temporal area described in monkeys. Another area exhibits color-selective responses, suggesting that it may be similar to V4 in non-human primates. A role for these areas in the perception of motion and color, respectively, is further supported by evidence for increases in activity not only during the presentation of the relevant stimulus, but also during periods when subjects experience motion or color afterimages.

The clinical description of selective visual deficits after localized damage to various regions of extrastriate cortex also supports functional specialization of extrastriate visual areas in humans. For example, a well-studied patient who suffered a stroke that damaged the extrastriate region thought to be comparable to area MT in the monkey was unable to appreciate the motion of objects. The neurologist who treated her noted that she had difficulty in pouring tea into a cup because the fluid seemed to be "frozen." In addition, she could not stop pouring at the right time because she was unable to perceive when the fluid level had risen to the brim. The patient also had trouble following a dialogue because she could not follow the movements of the speaker's mouth. Crossing the street was potentially terrifying because she couldn't judge the movement of approaching cars. As the patient related, "When I'm looking at the car first, it seems far away. But

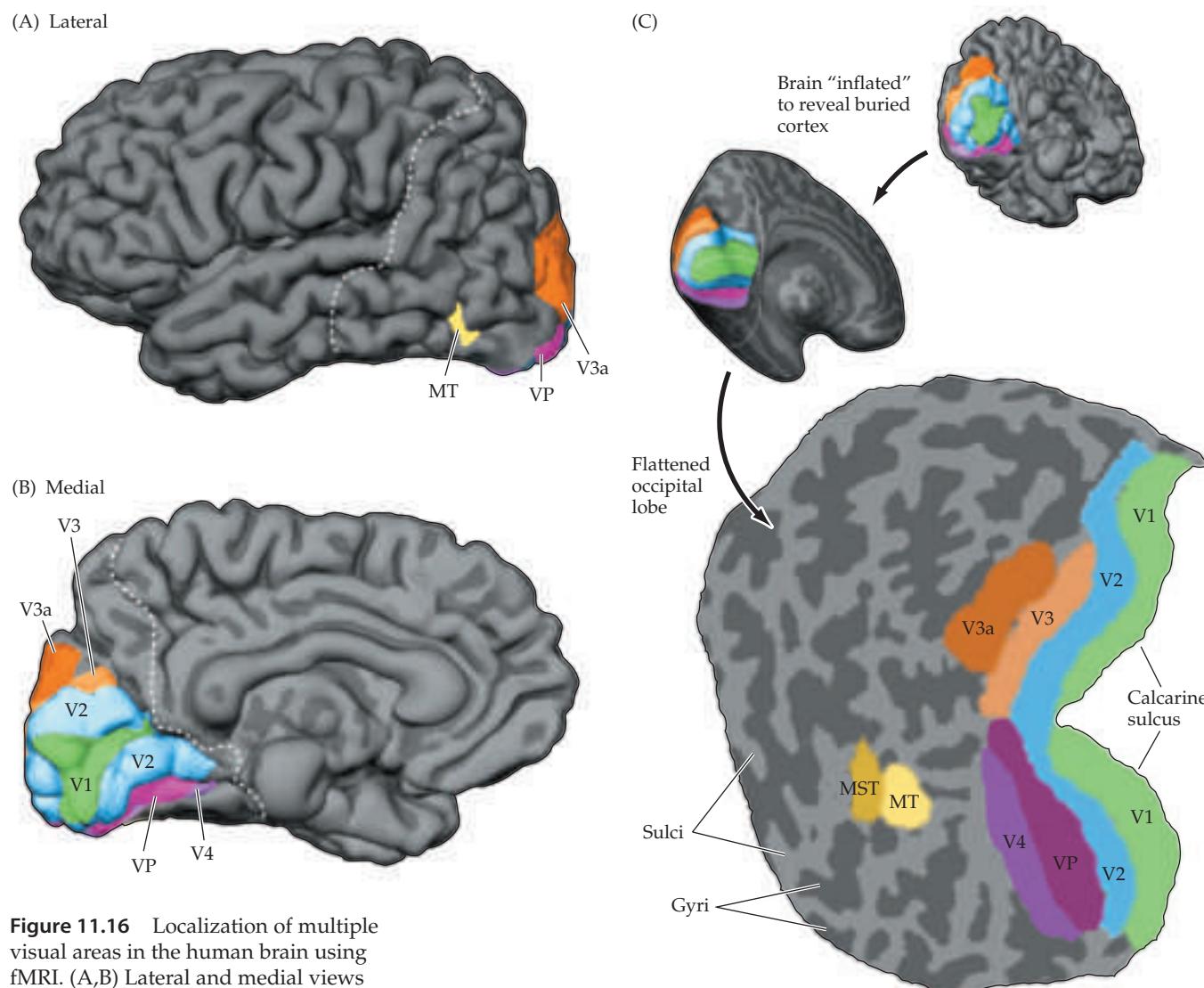


Figure 11.16 Localization of multiple visual areas in the human brain using fMRI. (A,B) Lateral and medial views (respectively) of the human brain, illustrating the location of primary visual cortex (V1) and additional visual areas V2, V3, VP (ventral posterior area), V4, MT (middle temporal area), and MST (medial superior temporal area). (C) Unfolded and flattened view of retinotopically defined visual areas in the occipital lobe. Dark grey areas correspond to cortical regions that were buried in sulci; light regions correspond to regions that were located on the surface of gyri. Visual areas in humans show a close resemblance to visual areas originally defined in monkeys (compare with Figure 11.15). (After Sereno et al., 1995.)

then, when I want to cross the road, suddenly the car is very near." Her ability to perceive other features of the visual scene, such as color and form, was intact.

Another example of a specific visual deficit as a result of damage to extrastriate cortex is **cerebral achromatopsia**. These patients lose the ability to see the world in color, although other aspects of vision remain in good working order. The normal colors of a visual scene are described as being replaced by "dirty" shades of gray, much like looking at a poor quality black-and-white movie. Achromatopsic individuals know the normal colors of objects—that a school bus is yellow, an apple red—but can no longer see them. Thus, when asked to draw objects from memory, they have no difficulty with shapes but are unable to appropriately color the objects they have represented. It is important to distinguish this condition from the color blindness that arises from the congenital absence of one or more cone pigments in the retina (see Chapter 10). In achromatopsia, the three types of cones are functioning normally; it is damage to specific extrastriate cortical areas that renders the patient unable to use the information supplied by the retina.