Enhanced DIGITAL VERSION Included

COSTANZO PHYSIOLOGY

7TH EDITION

LINDA S. COSTANZO

ELSEVIER

SEVENTH EDITION

Physiology

LINDA S. COSTANZO, PhD

Professor of Physiology and Biophysics
Virginia Commonwealth University School of Medicine
Richmond, Virginia

ELSEVIER

ELSEVIER

1600 John F. Kennedy Blvd.
Ste 1800
Philadelphia, PA 19103-2899

PHYSIOLOGY SEVENTH EDITION ISBN: 978-0-323-47881-6

Notices

Knowledge and best practice in this field are constantly changing. As new research and experience broaden our understanding, changes in research methods, professional practices, or medical treatment may become necessary.

Practitioners and researchers must always rely on their own experience and knowledge in evaluating and using any information, methods, compounds, or experiments described herein. In using such information or methods they should be mindful of their own safety and the safety of others, including parties for whom they have a professional responsibility.

With respect to any drug or pharmaceutical products identified, readers are advised to check the most current information provided (i) on procedures featured or (ii) by the manufacturer of each product to be administered, to verify the recommended dose or formula, the method and duration of administration, and contraindications. It is the responsibility of practitioners, relying on their own experience and knowledge of their patients, to make diagnoses, to determine dosages and the best treatment for each individual patient, and to take all appropriate safety precautions.

To the fullest extent of the law, neither the Publisher nor the authors, contributors, or editors, assume any liability for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions, or ideas contained in the material herein.

Previous editions copyrighted 2021, 2010, 2006, 2002, and 1998.

Library of Congress Cataloging-in-Publication Data

Names: Costanzo, Linda S., 1947- author.
Title: Physiology / Linda S. Costanzo.
Other titles: Physiology (Elsevier)
Description: Sixth edition. | Philadelphia, PA : Elsevier, [2018] | Includes index.
Identifiers: LCCN 2017002153 | ISBN 9780323478816 (pbk.)
Subjects: | MESH: Physiological Phenomena | Physiology
Classification: LCC QP31.2 | NLM QT 104 | DDC 612–dc23
LC record available at https://lccn.loc.gov/2017002153

Executive Content Strategist: Elyse O'Grady
Senior Content Development Specialist: Jennifer Ehlers
Publishing Services Manager: Catherine Jackson
Senior Project Manager: Daniel Fitzgerald
Designer: Renee Duenow

Cover image: Laguna Design/Nerve Cell, abstract artwork/Getty Images

Printed in China.

Last digit is the print number: 9 8 7 6 5 4 3 2 1

To
Heinz Valtin and **Arthur C. Guyton,**
who have written so well for students of physiology

Richard, Dan, Rebecca, Sheila, Elise, and **Max,**
who make everything worthwhile

Preface

Physiology is the foundation of medical practice. A firm grasp of its principles is essential for the medical student and the practicing physician. This book is intended for students of medicine and related disciplines who are engaged in the study of physiology. It can be used either as a companion to lectures and syllabi in discipline-based curricula or as a primary source in integrated or problem-based curricula. For advanced students, the book can serve as a reference in pathophysiology courses and in clinical clerkships.

In the sixth edition of this book, as in the previous editions, the important concepts in physiology are covered at the organ system and cellular levels. Chapters 1 and 2 present the underlying principles of cellular physiology and the autonomic nervous system. Chapters 3 through 10 present the major organ systems: neurophysiology and cardiovascular, respiratory, renal, acid-base, gastrointestinal, endocrine, and reproductive physiology. The relationships between organ systems are emphasized to underscore the integrative mechanisms for homeostasis.

This edition includes the following features designed to facilitate the study of physiology:

- **Text** that is easy to read and concise: Clear headings orient the student to the organization and hierarchy of the material. Complex physiologic information is presented systematically, logically, and in a stepwise manner. When a process occurs in a specific sequence, the steps are numbered in the text and often correlate with numbers shown in a companion figure. Bullets are used to separate and highlight the features of a process. Rhetorical questions are posed throughout the text to anticipate the questions that students may be asking; by first contemplating and then answering these questions, students learn to explain difficult concepts and rationalize unexpected or paradoxical findings. Chapter summaries provide a brief overview.
- **Tables and illustrations** that can be used in concert with the text or, because they are designed to stand alone, as a review: The tables summarize, organize, and make comparisons. Examples are (1) a table that compares the gastrointestinal hormones with respect to hormone family, site of and stimuli for secretion, and hormone actions; (2) a table that compares the pathophysiologic features of disorders of Ca^{2+} homeostasis; and (3) a table that compares the features of the action potential in different cardiac tissues. The illustrations are clearly labeled, often with main headings, and include simple diagrams, complex diagrams with numbered steps, and flow charts.
- **Equations and sample problems** that are integrated into the text: All terms and units in equations are defined, and each equation is restated in words to place it in a physiologic context. Sample problems are followed by complete numerical solutions and explanations that guide students through the proper steps in reasoning; by following the steps provided, students acquire the skills and confidence to solve similar or related problems.
- **Clinical physiology** presented in boxes: Each box features a fictitious patient with a classic disorder. The clinical findings and proposed treatment are explained in terms of underlying physiologic principles. An integrative approach to the patient is used to emphasize the relationships between organ systems. For example, the case of type I diabetes mellitus involves a disorder not only of the endocrine system but also of the renal, acid-base, respiratory, and cardiovascular systems.

- **Practice questions** in "Challenge Yourself" sections at the end of each chapter: Practice questions, which are designed for short answers (a word, a phrase, or a numerical solution), challenge the student to apply principles and concepts in problem solving rather than to recall isolated facts. The questions are posed in varying formats and are given in random order. They will be most helpful when used as a tool after studying each chapter and without referring to the text. In that way, the student can confirm his or her understanding of the material and can determine areas of weakness. Answers are provided at the end of the book.
- **Teaching videos on selected topics:** Because students may benefit from oral explanation of complex principles, brief teaching videos on selected topics are included to complement the written text.
- **Abbreviations and normal values** presented in appendices: As students refer to and use these common abbreviations and values throughout the book, they will find that their use becomes second nature.

This book embodies three beliefs that I hold about teaching: (1) even complex information can be transmitted clearly if the presentation is systematic, logical, and stepwise; (2) the presentation can be just as effective in print as in person; and (3) beginning medical students wish for nonreference teaching materials that are accurate and didactically strong but without the details that primarily concern experts. In essence, a book can "teach" if the teacher's voice is present, if the material is carefully selected to include essential information, and if great care is given to logic and sequence. This text offers a down-to-earth and professional presentation written *to* students and *for* students.

I hope that the readers of this book enjoy their study of physiology. Those who learn its principles well will be rewarded throughout their professional careers!

Linda S. Costanzo

Acknowledgments

I gratefully acknowledge the contributions of Elyse O'Grady, Jennifer Ehlers, and Dan Fitzgerald at Elsevier in preparing the sixth edition of *Physiology*. The artist, Matthew Chansky, revised existing figures and created new figures—all of which beautifully complement the text.

Colleagues at Virginia Commonwealth University have faithfully answered my questions, especially Drs. Clive Baumgarten, Diomedes Logothetis, Roland Pittman, and Raphael Witorsch. Sincere thanks also go to the medical students worldwide who have generously written to me about their experiences with earlier editions of the book.

My husband, Richard; our children, Dan and Rebecca; our daughter-in-law, Sheila; and our grandchildren, Elise and Max, have provided enthusiastic support and unqualified love, which give the book its spirit.

CHAPTER 1

Cellular Physiology

Understanding the functions of the organ systems requires profound knowledge of basic cellular mechanisms. Although each organ system differs in its overall function, all are undergirded by a common set of physiologic principles.

The following basic principles of physiology are introduced in this chapter: body fluids, with particular emphasis on the differences in composition of intracellular fluid and extracellular fluid; creation of these concentration differences by transport processes in cell membranes; the origin of the electrical potential difference across cell membranes, particularly in excitable cells such as nerve and muscle; generation of action potentials and their propagation in excitable cells; transmission of information between cells across synapses and the role of neurotransmitters; and the mechanisms that couple the action potentials to contraction in muscle cells.

These principles of cellular physiology constitute a set of recurring and interlocking themes. Once these principles are understood, they can be applied and integrated into the function of each organ system.

VOLUME AND COMPOSITION OF BODY FLUIDS

Distribution of Water in the Body Fluid Compartments

In the human body, water constitutes a high proportion of body weight. The total amount of fluid or water is called **total body water,** which accounts for 50% to 70% of body weight. For example, a 70-kilogram (kg) man whose total body water is 65% of his body weight has 45.5 kg or 45.5 liters (L) of water (1 kg water ≈ 1 L water). In general, total body water correlates inversely with body fat. Thus total body water is a higher percentage of body weight when body fat is low and a lower percentage when body fat is high. Because females have a higher percentage of adipose tissue than males, they tend to have less body water. The distribution of water among body fluid compartments is described briefly in this chapter and in greater detail in Chapter 6.

Total body water is distributed between two major body fluid compartments: intracellular fluid (ICF) and extracellular fluid (ECF) (Fig. 1.1). The **ICF** is contained within the cells and is two-thirds of total body water; the **ECF** is outside the cells and is one-third of total body water. ICF and ECF are separated by the cell membranes.

ECF is further divided into two compartments: plasma and interstitial fluid. **Plasma** is the fluid circulating in the blood vessels and is the smaller of the two ECF

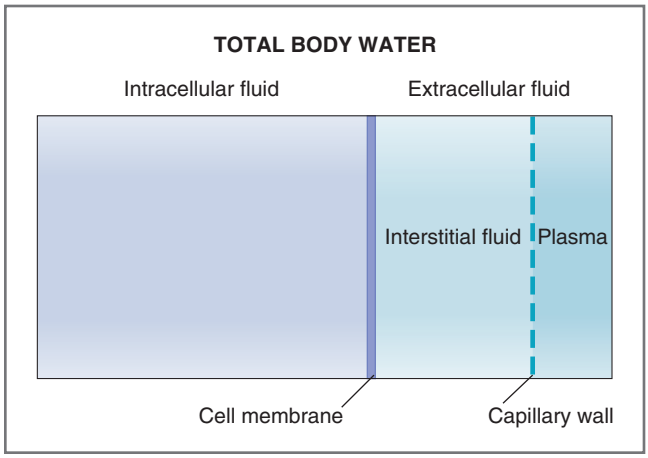


Fig. 1.1 **Body fluid compartments.**

subcompartments. **Interstitial fluid** is the fluid that actually bathes the cells and is the larger of the two subcompartments. Plasma and interstitial fluid are separated by the capillary wall. Interstitial fluid is an **ultrafiltrate** of plasma, formed by filtration processes across the capillary wall. Because the capillary wall is virtually impermeable to large molecules such as plasma proteins, interstitial fluid contains little, if any, protein.

The method for estimating the volume of the body fluid compartments is presented in Chapter 6.

Composition of Body Fluid Compartments

The composition of the body fluids is not uniform. ICF and ECF have vastly different concentrations of various solutes. There are also certain predictable differences in solute concentrations between plasma and interstitial fluid that occur as a result of the exclusion of protein from interstitial fluid.

Units for Measuring Solute Concentrations

Typically, **amounts** of solute are expressed in moles, equivalents, or osmoles. Likewise, **concentrations** of solutes are expressed in moles per liter (mol/L), equivalents per liter (Eq/L), or osmoles per liter (Osm/L). In biologic solutions, concentrations of solutes are usually quite low and are expressed in *milli*moles per liter (mmol/L), *milli*equivalents per liter (mEq/L), or *milli*osmoles per liter (mOsm/L).

One **mole** is 6×10^{23} molecules of a substance. One **millimole** is 1/1000 or 10^{-3} moles. A glucose concentration of 1 mmol/L has 1×10^{-3} moles of glucose in 1 L of solution.

An **equivalent** is used to describe the amount of charged (ionized) solute and is the number of moles of the solute multiplied by its valence. For example, one mole of potassium chloride (KCl) in solution dissociates into one equivalent of potassium (K^+) and one equivalent of chloride (Cl^-). Likewise, one mole of calcium chloride ($CaCl_2$) in solution dissociates into *two* equivalents of calcium (Ca^{2+}) and *two* equivalents of chloride (Cl^-); accordingly, a Ca^{2+} concentration of 1 mmol/L corresponds to 2 mEq/L.

One **osmole** is the number of particles into which a solute dissociates in solution. **Osmolarity** is the concentration of particles in solution expressed as osmoles per liter. If a solute does not dissociate in solution (e.g., glucose), then its osmolarity is equal to its molarity. If a solute dissociates into more than one particle in solution (e.g., NaCl), then its osmolarity equals the molarity multiplied by the number of particles in solution. For example, a solution containing 1 mmol/L NaCl is 2 mOsm/L because NaCl dissociates into two particles.

pH is a logarithmic term that is used to express hydrogen (H^+) concentration. Because the H^+ concentration of body fluids is very low (e.g., 40×10^{-9} Eq/L in arterial blood), it is more conveniently expressed as a logarithmic term, pH. The negative sign means that pH decreases as the concentration of H^+ increases, and pH increases as the concentration of H^+ decreases. Thus

$$pH = -\log_{10}[H^+]$$

SAMPLE PROBLEM. Two men, Subject A and Subject B, have disorders that cause excessive acid production in the body. The laboratory reports the acidity of Subject A's blood in terms of $[H^+]$ and the acidity of Subject B's blood in terms of pH. Subject A has an arterial $[H^+]$ of 65×10^{-9} Eq/L, and Subject B has an arterial pH of 7.3. *Which subject has the higher concentration of H^+ in his blood?*

SOLUTION. To compare the acidity of the blood of each subject, convert the $[H^+]$ for Subject A to pH as follows:

$$\begin{aligned} pH &= -\log_{10}[H^+] \\ &= -\log_{10}(65 \times 10^{-9}\ \text{Eq/L}) \\ &= -\log_{10}(6.5 \times 10^{-8}\ \text{Eq/L}) \\ \log_{10} 6.5 &= 0.81 \\ \log_{10} 10^{-8} &= -8.0 \\ \log_{10} 6.5 \times 10^{-8} &= 0.81 + (-8.0) = -7.19 \\ pH &= -(-7.19) = 7.19 \end{aligned}$$

Thus Subject A has a blood pH of 7.19 computed from the $[H^+]$, and Subject B has a reported blood pH of 7.3. Subject A has a lower blood pH, reflecting a higher $[H^+]$ and a more acidic condition.

Electroneutrality of Body Fluid Compartments

Each body fluid compartment must obey the **principle of macroscopic electroneutrality;** that is, each

compartment must have the same concentration, in mEq/L, of positive charges **(cations)** as of negative charges **(anions).** There can be no more cations than anions, or vice versa. Even when there is a potential difference across the cell membrane, charge balance still is maintained in the bulk (macroscopic) solutions. (Because potential differences are created by the separation of just a few charges adjacent to the membrane, this small separation of charges is not enough to measurably change bulk concentrations.)

Composition of Intracellular Fluid and Extracellular Fluid

The compositions of ICF and ECF are strikingly different, as shown in Table 1.1. The major cation in **ECF** is sodium (Na^+), and the balancing anions are chloride (Cl^-) and bicarbonate (HCO_3^-). The major cations in **ICF** are potassium (K^+) and magnesium (Mg^{2+}), and the balancing anions are proteins and organic phosphates. Other notable differences in composition involve Ca^{2+} and pH. Typically, ICF has a very low concentration of ionized Ca^{2+} ($\approx 10^{-7}$ mol/L), whereas the Ca^{2+} concentration in ECF is higher by approximately four orders of magnitude. ICF is more acidic (has a lower pH) than ECF. Thus substances found in high concentration in ECF are found in low concentration in ICF, and vice versa.

Remarkably, given all of the concentration differences for individual solutes, the total solute concentration **(osmolarity)** is the same in ICF and ECF. This equality is achieved because water flows freely across cell membranes. Any transient differences in osmolarity that occur between ICF and ECF are quickly dissipated by water movement into or out of cells to reestablish the equality.

TABLE 1.1 Approximate Compositions of Extracellular and Intracellular Fluids

Substance and Units	Extracellular Fluid	Intracellular Fluid[a]
Na^+ (mEq/L)	140	14
K^+ (mEq/L)	4	120
Ca^{2+}, ionized (mEq/L)	2.5[b]	1×10^{-4}
Cl^- (mEq/L)	105	10
HCO_3^- (mEq/L)	24	10
pH[c]	7.4	7.1
Osmolarity (mOsm/L)	290	290

[a]The major anions of intracellular fluid are proteins and organic phosphates.

[b]The corresponding total $[Ca^{2+}]$ in extracellular fluid is 5 mEq/L or 10 mg/dL.

[c]pH is $-\log_{10}$ of the $[H^+]$; pH 7.4 corresponds to $[H^+]$ of 40×10^{-9} Eq/L.

Creation of Concentration Differences Across Cell Membranes

The differences in solute concentration across cell membranes are created and maintained by energy-consuming transport mechanisms in the cell membranes.

The best known of these transport mechanisms is the Na^+-K^+ ATPase (Na^+-K^+ pump), which transports Na^+ from ICF to ECF and simultaneously transports K^+ from ECF to ICF. Both Na^+ and K^+ are transported against their respective electrochemical gradients; therefore an energy source, adenosine triphosphate (ATP), is required. The Na^+-K^+ ATPase is responsible for creating the large concentration gradients for Na^+ and K^+ that exist across cell membranes (i.e., the low intracellular Na^+ concentration and the high intracellular K^+ concentration).

Similarly, the intracellular Ca^{2+} concentration is maintained at a level much lower than the extracellular Ca^{2+} concentration. This concentration difference is established, in part, by a cell membrane Ca^{2+} ATPase that pumps Ca^{2+} against its electrochemical gradient. Like the Na^+-K^+ ATPase, the Ca^{2+} ATPase uses ATP as a direct energy source.

In addition to the transporters that use ATP directly, other transporters establish concentration differences across the cell membrane by utilizing the transmembrane Na^+ concentration gradient (established by the Na^+-K^+ ATPase) as an energy source. These transporters create concentration gradients for glucose, amino acids, Ca^{2+}, and H^+ without the direct utilization of ATP.

Clearly, cell membranes have the machinery to establish large concentration gradients. However, if cell membranes were freely permeable to all solutes, these gradients would quickly dissipate. Thus it is critically important that cell membranes are *not* freely permeable to all substances but, rather, have selective permeabilities that maintain the concentration gradients established by energy-consuming transport processes.

Directly or indirectly, the differences in composition between ICF and ECF underlie every important physiologic function, as the following examples illustrate: (1) The resting membrane potential of nerve and muscle critically depends on the difference in concentration of K^+ across the cell membrane; (2) The upstroke of the action potential of these same excitable cells depends on the differences in Na^+ concentration across the cell membrane; (3) Excitation-contraction coupling in muscle cells depends on the differences in Ca^{2+} concentration across the cell membrane and the membrane of the sarcoplasmic reticulum (SR); and (4) Absorption of essential nutrients depends on the transmembrane Na^+ concentration gradient (e.g., glucose absorption in the small intestine or glucose reabsorption in the renal proximal tubule).

Concentration Differences Between Plasma and Interstitial Fluids

As previously discussed, ECF consists of two subcompartments: interstitial fluid and plasma. The most significant difference in composition between these two compartments is the presence of proteins (e.g., albumin) in the plasma compartment. Plasma proteins do not readily cross capillary walls because of their large molecular size and therefore are excluded from interstitial fluid.

The exclusion of proteins from interstitial fluid has secondary consequences. The plasma proteins are negatively charged, and this negative charge causes a redistribution of small, permeant cations and anions across the capillary wall, called a **Gibbs-Donnan equilibrium.** The redistribution can be explained as follows: The plasma compartment contains the impermeant, negatively charged proteins. Because of the requirement for electroneutrality, the plasma compartment must have a slightly lower concentration of small anions (e.g., Cl^-) and a slightly higher concentration of small cations (e.g., Na^+ and K^+) than that of interstitial fluid. The small concentration difference for permeant ions is expressed in the **Gibbs-Donnan ratio,** which gives the plasma concentration relative to the interstitial fluid concentration for anions and interstitial fluid relative to plasma for cations. For example, the Cl^- concentration in plasma is slightly less than the Cl^- concentration in interstitial fluid (due to the effect of the impermeant plasma proteins); the Gibbs-Donnan ratio for Cl^- is 0.95, meaning that $[Cl^-]_{plasma}/[Cl^-]_{interstitial\ fluid}$ equals 0.95. For Na^+, the Gibbs-Donnan ratio is also 0.95, but Na^+, being positively charged, is oriented the opposite way, and $[Na^+]_{interstitial\ fluid}/[Na^+]_{plasma}$ equals 0.95. Generally, these minor differences in concentration for small cations and anions between plasma and interstitial fluid are ignored.

CHARACTERISTICS OF CELL MEMBRANES

Cell membranes are composed primarily of lipids and proteins. The lipid component consists of phospholipids, cholesterol, and glycolipids and is responsible for the high permeability of cell membranes to lipid-soluble substances such as carbon dioxide, oxygen, fatty acids, and steroid hormones. The lipid component of cell membranes is also responsible for the low permeability of cell membranes to water-soluble substances such as ions, glucose, and amino acids. The protein component of the membrane consists of transporters, enzymes, hormone receptors, cell-surface antigens, and ion and water channels.

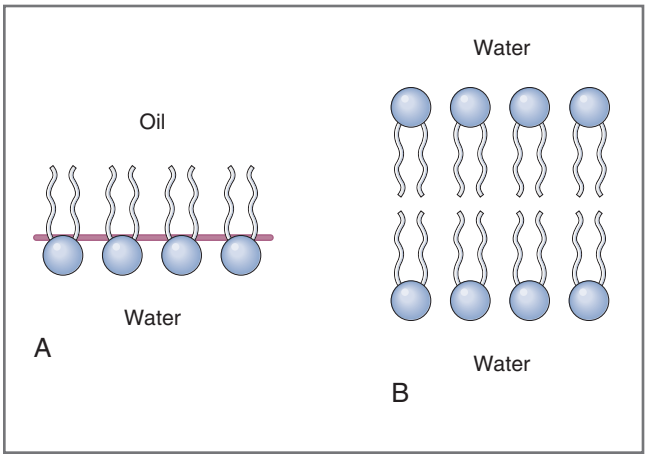


Fig. 1.2 Orientation of phospholipid molecules at oil and water interfaces. Depicted are the orientation of phospholipid at an oil-water interface **(A)** and the orientation of phospholipid in a bilayer, as occurs in the cell membrane **(B).**

Phospholipid Component of Cell Membranes

Phospholipids consist of a phosphorylated glycerol backbone ("head") and two fatty acid "tails" (Fig. 1.2). The glycerol backbone is **hydrophilic** (water soluble), and the fatty acid tails are **hydrophobic** (water insoluble). Thus phospholipid molecules have both hydrophilic and hydrophobic properties and are called **amphipathic.** At an oil-water interface (see Fig. 1.2A), molecules of phospholipids form a monolayer and orient themselves so that the glycerol backbone dissolves in the water phase and the fatty acid tails dissolve in the oil phase. In cell membranes (see Fig. 1.2B), phospholipids orient so that the lipid-soluble fatty acid tails face each other and the water-soluble glycerol heads point away from each other, dissolving in the aqueous solutions of the ICF or ECF. This orientation creates a **lipid bilayer.**

Protein Component of Cell Membranes

Proteins in cell membranes may be either integral or peripheral, depending on whether they span the membrane or whether they are present on only one side. The distribution of proteins in a phospholipid bilayer is illustrated in the **fluid mosaic model,** shown in Figure 1.3.

- **Integral membrane proteins** are embedded in, and anchored to, the cell membrane by **hydrophobic interactions.** To remove an integral protein from the cell membrane, its attachments to the lipid bilayer must be disrupted (e.g., by detergents). Some integral proteins are **transmembrane proteins,** meaning they span the lipid bilayer one or more times; thus

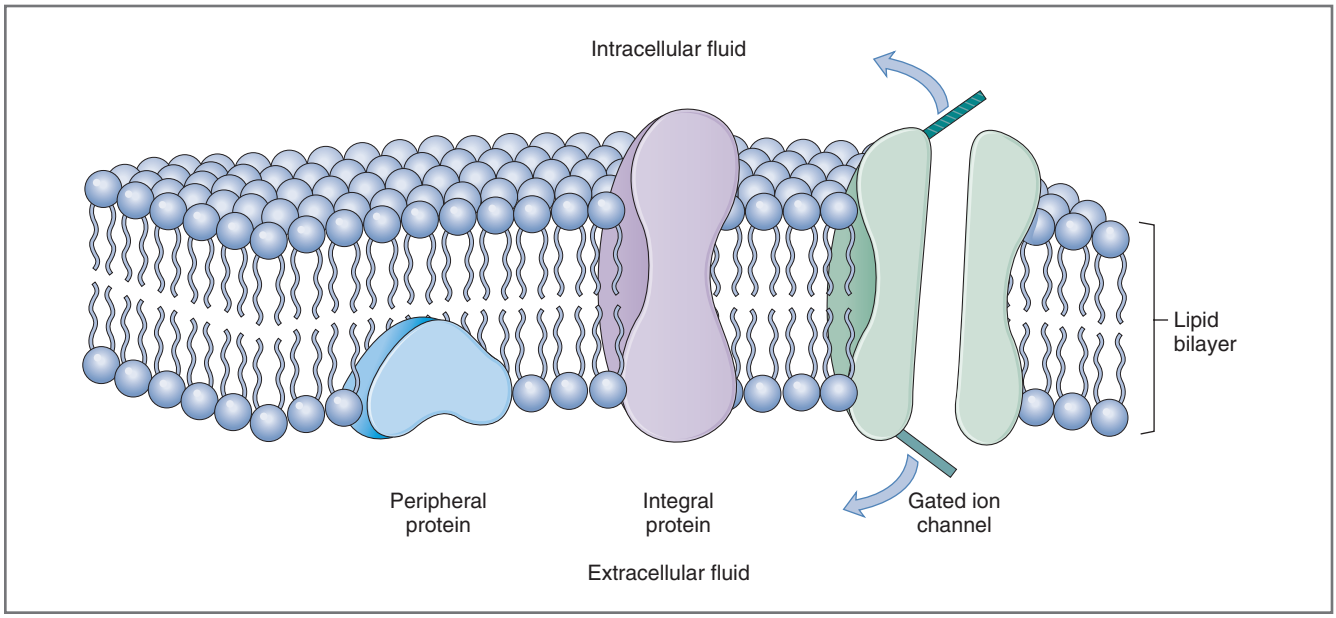


Fig. 1.3 **Fluid mosaic model for cell membranes.**

TABLE 1.2 Summary of Membrane Transport

Type of Transport	Active or Passive	Carrier-Mediated	Uses Metabolic Energy	Dependent on Na^+ Gradient
Simple diffusion	Passive; downhill	No	No	No
Facilitated diffusion	Passive; downhill	Yes	No	No
Primary active transport	Active; uphill	Yes	Yes; direct	No
Cotransport	Secondary active[a]	Yes	Yes; indirect	Yes (solutes move in same direction as Na^+ across cell membrane)
Countertransport	Secondary active[a]	Yes	Yes; indirect	Yes (solutes move in opposite direction as Na^+ across cell membrane)

[a]Na^+ is transported downhill, and one or more solutes are transported uphill.

transmembrane proteins are in contact with both ECF and ICF. Examples of transmembrane integral proteins are ligand-binding receptors (e.g., for hormones or neurotransmitters), transport proteins (e.g., Na^+-K^+ ATPase), pores, ion channels, cell adhesion molecules, and GTP-binding proteins (G proteins). A second category of integral proteins is embedded in the lipid bilayer of the membrane but does not span it. A third category of integral proteins is associated with membrane proteins but is not embedded in the lipid bilayer.

- **Peripheral membrane proteins** are *not* embedded in the membrane and are *not* covalently bound to cell membrane components. They are loosely attached to either the intracellular or extracellular side of the cell membrane by **electrostatic interactions** (e.g., with integral proteins) and can be removed with mild treatments that disrupt ionic or hydrogen bonds. One example of a peripheral membrane protein is **ankyrin,** which "anchors" the cytoskeleton of red blood cells to an integral membrane transport protein, the Cl^--HCO_3^- exchanger (also called band 3 protein).

TRANSPORT ACROSS CELL MEMBRANES

Several types of mechanisms are responsible for transport of substances across cell membranes (Table 1.2).

Substances may be transported down an electrochemical gradient (downhill) or against an electrochemical gradient (uphill). **Downhill** transport occurs by diffusion, either simple or facilitated, and requires no input of metabolic energy. **Uphill** transport occurs by active transport, which may be primary or secondary. Primary and secondary active transport processes

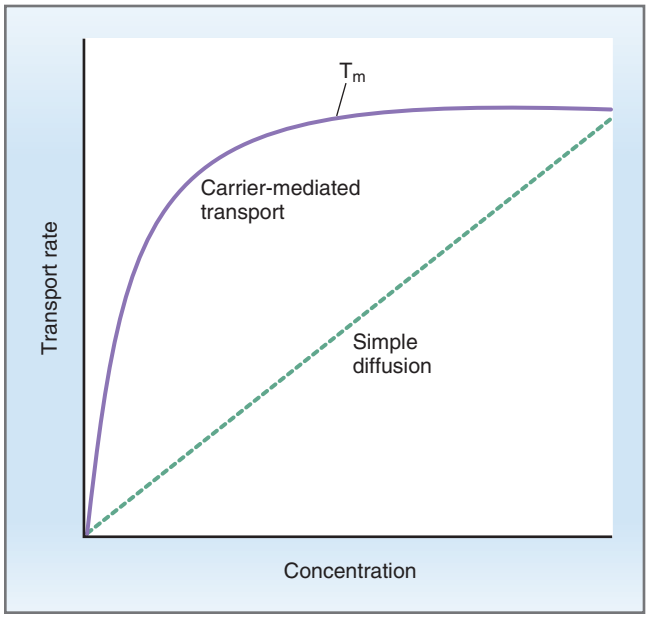


Fig. 1.4 Kinetics of carrier-mediated transport. T_m, Transport maximum.

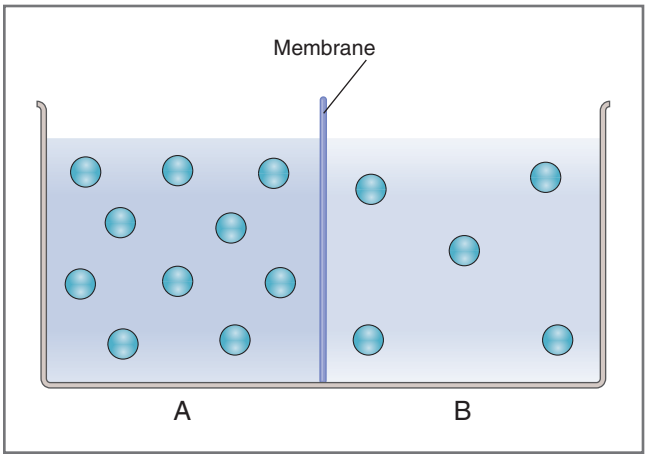


Fig. 1.5 Simple diffusion. The two solutions, **A** and **B,** are separated by a membrane, which is permeable to the solute *(circles).* Solution A initially contains a higher concentration of the solute than does Solution B.

are distinguished by their energy source. Primary active transport requires a *direct* input of metabolic energy; secondary active transport utilizes an *indirect* input of metabolic energy.

Further distinctions among transport mechanisms are based on whether the process involves a protein carrier. Simple diffusion is the only form of transport that is *not* carrier mediated. Facilitated diffusion, primary active transport, and secondary active transport all involve integral membrane proteins and are called **carrier-mediated transport.** All forms of carrier-mediated transport share the following three features: saturation, stereospecificity, and competition.

- **Saturation.** Saturability is based on the concept that carrier proteins have a limited number of binding sites for the solute. Figure 1.4 shows the relationship between the rate of carrier-mediated transport and solute concentration. At low solute concentrations, many binding sites are available and the rate of transport increases steeply as the concentration increases. However, at high solute concentrations, the available binding sites become scarce and the rate of transport levels off. Finally, when all of the binding sites are occupied, saturation is achieved at a point called the **transport maximum,** or $\mathbf{T_m}$**.** The kinetics of carrier-mediated transport are similar to Michaelis-Menten enzyme kinetics—both involve proteins with a limited number of binding sites. (The T_m is analogous to the V_{max} of enzyme kinetics.) T_m-limited glucose transport in the proximal tubule of the kidney is an example of saturable transport.
- **Stereospecificity.** The binding sites for solute on the transport proteins are stereospecific. For example, the transporter for glucose in the renal proximal tubule recognizes and transports the natural isomer D-glucose, but it does not recognize or transport the unnatural isomer L-glucose. In contrast, simple diffusion does not distinguish between the two glucose isomers because no protein carrier is involved.
- **Competition.** Although the binding sites for transported solutes are quite specific, they may recognize, bind, and even transport chemically related solutes. For example, the transporter for glucose is specific for D-glucose, but it also recognizes and transports a closely related sugar, D-galactose. Therefore the presence of D-galactose inhibits the transport of D-glucose by occupying some of the binding sites and making them unavailable for glucose.

Simple Diffusion

Diffusion of Nonelectrolytes

Simple diffusion occurs as a result of the random thermal motion of molecules, as shown in Figure 1.5. Two solutions, A and B, are separated by a membrane that is permeable to the solute. The solute concentration in A is initially twice that of B. The solute molecules are in constant motion, with equal probability that a given molecule will cross the membrane to the other solution. However, because there are twice as many solute molecules in Solution A as in Solution B, there will be greater movement of molecules from A to B than from B to A. In other words, there will be **net diffusion** of the solute from A to B, which will continue until the solute concentrations of the two solutions become equal (although the random movement of molecules will go on forever).

Net diffusion of the solute is called **flux,** or **flow (J),** and depends on the following variables: size of the concentration gradient, partition coefficient, diffusion coefficient, thickness of the membrane, and surface area available for diffusion.

CONCENTRATION GRADIENT ($C_A - C_B$)

The concentration gradient across the membrane is the driving force for net diffusion. The larger the difference in solute concentration between Solution A and Solution B, the greater the driving force and the greater the net diffusion. It also follows that, if the concentrations in the two solutions are equal, there is no driving force and no net diffusion.

PARTITION COEFFICIENT (K)

The partition coefficient, by definition, describes the solubility of a solute in oil relative to its solubility in water. The greater the relative solubility in oil, the higher the partition coefficient and the more easily the solute can dissolve in the cell membrane's lipid bilayer. Nonpolar solutes tend to be soluble in oil and have high values for partition coefficient, whereas polar solutes tend to be insoluble in oil and have low values for partition coefficient. The partition coefficient can be measured by adding the solute to a mixture of olive oil and water and then measuring its concentration in the oil phase relative to its concentration in the water phase. Thus

$$K = \frac{\text{Concentration in olive oil}}{\text{Concentration in water}}$$

DIFFUSION COEFFICIENT (D)

The diffusion coefficient depends on such characteristics as size of the solute molecule and the viscosity of the medium. It is defined by the Stokes-Einstein equation (see later). The diffusion coefficient correlates *inversely* with the molecular radius of the solute and the viscosity of the medium. Thus small solutes in nonviscous solutions have the largest diffusion coefficients and diffuse most readily; large solutes in viscous solutions have the smallest diffusion coefficients and diffuse least readily. Thus

$$D = \frac{KT}{6\pi r\eta}$$

where

D = Diffusion coefficient
K = Boltzmann constant
T = Absolute temperature (K)
r = Molecular radius
η = Viscosity of the medium

THICKNESS OF THE MEMBRANE (ΔX)

The thicker the cell membrane, the greater the distance the solute must diffuse and the lower the rate of diffusion.

SURFACE AREA (A)

The greater the surface area of membrane available, the higher the rate of diffusion. For example, lipid-soluble gases such as oxygen and carbon dioxide have particularly high rates of diffusion across cell membranes. These high rates can be attributed to the large surface area for diffusion provided by the lipid component of the membrane.

To simplify the description of diffusion, several of the previously cited characteristics can be combined into a single term called **permeability (P).** Permeability includes the partition coefficient, the diffusion coefficient, and the membrane thickness. Thus

$$P = \frac{KD}{\Delta x}$$

By combining several variables into permeability, the rate of net diffusion is simplified to the following expression:

$$J = PA(C_A - C_B)$$

where

J = Net rate of diffusion (mmol/s)
P = Permeability (cm/s)
A = Surface area for diffusion (cm^2)
C_A = Concentration in Solution A (mmol/L)
C_B = Concentration in Solution B (mmol/L)

SAMPLE PROBLEM. Solution A and Solution B are separated by a membrane whose permeability to urea is 2×10^{-5} cm/s and whose surface area is 1 cm^2. The concentration of urea in A is 10 mg/mL, and the concentration of urea in B is 1 mg/mL. The partition coefficient for urea is 10^{-3}, as measured in an oil-water mixture. *What are the initial rate and direction of net diffusion of urea?*

SOLUTION. Note that the partition coefficient is extraneous information because the value for permeability, which already includes the partition coefficient, is given. Net flux can be calculated by substituting the following values in the equation for net diffusion: Assume that 1 mL of water = 1 cm^3. Thus

$$J = PA(C_A - C_B)$$

where

$$J = 2 \times 10^{-5} \text{ cm/s} \times 1 \text{ cm}^2 \times (10 \text{ mg/mL} - 1 \text{ mg/mL})$$

$$J = 2 \times 10^{-5} \text{ cm/s} \times 1 \text{ cm}^2 \times (10 \text{ mg/cm}^3 - 1 \text{ mg/cm}^3)$$

$$= 1.8 \times 10^{-4} \text{ mg/s}$$

The *magnitude* of net flux has been calculated as 1.8×10^{-4} mg/s. The *direction* of net flux can be determined intuitively because net flux will occur from the area of high concentration (Solution A) to the area of low concentration (Solution B). Net diffusion will continue until the urea concentrations of the two solutions become equal, at which point the driving force will be zero.

Diffusion of Electrolytes

Thus far, the discussion concerning diffusion has assumed that the solute is a nonelectrolyte (i.e., it is uncharged). However, if the diffusing solute is an **ion** or an **electrolyte,** there are two additional consequences of the presence of charge on the solute.

First, if there is a potential difference across the membrane, that potential difference will alter the net rate of diffusion of a charged solute. (A potential difference does not alter the rate of diffusion of a nonelectrolyte.) For example, the diffusion of K^+ ions will be slowed if K^+ is diffusing into an area of positive charge, and it will be accelerated if K^+ is diffusing into an area of negative charge. This effect of potential difference can either add to or negate the effects of differences in concentrations, depending on the orientation of the potential difference and the charge on the diffusing ion. If the concentration gradient and the charge effect are oriented in the same direction across the membrane, they will combine; if they are oriented in opposite directions, they may cancel each other out.

Second, when a charged solute diffuses down a concentration gradient, that diffusion can *itself* generate a potential difference across a membrane called a **diffusion potential.** The concept of diffusion potential will be discussed more fully in a following section.

Facilitated Diffusion

Like simple diffusion, facilitated diffusion occurs down an electrochemical potential gradient; thus it requires no input of metabolic energy. Unlike simple diffusion, however, facilitated diffusion uses a membrane carrier and exhibits all the characteristics of carrier-mediated transport: saturation, stereospecificity, and competition. At low solute concentration, facilitated diffusion typically proceeds faster than simple diffusion (i.e., is facilitated) because of the function of the carrier. However, at higher concentrations, the carriers will become saturated and facilitated diffusion will level off. (In contrast, simple diffusion will proceed as long as there is a concentration gradient for the solute.)

An excellent example of facilitated diffusion is the transport of **D-glucose** into skeletal muscle and adipose cells by the **GLUT4** transporter. Glucose transport can proceed as long as the blood concentration of glucose is higher than the intracellular concentration of glucose and as long as the carriers are not saturated. Other monosaccharides such as D-galactose, 3-*O*-methyl glucose, and phlorizin competitively inhibit the transport of glucose because they bind to transport sites on the carrier. The competitive solute may itself be transported (e.g., D-galactose), or it may simply occupy the binding sites and prevent the attachment of glucose (e.g., phlorizin). As noted previously, the nonphysiologic stereoisomer, L-glucose, is not recognized by the carrier for facilitated diffusion and therefore is not bound or transported.

Primary Active Transport

In active transport, one or more solutes are moved against an electrochemical potential gradient (uphill). In other words, solute is moved from an area of low concentration (or low electrochemical potential) to an area of high concentration (or high electrochemical potential). Because movement of a solute *uphill* is work, metabolic energy in the form of ATP must be provided. In the process, ATP is hydrolyzed to adenosine diphosphate (ADP) and inorganic phosphate (P_i), releasing energy from the terminal high-energy phosphate bond of ATP. When the terminal phosphate is released, it is transferred to the transport protein, initiating a cycle of phosphorylation and dephosphorylation. When the ATP energy source is directly coupled to the transport process, it is called *primary* active transport. Three examples of primary active transport in physiologic systems are the Na^+-K^+ ATPase present in all cell membranes, the Ca^{2+} ATPase present in SR and endoplasmic reticulum, and the H^+-K^+ ATPase present in gastric parietal cells and renal α-intercalated cells.

Na^+-K^+ ATPase (Na^+-K^+ Pump)

Na^+-K^+ ATPase is present in the membranes of all cells. It pumps Na^+ from ICF to ECF and K^+ from ECF to ICF (Fig. 1.6). Each ion moves against its respective electrochemical gradient. The stoichiometry can vary but, typically, for every three Na^+ ions pumped out of the cell, two K^+ ions are pumped into the cell. This stoichiometry of three Na^+ ions per two K^+ ions means that, for each cycle of the Na^+-K^+ ATPase, more positive charge is pumped out of the cell than is pumped into the cell. Thus the transport process is termed **electrogenic** because it creates a charge separation and a potential difference. The Na^+-K^+ ATPase is responsible

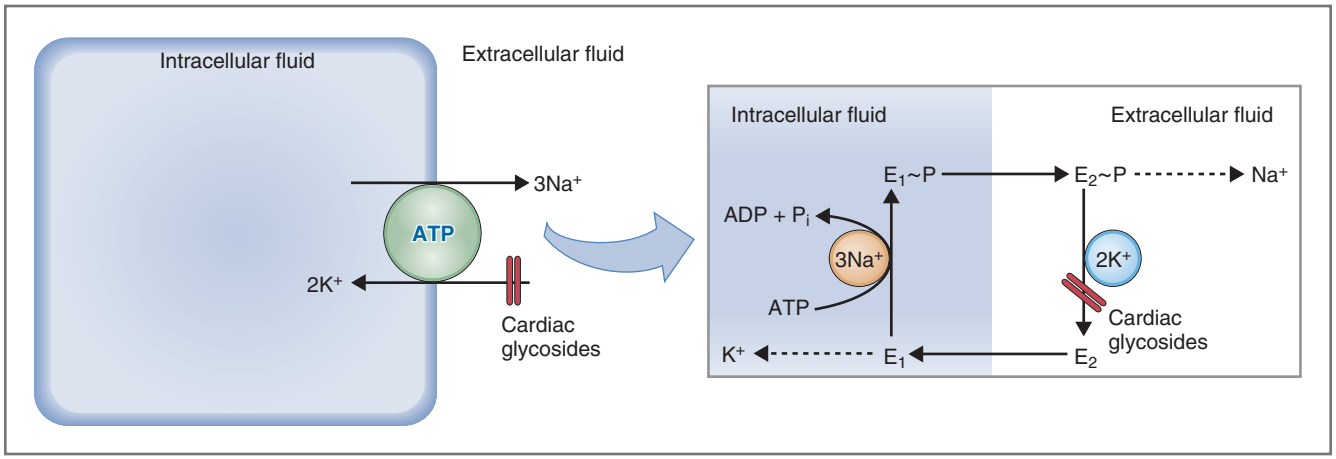


Fig. 1.6 Na^+-K^+ pump of cell membranes. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; *E,* Na^+-K^+ ATPase; *E~P,* phosphorylated Na^+-K^+ ATPase; P_i, inorganic phosphate.

for maintaining concentration gradients for both Na^+ and K^+ across cell membranes, keeping the intracellular Na^+ concentration low and the intracellular K^+ concentration high.

The Na^+-K^+ ATPase consists of α and β subunits. The α subunit contains the ATPase activity, as well as the binding sites for the transported ions, Na^+ and K^+. The Na^+-K^+ ATPase switches between two major conformational states, E_1 and E_2. In the **E_1 state,** the binding sites for Na^+ and K^+ face the ICF and the enzyme has a high affinity for Na^+. In the **E_2 state,** the binding sites for Na^+ and K^+ face the ECF and the enzyme has a high affinity for K^+. The enzyme's ion-transporting function (i.e., pumping Na^+ out of the cell and K^+ into the cell) is based on cycling between the E_1 and E_2 states and is powered by ATP hydrolysis.

The **transport cycle** is illustrated in Figure 1.6. The cycle begins with the enzyme in the E_1 state, bound to ATP. In the E_1 state, the ion-binding sites face the ICF, and the enzyme has a high affinity for Na^+; three Na^+ ions bind, ATP is hydrolyzed, and the terminal phosphate of ATP is transferred to the enzyme, producing a high-energy state, $E_1 \sim P$. Now, a major conformational change occurs, and the enzyme switches from $E_1 \sim P$ to $E_2 \sim P$. In the E_2 state, the ion-binding sites face the ECF, the affinity for Na^+ is low, and the affinity for K^+ is high. The three Na^+ ions are released from the enzyme to ECF, two K^+ ions are bound, and inorganic phosphate is released from E_2. The enzyme now binds intracellular ATP, and another major conformational change occurs that returns the enzyme to the E_1 state; the two K^+ ions are released to ICF, and the enzyme is ready for another cycle.

Cardiac glycosides (e.g., **ouabain** and **digitalis**) are a class of drugs that inhibits Na^+-K^+ ATPase. Treatment with this class of drugs causes certain predictable changes in intracellular ionic concentration: The intracellular Na^+ concentration will increase, and the intracellular K^+ concentration will decrease. Cardiac glycosides inhibit the Na^+-K^+ ATPase by binding to the $E_2 \sim P$ form near the K^+-binding site on the extracellular side, thereby preventing the conversion of $E_2 \sim P$ back to E_1. By disrupting the cycle of phosphorylation-dephosphorylation, these drugs disrupt the entire enzyme cycle and its transport functions.

Ca^{2+} ATPase (Ca^{2+} Pump)

Most **cell (plasma) membranes** contain a Ca^{2+} ATPase, or plasma-membrane Ca^{2+} ATPase **(PMCA),** whose function is to extrude Ca^{2+} from the cell against an electrochemical gradient; one Ca^{2+} ion is extruded for each ATP hydrolyzed. PMCA is responsible, in part, for maintaining the very low intracellular Ca^{2+} concentration. In addition, the **sarcoplasmic reticulum (SR)** of muscle cells and the **endoplasmic reticulum** of other cells contain variants of Ca^{2+} ATPase that pump two Ca^{2+} ions (for each ATP hydrolyzed) from ICF into the interior of the SR or endoplasmic reticulum (i.e., Ca^{2+} sequestration). These variants are called SR and endoplasmic reticulum Ca^{2+} ATPase **(SERCA).** Ca^{2+} ATPase functions similarly to Na^+-K^+ ATPase, with E_1 and E_2 states that have, respectively, high and low affinities for Ca^{2+}. For PMCA, the E_1 state binds Ca^{2+} on the intracellular side, a conformational change to the E_2 state occurs, and the E_2 state releases Ca^{2+} to ECF. For SERCA, the E_1 state binds Ca^{2+} on the intracellular side and the E_2 state releases Ca^{2+} to the lumen of the SR or endoplasmic reticulum.

H^+-K^+ ATPase (H^+-K^+ Pump)

H^+-K^+ ATPase is found in the parietal cells of the gastric mucosa and in the α-intercalated cells of the renal collecting duct. In the stomach, it pumps H^+ from the ICF of the parietal cells into the lumen of the stomach, where it acidifies the gastric contents. **Omeprazole,** an inhibitor of gastric H^+-K^+ ATPase, can be used therapeutically to reduce the secretion of H^+ in the treatment of some types of peptic ulcer disease.

Secondary Active Transport

Secondary active transport processes are those in which the transport of two or more solutes is coupled. One of the solutes, usually Na^+, moves down its electrochemical gradient (downhill), and the other solute moves against its electrochemical gradient (uphill). The downhill movement of Na^+ provides energy for the uphill movement of the other solute. Thus metabolic energy, as ATP, is not used directly, but it is supplied indirectly in the Na^+ concentration gradient across the cell membrane. (The Na^+-K^+ ATPase, utilizing ATP, creates and maintains this Na^+ gradient.) The name *secondary* active transport therefore refers to the *indirect* utilization of ATP as an energy source.

Inhibition of the Na^+-K^+ ATPase (e.g., by treatment with ouabain) diminishes the transport of Na^+ from ICF to ECF, causing the intracellular Na^+ concentration to increase and thereby decreasing the size of the transmembrane Na^+ gradient. Thus indirectly, all secondary active transport processes are diminished by inhibitors of the Na^+-K^+ ATPase because their energy source, the Na^+ gradient, is diminished.

There are two types of secondary active transport, distinguishable by the direction of movement of the uphill solute. If the uphill solute moves in the same direction as Na^+, it is called **cotransport,** or **symport.** If the uphill solute moves in the opposite direction of Na^+, it is called **countertransport, antiport,** or **exchange.**

Cotransport

Cotransport (symport) is a form of secondary active transport in which all solutes are transported in the **same direction** across the cell membrane. Na^+ moves *into* the cell on the carrier down its electrochemical gradient; the solutes, cotransported with Na^+, also move *into* the cell. Cotransport is involved in several critical physiologic processes, particularly in the absorbing epithelia of the small intestine and the renal tubule. For example, **Na^+-glucose cotransport** (SGLT) and **Na^+-amino acid cotransport** are present in the luminal membranes of the epithelial cells of both small intestine and renal proximal tubule. Another example of cotransport involving the renal tubule is **Na^+-K^+-$2Cl^-$ cotransport,** which is present in the luminal membrane of epithelial cells of the thick ascending limb. In each example, the Na^+ gradient established by the Na^+-K^+ ATPase is used to transport solutes such as glucose, amino acids, K^+, or Cl^- against electrochemical gradients.

Figure 1.7 illustrates the principles of cotransport using the example of Na^+-glucose cotransport (**SGLT1,** or Na^+-glucose transport protein 1) in intestinal epithelial cells. The cotransporter is present in the luminal membrane of these cells and can be visualized as having

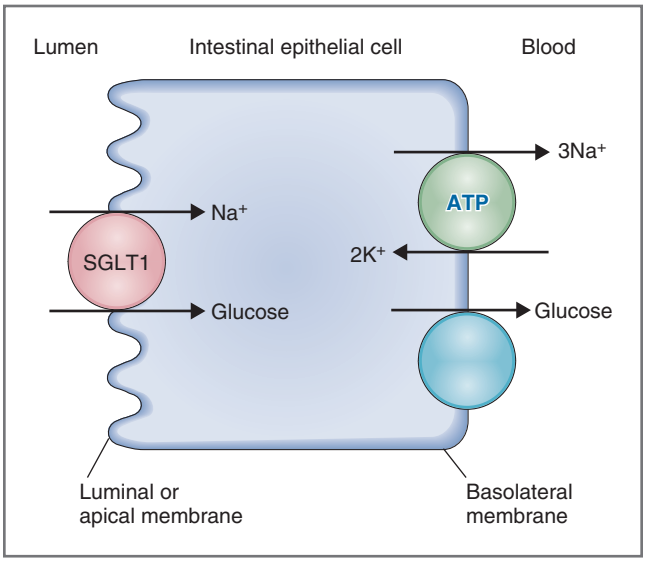


Fig. 1.7 Na^+-glucose cotransport in an intestinal epithelial cell. *ATP,* Adenosine triphosphate; *SGLT1,* Na^+-glucose transport protein 1.

two specific recognition sites, one for Na^+ ions and the other for glucose. When both Na^+ and glucose are present in the lumen of the small intestine, they bind to the transporter. In this configuration, the cotransport protein rotates and releases both Na^+ and glucose to the interior of the cell. (Subsequently, both solutes are transported out of the cell across the basolateral membrane—Na^+ by the Na^+-K^+ ATPase and glucose by facilitated diffusion.) If either Na^+ or glucose is missing from the intestinal lumen, the cotransporter cannot rotate. Thus both solutes are required, and neither can be transported in the absence of the other (Box 1.1).

Finally, the role of the intestinal Na^+-glucose cotransport process can be understood in the context of overall intestinal absorption of carbohydrates. Dietary carbohydrates are digested by gastrointestinal enzymes to an absorbable form, the monosaccharides. One of these monosaccharides is glucose, which is absorbed across the intestinal epithelial cells by a combination of Na^+-glucose cotransport in the luminal membrane and facilitated diffusion of glucose in the basolateral membrane. Na^+-glucose cotransport is the active step, allowing glucose to be absorbed into the blood against an electrochemical gradient.

Countertransport

Countertransport (antiport or exchange) is a form of secondary active transport in which solutes move in *opposite directions* across the cell membrane. Na^+ moves *into* the cell on the carrier down its electrochemical gradient; the solutes that are countertransported or exchanged for Na^+ move *out of* the cell. Countertransport is illustrated by Ca^{2+}-Na^+ exchange (Fig. 1.8) and by Na^+-H^+ exchange. As with cotransport, each process

BOX 1.1 Clinical Physiology: Glucosuria Due to Diabetes Mellitus

DESCRIPTION OF CASE. At his annual physical examination, a 14-year-old boy reports symptoms of frequent urination and severe thirst. A dipstick test of his urine shows elevated levels of glucose. The physician orders a glucose tolerance test, which indicates that the boy has type I diabetes mellitus. He is treated with insulin by injection, and his dipstick test is subsequently normal.

EXPLANATION OF CASE. Although type I diabetes mellitus is a complex disease, this discussion is limited to the symptom of frequent urination and the finding of glucosuria (glucose in the urine). Glucose is normally handled by the kidney in the following manner: Glucose in the blood is filtered across the glomerular capillaries. The epithelial cells, which line the renal proximal tubule, then reabsorb all of the filtered glucose so that no glucose is excreted in the urine. Thus a normal dipstick test would show no glucose in the urine. If the epithelial cells in the proximal tubule do not reabsorb all of the filtered glucose back into the blood, the glucose that escapes reabsorption is excreted. The cellular mechanism for this glucose reabsorption is the Na^+-glucose cotransporter in the luminal membrane of the proximal tubule cells. Because this is a carrier-mediated transporter, there is a finite number of binding sites for glucose. Once these binding sites are fully occupied, saturation of transport occurs (transport maximum).

In this patient with type I diabetes mellitus, the hormone insulin is not produced in sufficient amounts by the pancreatic β cells. Insulin is required for normal uptake of glucose into liver, muscle, and other cells. Without insulin, the blood glucose concentration increases because glucose is not taken up by the cells. When the blood glucose concentration increases to high levels, more glucose is filtered by the renal glomeruli and the amount of glucose filtered exceeds the capacity of the Na^+-glucose cotransporter. The glucose that cannot be reabsorbed because of saturation of this transporter is then "spilled" in the urine.

TREATMENT. Treatment of the patient with type I diabetes mellitus consists of administering exogenous insulin by injection. Whether secreted normally from the pancreatic β cells or administered by injection, insulin lowers the blood glucose concentration by promoting glucose uptake into cells. When this patient received insulin, his blood glucose concentration was reduced; thus the amount of glucose filtered was reduced, and the Na^+-glucose cotransporters were no longer saturated. All of the filtered glucose could be reabsorbed, and therefore no glucose was excreted, or "spilled," in the urine.

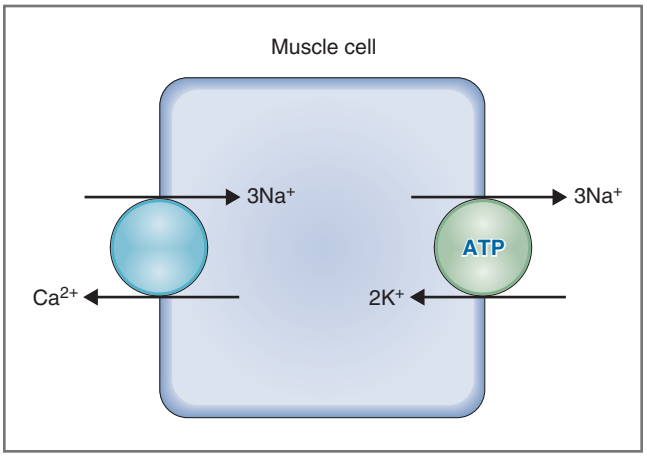


Fig. 1.8 Ca^{2+}-Na^+ countertransport (exchange) in a muscle cell. *ATP,* Adenosine triphosphate.

uses the Na^+ gradient established by the Na^+-K^+ ATPase as an energy source; Na^+ moves downhill and Ca^{2+} or H^+ moves uphill.

Ca^{2+}-Na^+ exchange is one of the transport mechanisms, along with the Ca^{2+} ATPase, that helps maintain the intracellular Ca^{2+} concentration at very low levels ($\approx 10^{-7}$ molar). To accomplish Ca^{2+}-Na^+ exchange, active transport must be involved because Ca^{2+} moves out of the cell against its electrochemical gradient. Figure 1.8 illustrates the concept of Ca^{2+}-Na^+ exchange in a muscle cell membrane. The exchange protein has recognition sites for both Ca^{2+} and Na^+. The protein must bind Ca^{2+} on the intracellular side of the membrane and, simultaneously, bind Na^+ on the extracellular side. In this configuration, the exchange protein rotates and delivers Ca^{2+} to the exterior of the cell and Na^+ to the interior of the cell.

The stoichiometry of Ca^{2+}-Na^+ exchange varies between different cell types and may even vary for a single cell type under different conditions. Usually, however, three Na^+ ions enter the cell for each Ca^{2+} ion extruded from the cell. With this stoichiometry of three Na^+ ions per one Ca^{2+} ion, three positive charges move into the cell in exchange for two positive charges leaving the cell, making the Ca^{2+}-Na^+ exchanger **electrogenic.**

Osmosis

Osmosis is the flow of water across a semipermeable membrane because of differences in solute concentration. Concentration differences of impermeant solutes establish osmotic pressure differences, and this osmotic pressure difference causes water to flow by osmosis. Osmosis of water is *not* diffusion of water: Osmosis occurs because of a pressure difference, whereas diffusion occurs because of a concentration (or activity) difference of water.

Osmolarity

The osmolarity of a solution is its concentration of osmotically active particles, expressed as osmoles per liter or milliosmoles per liter. To calculate osmolarity, it is necessary to know the concentration of solute and whether the solute dissociates in solution. For example, glucose does not dissociate in solution; theoretically, NaCl dissociates into two particles and $CaCl_2$ dissociates into three particles. The symbol "g" gives the number of particles in solution and also takes into account whether there is complete or only partial dissociation. Thus if NaCl is completely dissociated into two particles, g equals 2.0; if NaCl dissociates only partially, then g falls between 1.0 and 2.0. Osmolarity is calculated as follows:

$$\text{Osmolarity} = g\ C$$

where

Osmolarity = Concentration of particles (mOsm/L)
g = Number of particles per mole in solution (Osm/mol)
C = Concentration (mmol/L)

If two solutions have the same calculated osmolarity, they are called **isosmotic.** If two solutions have different calculated osmolarities, the solution with the higher osmolarity is called **hyperosmotic** and the solution with the lower osmolarity is called **hyposmotic.**

Osmolality

Osmolality is similar to osmolarity, except that it is the concentration of osmotically active particles, expressed as osmoles (or milliosmoles) *per kilogram of water.* Because 1 kg of water is approximately equivalent to 1 L of water, osmolarity and osmolality will have essentially the same numerical value.

SAMPLE PROBLEM. Solution A is 2 mmol/L urea, and Solution B is 1 mmol/L NaCl. Assume that g_{NaCl} = 1.85. *Are the two solutions isosmotic?*

SOLUTION. Calculate the osmolarities of both solutions to compare them. Solution A contains urea, which does not dissociate in solution. Solution B contains NaCl, which dissociates partially in solution but not completely (i.e., g < 2.0). Thus

$$\text{Osmolarity}_A = 1\ \text{Osm/mol} \times 2\ \text{mmol/L}$$
$$= 2\ \text{mOsm/L}$$
$$\text{Osmolarity}_B = 1.85\ \text{Osm/mol} \times 1\ \text{mmol/L}$$
$$= 1.85\ \text{mOsm/L}$$

The two solutions do not have the same calculated osmolarity; therefore they are *not isosmotic.* Solution A has a higher osmolarity than Solution B and is hyperosmotic; Solution B is hyposmotic.

Osmotic Pressure

Osmosis is the flow of water across a semipermeable membrane due to a difference in solute concentration. The difference in solute concentration creates an osmotic pressure difference across the membrane and that pressure difference is the driving force for osmotic water flow.

Figure 1.9 illustrates the concept of osmosis. Two aqueous solutions, open to the atmosphere, are shown in Figure 1.9A. The membrane separating the solutions is permeable to water but is impermeable to the solute. Initially, solute is present only in Solution 1. The solute in Solution 1 produces an osmotic pressure and causes, by the interaction of solute with pores in the membrane, a reduction in hydrostatic pressure of Solution 1. The resulting hydrostatic pressure difference across the membrane then causes water to flow from Solution 2 into Solution 1. With time, water flow causes the volume of Solution 1 to increase and the volume of Solution 2 to decrease.

Figure 1.9B shows a similar pair of solutions; however, the preparation has been modified so that water flow into Solution 1 is prevented by applying pressure to a piston. *The pressure required to stop the flow of water is the osmotic pressure of Solution 1.*

The osmotic pressure (π) of Solution 1 depends on two factors: the concentration of osmotically active particles and whether the solute remains in Solution 1 (i.e., whether the solute can cross the membrane or not). Osmotic pressure is calculated by the **van't Hoff equation** (as follows), which converts the concentration of particles to a pressure, taking into account whether the solute is retained in the original solution.

Thus

$$\pi = g\ C\ \sigma\ R\ T$$

where

π = Osmotic pressure (atm or mm Hg)
g = Number of particles per mole in solution (Osm/mol)
C = Concentration (mmol/L)
σ = Reflection coefficient (varies from 0 to 1)
R = Gas constant (0.082 L – atm/mol – K)
T = Absolute temperature (K)

The **reflection coefficient (σ)** is a dimensionless number ranging between 0 and 1 that describes the

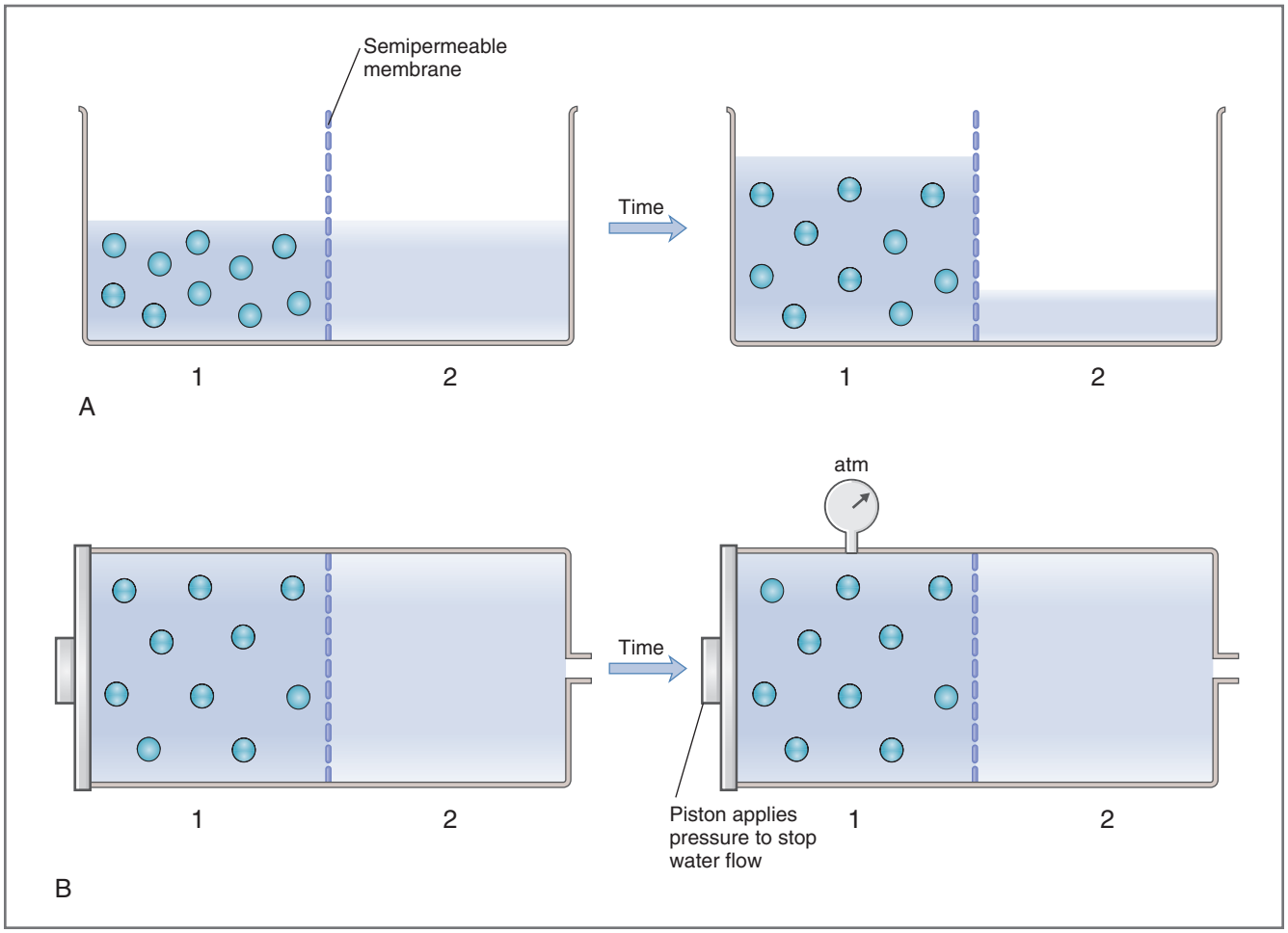


Fig. 1.9 Osmosis across a semipermeable membrane. A, Solute *(circles)* is present on one side of a semipermeable membrane; with time, the osmotic pressure created by the solute causes water to flow from Solution 2 to Solution 1. The resulting volume changes are shown. **B,** The solutions are closed to the atmosphere, and a piston is applied to stop the flow of water into Solution 1. The pressure needed to stop the flow of water is the effective osmotic pressure of Solution 1. *atm,* Atmosphere.

ease with which a solute crosses a membrane. Reflection coefficients can be described for the following three conditions (Fig. 1.10):

- **σ = 1.0** (see Fig. 1.10A). If the membrane is impermeable to the solute, σ is 1.0, and the solute will be retained in the original solution and exert its full osmotic effect. In this case, the effective osmotic pressure will be maximal and will cause maximal water flow. For example, **serum albumin** and **intracellular proteins** are solutes where σ = 1.
- **σ = 0** (see Fig. 1.10C). If the membrane is freely permeable to the solute, σ is 0, and the solute will diffuse across the membrane down its concentration gradient until the solute concentrations of the two solutions are equal. In other words, the solute behaves as if it were water. In this case, there will be *no* effective osmotic pressure difference across the membrane and therefore no driving force for osmotic water flow. Refer again to the van't Hoff equation and notice that, when σ = 0, the calculated effective osmotic pressure becomes zero. **Urea** is an example of a solute where σ = 0 (or nearly 0).
- **σ = a value between 0 and 1** (see Fig. 1.10B). Most solutes are neither impermeant (σ = 1) nor freely permeant (σ = 0) across membranes, but the reflection coefficient falls somewhere between 0 and 1. In such cases, the effective osmotic pressure lies between its maximal possible value (when the solute is completely impermeable) and zero (when the solute is freely permeable). Refer once again to the van't Hoff equation and notice that, when σ is between 0 and 1, the calculated effective osmotic pressure will be less than its maximal possible value but greater than zero.

When two solutions separated by a semipermeable membrane have the same effective osmotic pressure,

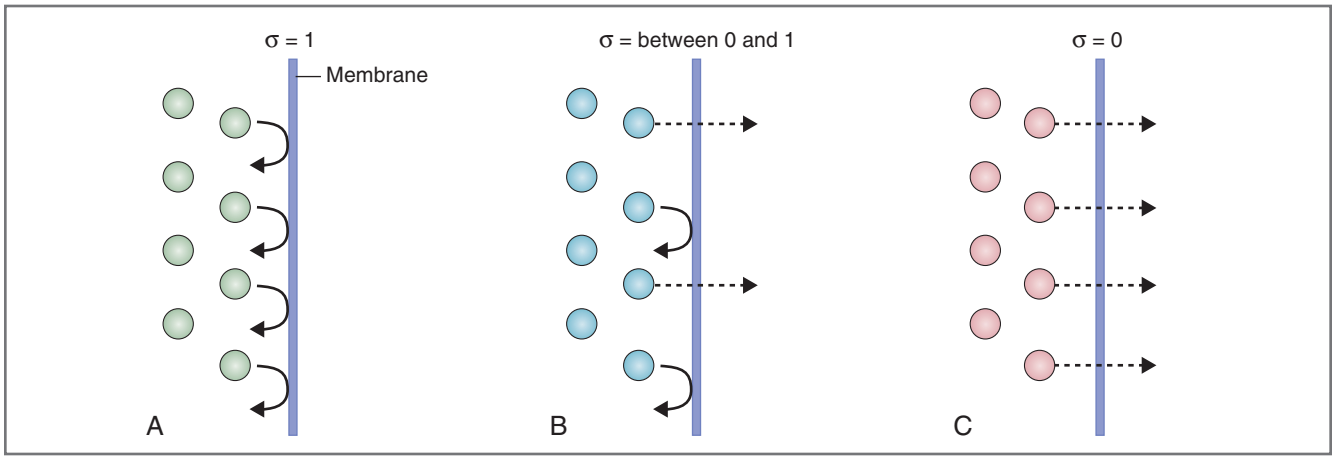


Fig. 1.10 **Reflection coefficient (σ).**

they are **isotonic;** that is, no water will flow between them because there is no effective osmotic pressure difference across the membrane. When two solutions have different effective osmotic pressures, the solution with the lower effective osmotic pressure is **hypotonic** and the solution with the higher effective osmotic pressure is **hypertonic.** *Water will flow from the hypotonic solution into the hypertonic solution (Box 1.2).*

SAMPLE PROBLEM. A solution of 1 mol/L NaCl is separated from a solution of 2 mol/L urea by a semipermeable membrane. Assume that NaCl is completely dissociated, that $\sigma_{NaCl} = 0.3$, and $\sigma_{urea} = 0.05$. *Are the two solutions isosmotic and/or isotonic? Is there net water flow, and what is its direction?*

SOLUTION

Step 1. To determine whether the solutions are isosmotic, simply calculate the osmolarity of each solution (g × C) and compare the two values. It was stated that NaCl is completely dissociated (i.e., separated into two particles); thus for NaCl, g = 2.0. Urea does not dissociate in solution; thus for urea, g = 1.0.

$$\text{NaCl: Osmolarity} = g\ C = 2.0 \times 1\ \text{mol/L} = 2\ \text{Osm/L}$$

$$\text{Urea: Osmolarity} = g\ C = 1.0 \times 2\ \text{mol/L} = 2\ \text{Osm/L}$$

Each solution has an osmolarity of 2 Osm/L—they are indeed isosmotic.

Step 2. To determine whether the solutions are isotonic, the effective osmotic pressure of each solution must be determined. Assume that at 37°C (310 K), RT = 25.45 L-atm/mol. Thus

$$\text{NaCl: } \pi = g\ C\ \sigma\ RT = 2 \times 1\ \text{mol/L} \times 0.3 \times RT = 0.6\ RT = 15.3\ \text{atm}$$

$$\text{Urea: } \pi = g\ C\ \sigma\ RT = 1 \times 2\ \text{mol/L} \times 0.05 \times RT = 0.1\ RT = 2.5\ \text{atm}$$

Although the two solutions have the same calculated osmolarities and are isosmotic (Step 1), they have different effective osmotic pressures and they are not isotonic (Step 2). This difference occurs because the reflection coefficient for NaCl is much higher than the reflection coefficient for urea and, thus NaCl creates the greater *effective* osmotic pressure. Water will flow from the urea solution into the NaCl solution, from the hypotonic solution to the hypertonic solution.

DIFFUSION POTENTIALS AND EQUILIBRIUM POTENTIALS

Ion Channels

Ion channels are integral, membrane-spanning proteins that, when open, permit the passage of certain ions. Thus ion channels are **selective** and allow ions with specific characteristics to move through them. This selectivity is based on both the size of the channel and the charges lining it. For example, channels lined with negative charges typically permit the passage of cations but exclude anions; channels lined with positive charges permit the passage of anions but exclude cations. Channels also discriminate on the basis of size. For example, a cation-selective channel lined with negative charges might permit the passage of Na^+ but exclude K^+; another

BOX 1.2 Clinical Physiology: Hyposmolarity With Brain Swelling

DESCRIPTION OF CASE. A 72-year-old man was diagnosed recently with oat cell carcinoma of the lung. He tried to stay busy with consulting work, but the disease sapped his energy. One evening, his wife noticed that he seemed confused and lethargic, and suddenly he suffered a grand mal seizure. In the emergency department, his plasma Na^+ concentration was 113 mEq/L (normal, 140 mEq/L) and his plasma osmolarity was 230 mOsm/L (normal, 290 mOsm/L). He was treated immediately with an infusion of hypertonic NaCl and was released from the hospital a few days later, with strict instructions to limit his water intake.

EXPLANATION OF CASE. The man's oat cell carcinoma autonomously secretes antidiuretic hormone (ADH), which causes syndrome of inappropriate antidiuretic hormone (SIADH). In SIADH, the high circulating levels of ADH cause excessive water reabsorption by the principal cells of the late distal tubule and collecting ducts. The excess water that is reabsorbed and retained in the body dilutes the Na^+ concentration and osmolarity of the ECF. The decreased osmolarity means there is also decreased effective osmotic pressure of ECF and, briefly, osmotic pressure of ECF is less than osmotic pressure of ICF. The effective osmotic pressure difference across cell membranes causes osmotic water flow from ECF to ICF, which results in cell swelling. Because the brain is contained in a fixed structure (the skull), swelling of brain cells can cause seizure.

TREATMENT. Treatment of the patient with hypertonic NaCl infusion was designed to quickly raise his ECF osmolarity and osmotic pressure, which would eliminate the effective osmotic pressure difference across the brain cell membranes and stop osmotic water flow and brain cell swelling.

cation-selective channel (e.g., nicotinic receptor on the motor end plate) might have less selectivity and permit the passage of several different small cations.

Ion channels are controlled by **gates,** and, depending on the position of the gates, the channels may be open or closed. When a channel is open, the ions for which it is selective can flow through it by passive diffusion, down the existing electrochemical gradient. In the open state, there is a continuous path between ECF and ICF, through which ions can flow. When the channel is closed, the ions cannot flow through it, no matter what the size of the electrochemical gradient. The **conductance** of a channel depends on the probability that it is open. The higher the probability that the channel is open, the higher is its conductance or permeability.

The gates on ion channels are controlled by three types of **sensors.** One type of gate has sensors that respond to changes in membrane potential (i.e., voltage-gated channels); a second type of gate responds to changes in signaling molecules (i.e., second messenger–gated channels); and a third type of gate responds to changes in ligands such as hormones or neurotransmitters (i.e., ligand-gated channels).

- **Voltage-gated channels** have gates that are controlled by changes in membrane potential. For example, the **activation gate on the nerve Na^+ channel** is *opened* by depolarization of the nerve cell membrane; opening of this channel is responsible for the upstroke of the action potential. Interestingly, another gate on the Na^+ channel, an **inactivation gate,** is *closed* by depolarization. Because the activation gate responds more rapidly to depolarization than the inactivation gate, the Na^+ channel first opens and then closes. This difference in response times of the two gates accounts for the shape and time course of the action potential.
- **Second messenger–gated channels** have gates that are controlled by changes in levels of intracellular signaling molecules such as cyclic adenosine monophosphate (cAMP) or inositol 1,4,5-triphosphate (IP_3). Thus the sensors for these gates are on the intracellular side of the ion channel. For example, the gates on Na^+ channels in cardiac sinoatrial node are opened by increased intracellular cAMP.
- **Ligand-gated channels** have gates that are controlled by hormones and neurotransmitters. The sensors for these gates are located on the extracellular side of the ion channel. For example, the **nicotinic receptor** on the **motor end plate** is actually an ion channel that opens when acetylcholine (ACh) binds to it; when open, it is permeable to Na^+ and K^+ ions.

Diffusion Potentials

A diffusion potential is the potential difference generated across a membrane when a charged solute (an ion) diffuses down its concentration gradient. Therefore a **diffusion potential is caused by diffusion of ions.** It follows, then, that a diffusion potential can be generated *only* if the membrane is permeable to that ion. Furthermore, if the membrane is not permeable to the ion, no diffusion potential will be generated no matter how large a concentration gradient is present.

The **magnitude** of a diffusion potential, measured in millivolts (mV), depends on the size of the concentration gradient, where the concentration gradient is the driving force. The **sign** of the diffusion potential depends on the charge of the diffusing ion. Finally, as noted, diffusion potentials are created by the

movement of only a few ions, and they do not cause changes in the concentration of ions in bulk solution.

Equilibrium Potentials

The concept of equilibrium potential is simply an extension of the concept of diffusion potential. If there is a concentration difference for an ion across a membrane and the membrane is permeable to that ion, a potential difference (the diffusion potential) is created. Eventually, net diffusion of the ion slows and then stops because of that potential difference. In other words, if a cation diffuses down its concentration gradient, it carries a positive charge across the membrane, which will retard and eventually stop further diffusion of the cation. If an anion diffuses down its concentration gradient, it carries a negative charge, which will retard and then stop further diffusion of the anion. The **equilibrium potential** is the diffusion potential that exactly balances or opposes the tendency for diffusion down the concentration difference. At **electrochemical equilibrium,** the chemical and electrical driving forces acting on an ion are equal and opposite, and no further net diffusion occurs.

The following examples of a diffusing cation and a diffusing anion illustrate the concepts of equilibrium potential and electrochemical equilibrium.

Example of Na^+ Equilibrium Potential

Figure 1.11 shows two solutions separated by a theoretical membrane that is permeable to Na^+ but not to Cl^-. The NaCl concentration is higher in Solution 1 than in Solution 2. The permeant ion, Na^+, will diffuse down its concentration gradient from Solution 1 to Solution 2, but the impermeant ion, Cl^-, will not accompany it. As a result of the net movement of positive charge to Solution 2, an **Na^+ diffusion potential** develops and Solution 2 becomes positive with respect to Solution 1. The positivity in Solution 2 opposes further diffusion of Na^+, and eventually it is large enough to prevent further net diffusion. The potential difference that exactly balances the tendency of Na^+ to diffuse down its concentration gradient is the **Na^+ equilibrium potential.** When the chemical and electrical driving forces on Na^+ are equal and opposite, Na^+ is said to be at **electrochemical equilibrium.** This diffusion of a *few* Na^+ ions, sufficient to create the diffusion potential, does not produce any change in Na^+ concentration in the bulk solutions.

Example of Cl^- Equilibrium Potential

Figure 1.12 shows the same pair of solutions as in Figure 1.11; however, in Figure 1.12, the theoretical membrane is permeable to Cl^- rather than to Na^+. Cl^- will diffuse from Solution 1 to Solution 2 down its concentration gradient, but Na^+ will not accompany it. A diffusion potential will be established, and Solution 2 will become negative relative to Solution 1. The potential difference that exactly balances the tendency of Cl^- to diffuse down its concentration gradient is the **Cl^- equilibrium potential.** When the chemical and electrical driving forces on Cl^- are equal and opposite, then Cl^- is at **electrochemical equilibrium.** Again, diffusion of these few Cl^- ions will not change the Cl^- concentration in the bulk solutions.

Nernst Equation

The Nernst equation is used to calculate the equilibrium potential for an ion at a given concentration difference across a membrane, assuming that the membrane is permeable to that ion. By definition, the equilibrium potential is calculated for *one ion at a time.* Thus

$$E_x = \frac{-2.3RT}{zF} \log_{10} \frac{[C_i]}{[C_e]}$$

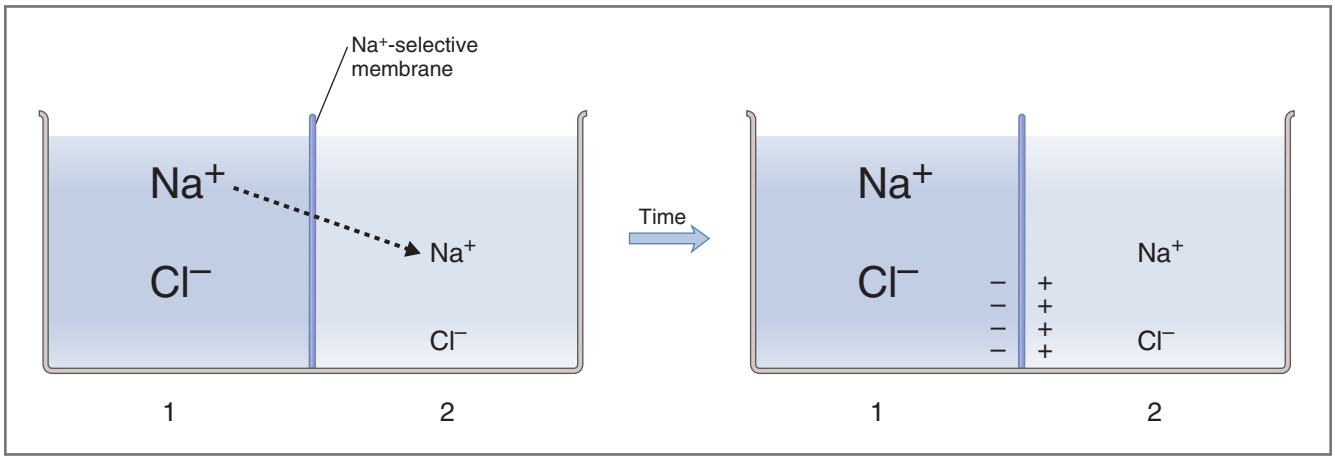


Fig. 1.11 Generation of an Na^+ diffusion potential.

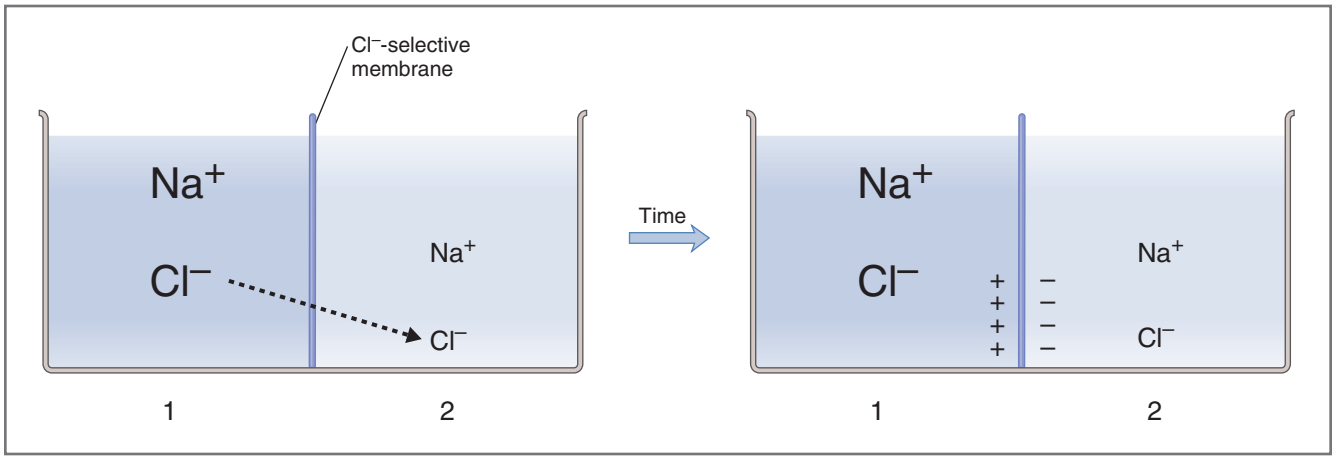


Fig. 1.12 **Generation of a Cl^- diffusion potential.**

where

E_X = Equilibrium potential (mV) for a given ion, X

$\frac{2.3RT}{F}$ = Constant (60 mV at 37°C)

z = Charge on the ion (+1 for Na^+; + 2 for Ca^{2+}; − 1 for Cl^-)

C_i = Intracellular concentration of X (mmol/L)

C_e = Extracellular concentration of X (mmol/L)

In words, the Nernst equation converts a concentration difference for an ion into a voltage. This conversion is accomplished by the various constants: R is the gas constant, T is the absolute temperature, and F is Faraday constant; multiplying by 2.3 converts natural logarithm to $\log_{10}$.

By convention, *membrane potential is expressed as intracellular potential relative to extracellular potential.* Hence, a transmembrane potential difference of −70 mV means 70 mV, cell interior negative.

Typical values for equilibrium potential for common ions in skeletal muscle, calculated as previously described and assuming typical concentration gradients across cell membranes, are as follows:

$$E_{Na^+} = +65 \text{ mV}$$
$$E_{Ca^{2+}} = +120 \text{ mV}$$
$$E_{K^+} = -95 \text{ mV}$$
$$E_{Cl^-} = -90 \text{ mV}$$

It is useful to keep these values in mind when considering the concepts of resting membrane potential and action potentials.

SAMPLE PROBLEM. *If the intracellular $[Ca^{2+}]$ is 10^{-7} mol/L and the extracellular $[Ca^{2+}]$ is 2×10^{-3} mol/L, at what potential difference across the cell membrane will Ca^{2+} be at electrochemical equilibrium?* Assume that 2.3RT/F = 60 mV at body temperature (37°C).

SOLUTION. Another way of posing the question is to ask what the membrane potential will be, given this concentration gradient across the membrane, if Ca^{2+} is the only permeant ion. Remember, Ca^{2+} is divalent, so z = +2. Thus

$$E_{Ca^{2+}} = \frac{-60 \text{ mV}}{z} \log_{10} \frac{C_i}{C_e}$$
$$= \frac{-60 \text{ mV}}{+2} \log_{10} \frac{10^{-7} \text{ mol/L}}{2 \times 10^{-3} \text{ mol/L}}$$
$$= -30 \text{ mV} \log_{10} 5 \times 10^{-5}$$
$$= -30 \text{ mV } (-4.3)$$
$$= +129 \text{ mV}$$

Because this is a log function, it is not necessary to remember which concentration goes in the numerator. Simply complete the calculation either way to arrive at 129 mV, and then determine the correct sign with an intuitive approach. The intuitive approach depends on the knowledge that, because the $[Ca^{2+}]$ is much higher in ECF than in ICF, Ca^{2+} will tend to diffuse down this concentration gradient from ECF into ICF, making the inside of the cell positive. Thus Ca^{2+} will be at electrochemical equilibrium when the membrane potential is +129 mV (cell interior positive).

Be aware that the equilibrium potential has been calculated at a given concentration gradient for Ca^{2+} ions. With a different concentration gradient, the calculated equilibrium potential would be different.

Driving Force

When dealing with uncharged solutes, the driving force for net diffusion is simply the concentration difference of the solute across the cell membrane. However, when dealing with charged solutes (i.e., ions), the driving force for net diffusion must consider both concentration difference and electrical potential difference across the cell membrane.

The **driving force** on a given ion is the difference between the actual, measured membrane potential (E_m) and the ion's calculated equilibrium potential (E_X). In other words, it is the difference between the actual E_m and the value the ion would "like" the membrane potential to be. (The ion would "like" the membrane potential to be its equilibrium potential, as calculated by the Nernst equation.) The driving force on a given ion, X, is therefore calculated as:

$$\text{Net driving force (mV)} = E_m - E_x$$

where

$$\text{Driving force} = \text{Driving force (mV)}$$
$$E_m = \text{Actual membrane potential (mV)}$$
$$E_X = \text{Equilibrium potential for X (mV)}$$

When the driving force is negative (i.e., E_m is more negative than the ion's equilibrium potential), that ion X will enter the cell if it is a cation and will leave the cell if it is an anion. In other words, ion X "thinks" the membrane potential is too negative and tries to bring the membrane potential toward its equilibrium potential by diffusing in the appropriate direction across the cell membrane. Conversely, if the driving force is positive (E_m is more positive than the ion's equilibrium potential), then ion X will leave the cell if it is a cation and will enter the cell if it is an anion; in this case, ion X "thinks" the membrane potential is too positive and tries to bring the membrane potential toward its equilibrium potential by diffusing in the appropriate direction across the cell membrane. Finally, if E_m is equal to the ion's equilibrium potential, then the driving force on the ion is zero, and the ion is, by definition, at electrochemical equilibrium; since there is no driving force, there will be no net movement of the ion in either direction.

Ionic Current

Ionic current (I_X), or current flow, occurs when there is movement of an ion across the cell membrane. Ions will move across the cell membrane through ion channels when two conditions are met: (1) there is a driving force on the ion, and (2) the membrane has a conductance to that ion (i.e., its ion channels are open). Thus

$$I_X = G_X(E_m - E_X)$$

where

$$I_X = \text{ionic current (mAmp)}$$
$$G_X = \text{ionic conductance (1/ohm), where conductance is the reciprocal of resistance}$$
$$E_m - E_X = \text{driving force on ion X (mV)}$$

You will notice that the equation for ionic current is simply a rearrangement of Ohm's law, where V = IR or I = V/R (where V is the same thing as E). Because conductance (G) is the reciprocal of resistance (R), $I = G \times V$.

The **direction of ionic current** is determined by the direction of the driving force, as described in the previous section. The **magnitude of ionic current** is determined by the size of the driving force and the conductance of the ion. For a given conductance, the greater the driving force, the greater the current flow. For a given driving force, the greater the conductance, the greater the current flow. Lastly, if either the driving force or the conductance of an ion is zero, there can be no net diffusion of that ion across the cell membrane and no current flow.

RESTING MEMBRANE POTENTIAL

The resting membrane potential is the potential difference that exists across the membrane of excitable cells such as nerve and muscle in the period between action potentials (i.e., at rest). As stated previously, in expressing the membrane potential, it is conventional to refer the intracellular potential to the extracellular potential.

The resting membrane potential is established by diffusion potentials, which result from the concentration differences for various ions across the cell membrane. (Recall that these concentration differences have been established by primary and secondary active transport mechanisms.) *Each permeant ion attempts to drive the membrane potential toward its own equilibrium potential.* Ions with the highest permeabilities or conductances at rest will make the greatest contributions to the resting membrane potential, and those with the lowest permeabilities will make little or no contribution.

The resting membrane potential of most excitable cells falls in the range of **–70 to –80 mV.** These values can best be explained by the concept of relative permeabilities of the cell membrane. Thus the resting membrane potential is *close to* the equilibrium potentials for K^+ and Cl^- because the permeability to these ions at rest is high. The resting membrane potential is *far from*

the equilibrium potentials for Na^+ and Ca^{2+} because the permeability to these ions at rest is low.

One way of evaluating the contribution each ion makes to the membrane potential is by using the **chord conductance equation,** which weights the equilibrium potential for each ion (calculated by the Nernst equation) by its relative conductance. Ions with the highest conductance drive the membrane potential toward their equilibrium potentials, whereas those with low conductance have little influence on the membrane potential. (An alternative approach to the same question applies the **Goldman equation,** which considers the contribution of each ion by its relative permeability rather than by its conductance.) The chord conductance equation is written as follows:

$$E_m = \frac{g_{K^+}}{g_T} E_{K^+} + \frac{g_{Na^+}}{g_T} E_{Na^+} + \frac{g_{Cl^-}}{g_T} E_{Cl^-} + \frac{g_{Ca^{2+}}}{g_T} E_{Ca^{2+}}$$

where

E_m = Membrane potential (mV)

g_{K^+} etc. = K^+ conductance etc. (mho, reciprocal of resistance)

g_T = Total conductance (mho)

E_{K^+} etc. = K^+ equilibrium potential etc. (mV)

At rest, the membranes of excitable cells are far more permeable to K^+ and Cl^- than to Na^+ and Ca^{2+}. These differences in permeability account for the resting membrane potential.

What role, if any, does the Na^+-K^+ ATPase play in creating the resting membrane potential? The answer has two parts. First, there is a small *direct* electrogenic contribution of the Na^+-K^+ ATPase, which is based on the stoichiometry of three Na^+ ions pumped out of the cell for every two K^+ ions pumped into the cell. Second, the more important *indirect* contribution is in maintaining the concentration gradient for K^+ across the cell membrane, which then is responsible for the K^+ diffusion potential that drives the membrane potential toward the K^+ equilibrium potential. Thus the Na^+-K^+ ATPase is necessary to create and maintain the K^+ concentration gradient, which establishes the resting membrane potential. (A similar argument can be made for the role of the Na^+-K^+ ATPase in the upstroke of the action potential, where it maintains the ionic gradient for Na^+ across the cell membrane.)

ACTION POTENTIALS

The action potential is a phenomenon of excitable cells such as nerve and muscle and consists of a rapid depolarization (upstroke) followed by repolarization of the membrane potential. Action potentials are the basic mechanism for transmission of information in the nervous system and in all types of muscle.

Terminology

The following terminology will be used for discussion of the action potential, the refractory periods, and the propagation of action potentials:

- **Depolarization** is the process of making the membrane potential *less negative.* As noted, the usual resting membrane potential of excitable cells is oriented with the cell interior negative. Depolarization makes the interior of the cell less negative, or it may even cause the cell interior to become positive. Such a change in membrane potential should not be described as "increasing" or "decreasing" because those terms are ambiguous. (For example, when the membrane potential depolarizes, or becomes less negative, has the membrane potential increased or decreased?)
- **Hyperpolarization** is the process of making the membrane potential *more negative.* As with depolarization, the terms "increasing" or "decreasing" should not be used to describe a change that makes the membrane potential more negative.
- **Inward current** is the flow of positive charge into the cell. Thus inward currents *depolarize* the membrane potential. An example of an inward current is the flow of Na^+ into the cell during the upstroke of the action potential.
- **Outward current** is the flow of positive charge out of the cell. Outward currents *hyperpolarize* the membrane potential. An example of an outward current is the flow of K^+ out of the cell during the repolarization phase of the action potential.
- **Threshold potential** is the membrane potential at which occurrence of the action potential is inevitable. Because the threshold potential is less negative than the resting membrane potential, an inward current is required to depolarize the membrane potential to threshold. At threshold potential, net inward current (e.g., inward Na^+ current) becomes larger than net outward current (e.g., outward K^+ current), and the resulting depolarization becomes self-sustaining, giving rise to the upstroke of the action potential. If net inward current is less than net outward current, the membrane will not be depolarized to threshold and no action potential will occur (see all-or-none response).
- **Overshoot** is that portion of the action potential where the membrane potential is positive (cell interior positive).

- **Undershoot,** or **hyperpolarizing afterpotential,** is that portion of the action potential, following repolarization, where the membrane potential is actually more negative than it is at rest.
- **Refractory period** is a period during which another normal action potential cannot be elicited in an excitable cell. Refractory periods can be absolute or relative. (In cardiac muscle cells, there is an additional category called effective refractory period.)

Characteristics of Action Potentials

Action potentials have three basic characteristics: stereotypical size and shape, propagation, and all-or-none response.

- **Stereotypical size and shape.** Each *normal* action potential for a given cell type looks identical, depolarizes to the same potential, and repolarizes back to the same resting potential.
- **Propagation.** An action potential at one site causes depolarization at adjacent sites, bringing those adjacent sites to threshold. Propagation of action potentials from one site to the next is *nondecremental.*
- **All-or-none response.** An action potential either occurs or does not occur. If an excitable cell is depolarized to threshold in a *normal* manner, then the occurrence of an action potential is inevitable. On the other hand, if the membrane is not depolarized to threshold, no action potential can occur. Indeed, if the stimulus is applied during the refractory period, then either no action potential occurs, or the action potential will occur but not have the stereotypical size and shape.

Ionic Basis of the Action Potential

The action potential is a fast depolarization (the upstroke), followed by repolarization back to the resting membrane potential. Figure 1.13 illustrates the events of the action potential in nerve and skeletal muscle, which occur in the following steps:

1. **Resting membrane potential.** At rest, the membrane potential is approximately –70 mV (cell interior

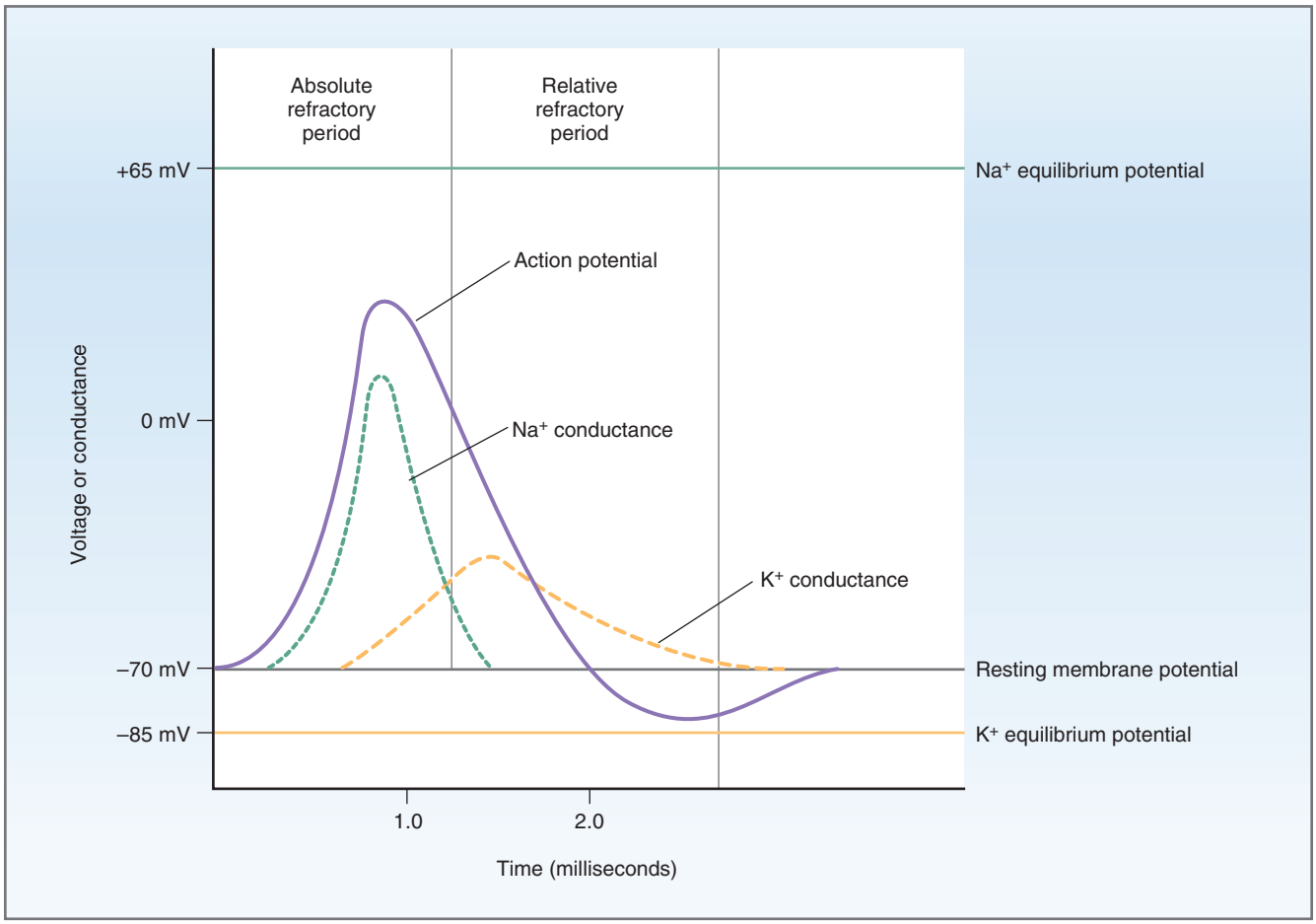


Fig. 1.13 **Time course of voltage and conductance changes during the action potential of nerve.**

negative). The **K^+ conductance or permeability is high** and K^+ channels are almost fully open, allowing K^+ ions to diffuse out of the cell down the existing concentration gradient. This diffusion creates a K^+ diffusion potential, which drives the membrane potential toward the K^+ equilibrium potential. The conductance to Cl^- (not shown) also is high, and, at rest, Cl^- also is near electrochemical equilibrium. At rest, the **Na^+ conductance is low,** and thus the resting membrane potential is far from the Na^+ equilibrium potential, and Na^+ is far from electrochemical equilibrium.

2. **Upstroke of the action potential.** An inward current, usually the result of current spread from action potentials at neighboring sites, causes depolarization of the nerve cell membrane to threshold, which occurs at approximately –60 mV. This initial depolarization causes rapid opening of the **activation gates** of the Na^+ channel, and the Na^+ conductance promptly increases and becomes even higher than the K^+ conductance (Fig. 1.14). The increase in Na^+ conductance results in an **inward Na^+ current;** the membrane potential is further depolarized toward, but does not quite reach, the Na^+ equilibrium potential of +65 mV. **Tetrodotoxin** (a toxin from the Japanese puffer fish) and the local anesthetic **lidocaine** block these voltage-sensitive Na^+ channels and prevent the occurrence of nerve action potentials.

3. **Repolarization of the action potential.** The upstroke is terminated, and the membrane potential repolarizes to the resting level as a result of two events. First, the inactivation gates on the Na^+ channels respond to depolarization by closing, but their response is slower than the opening of the activation gates. Thus after a delay, the **inactivation gates** close, which closes the Na^+ channels and terminates the upstroke. Second, depolarization opens K^+ channels and increases K^+ conductance to a value even higher than occurs at rest. The combined effect of closing of the Na^+ channels and greater opening of the K^+ channels makes the K^+ conductance much higher than the Na^+ conductance. Thus an **outward K^+ current** results, and the membrane is repolarized. **Tetraethylammonium (TEA)** blocks these voltage-gated K^+ channels, the outward K^+ current, and repolarization.

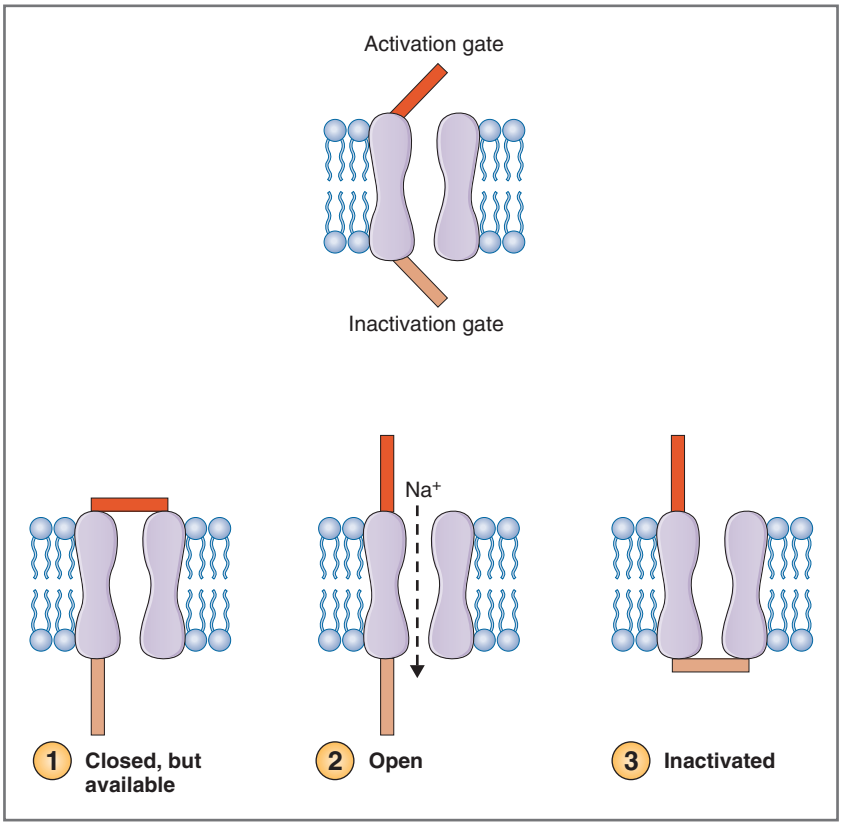


Fig. 1.14 States of activation and inactivation gates on the nerve Na^+ channel. *1,* In the closed but available state, at the resting membrane potential, the activation gate is closed, the inactivation gate is open, and the channel is closed (but available, if depolarization occurs). *2,* In the open state, during the upstroke of the action potential, both the activation and inactivation gates are open and the channel is open. *3,* In the inactivated state, at the peak of the action potential, the activation gate is open, the inactivation gate is closed, and the channel is closed.

4. **Hyperpolarizing afterpotential (undershoot).** For a brief period following repolarization, the K^+ conductance is higher than at rest and the membrane potential is driven even closer to the K^+ equilibrium potential (hyperpolarizing afterpotential). Eventually, the K^+ conductance returns to the resting level, and the membrane potential depolarizes slightly, back to the resting membrane potential. The membrane is now ready, if stimulated, to generate another action potential.

The Nerve Na^+ Channel

A voltage-gated Na^+ channel is responsible for the upstroke of the action potential in nerve and skeletal muscle. This channel is an integral membrane protein, consisting of a large α subunit and two β subunits. The α subunit has four domains, each of which has six transmembrane α-helices. The repeats of transmembrane α-helices surround a central pore, through which Na^+ ions can flow (if the channel's gates are open). A conceptual model of the Na^+ channel demonstrating the function of the activation and inactivation gates is shown in Figure 1.14. The basic assumption of this model is that in order for Na^+ to move through the channel, *both gates on the channel must be open.* Recall how these gates respond to changes in voltage. The activation gates open quickly in response to depolarization. The inactivation gates close in response to depolarization, but slowly, after a time delay. Thus when depolarization occurs, the activation gates open quickly, followed by slower closing of the inactivation gates. The figure shows three combinations of the gates' positions and the resulting effect on Na^+ channel opening.

1. **Closed, but available.** At the **resting membrane potential,** the activation gates are closed and the inactivation gates are open. Thus the Na^+ channels are closed. However, they are "available" to fire an action potential *if* depolarization occurs. (Depolarization would open the activation gates and, because the inactivation gates are already open, the Na^+ channels would then be open.)
2. **Open.** During the **upstroke** of the action potential, depolarization quickly opens the activation gates and both the activation and inactivation gates are briefly open. Na^+ can flow through the channels into the cell, causing further depolarization.
3. **Inactivated.** At the peak of the action potential, the slow inactivation gates finally close in response to depolarization; now the Na^+ channels are closed, the upstroke is terminated, and repolarization begins.

How do the Na^+ channels return to the closed, but available state? In other words, *how do they recover,* so that they are ready to fire another action potential? Repolarization back to the resting membrane potential causes the inactivation gates to open. The Na^+ channels now return to the closed, but available state and are ready and "available" to fire another action potential if depolarization occurs.

Refractory Periods

During the refractory periods, excitable cells are *incapable* of producing normal action potentials (see Fig. 1.13). The refractory period includes an absolute refractory period and a relative refractory period.

Absolute Refractory Period

The absolute refractory period overlaps with almost the entire duration of the action potential. During this period, no matter how great the stimulus, another action potential cannot be elicited. The basis for the absolute refractory period is closure of the inactivation gates of the Na^+ channel in response to depolarization. These inactivation gates are in the closed position until the cell is repolarized back to the resting membrane potential and the Na^+ channels have recovered to the "closed, but available" state (see Fig. 1.14).

Relative Refractory Period

The relative refractory period begins at the end of the absolute refractory period and overlaps primarily with the period of the hyperpolarizing afterpotential. During this period, an action potential can be elicited, but only if a greater than usual depolarizing (inward) current is applied. The basis for the relative refractory period is the higher K^+ conductance than is present at rest. Because the membrane potential is closer to the K^+ equilibrium potential, more inward current is needed to bring the membrane to threshold for the next action potential to be initiated.

Accommodation

When a nerve or muscle cell is depolarized slowly or is held at a depolarized level, the usual threshold potential may pass without an action potential having been fired. This process, called accommodation, occurs because depolarization closes inactivation gates on the Na^+ channels. If depolarization occurs slowly enough, the Na^+ channels close and remain closed. The upstroke of the action potential cannot occur because there are insufficient available Na^+ channels to carry inward current. An example of accommodation is seen in persons who have an elevated serum K^+ concentration, or **hyperkalemia.** At rest, nerve and muscle cell membranes are very permeable to K^+; an increase in extracellular K^+ concentration causes depolarization of the resting membrane (as dictated by the Nernst equation). This depolarization brings the cell membrane closer to

BOX 1.3 Clinical Physiology: Hyperkalemia With Muscle Weakness

DESCRIPTION OF CASE. A 48-year-old woman with insulin-dependent diabetes mellitus reports to her physician that she is experiencing severe muscle weakness. She is being treated for hypertension with propranolol, a β-adrenergic blocking agent. Her physician immediately orders blood studies, which reveal a serum [K^+] of 6.5 mEq/L (normal, 4.5 mEq/L) and elevated BUN (blood urea nitrogen). The physician tapers off the dosage of propranolol, with eventual discontinuation of the drug. He adjusts her insulin dosage. Within a few days, the patient's serum [K^+] has decreased to 4.7 mEq/L, and she reports that her muscle strength has returned to normal.

EXPLANATION OF CASE. This diabetic patient has severe hyperkalemia caused by several factors: (1) Because her insulin dosage is insufficient, the lack of adequate insulin has caused a shift of K^+ out of cells into blood (insulin promotes K^+ uptake into cells). (2) Propranolol, the β-blocking agent used to treat the woman's hypertension, also shifts K^+ out of cells into blood. (3) Elevated BUN suggests that the woman is developing renal failure; her failing kidneys are unable to excrete the extra K^+ that is accumulating in her blood. These mechanisms involve concepts related to renal physiology and endocrine physiology.

It is important to understand that this woman has a severely elevated blood [K^+] (hyperkalemia) and that her muscle weakness results from this hyperkalemia. The basis for this weakness can be explained as follows: The resting membrane potential of muscle cells is determined by the concentration gradient for K^+ across the cell membrane (Nernst equation). At rest, the cell membrane is very permeable to K^+, and K^+ diffuses out of the cell down its concentration gradient, creating a K^+ diffusion potential. This K^+ diffusion potential is responsible for the resting membrane potential, which is cell interior negative. The larger the K^+ concentration gradient, the greater the negativity in the cell. When the blood [K^+] is elevated, the concentration gradient across the cell membrane is less than normal; resting membrane potential will therefore be less negative (i.e., depolarized).

It might be expected that this depolarization would make it easier to generate action potentials in the muscle because the resting membrane potential would be closer to threshold. A more important effect of depolarization, however, is that it closes the inactivation gates on Na^+ channels. When these inactivation gates are closed, no action potentials can be generated, even if the activation gates are open. Without action potentials in the muscle, there can be no contraction.

TREATMENT. Treatment of this patient is based on shifting K^+ back into the cells by increasing the woman's insulin dosages and by discontinuing propranolol. By reducing the woman's blood [K^+] to normal levels, the resting membrane potential of her skeletal muscle cells will return to normal, the inactivation gates on the Na^+ channels will be open at the resting membrane potential (as they should be), and normal action potentials can occur.

threshold and would seem to make it more likely to fire an action potential. However, the cell is actually *less* likely to fire an action potential because this sustained depolarization closes the inactivation gates on the Na^+ channels (Box 1.3).

Propagation of Action Potentials

Propagation of action potentials down a nerve or muscle fiber occurs by the spread of **local currents** from active regions to adjacent inactive regions. Figure 1.15 shows a nerve cell body with its dendritic tree and an axon. At rest, the entire nerve axon is at the resting membrane potential, with the cell interior negative. Action potentials are initiated in the initial segment of the axon, nearest the nerve cell body. They propagate down the axon by spread of local currents, as illustrated in the figure.

In Figure 1.15A the initial segment of the nerve axon is depolarized to threshold and fires an action potential (the active region). As the result of an inward Na^+ current, at the peak of the action potential, the polarity of the membrane potential is reversed and the cell interior becomes positive. The adjacent region of the axon remains inactive, with its cell interior negative.

Figure 1.15B illustrates the spread of local current from the depolarized active region to the adjacent inactive region. At the active site, positive charges inside the cell flow toward negative charges at the adjacent inactive site. This current flow causes the adjacent region to depolarize to threshold.

In Figure 1.15C the adjacent region of the nerve axon, having been depolarized to threshold, now fires an action potential. The polarity of its membrane potential is reversed, and the cell interior becomes positive. At this time, the *original* active region has been repolarized back to the resting membrane potential and restored to its inside-negative polarity. The process continues, transmitting the action potential sequentially down the axon.

Conduction Velocity

The speed at which action potentials are conducted along a nerve or muscle fiber is the conduction velocity. This property is of great physiologic importance because

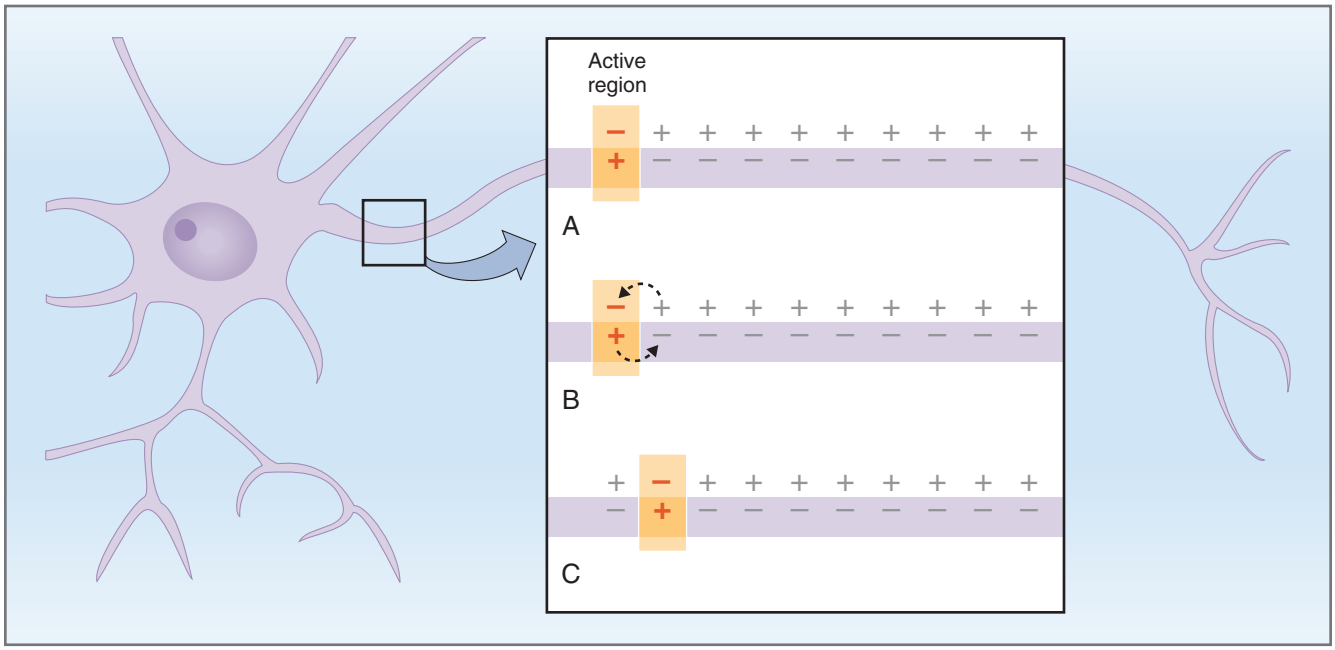


Fig. 1.15 Spread of depolarization down a nerve fiber by local currents. A, The initial segment of the axon has fired an action potential, and the potential difference across the cell membrane has reversed to become inside positive. The adjacent area is inactive and remains at the resting membrane potential, inside negative. **B,** At the active site, positive charges inside the nerve flow to the adjacent inactive area. **C,** Local current flow causes the adjacent area to be depolarized to threshold and to fire action potentials; the original active region has repolarized back to the resting membrane potential.

it determines the speed at which information can be transmitted in the nervous system. To understand conduction velocity in excitable tissues, two major concepts must be explained: the time constant and the length constant. These concepts, called **cable properties,** explain how nerves and muscles act as cables to conduct electrical activity.

The **time constant (τ)** is the amount of time it takes following the injection of current for the potential to change to 63% of its final value. In other words, the time constant indicates how quickly a cell membrane depolarizes in response to an inward current or how quickly it hyperpolarizes in response to an outward current. Thus

$$\tau = R_m C_m$$

where

τ = Time constant
R_m = Membrane resistance
C_m = Membrane capacitance

Two factors affect the time constant. The first factor is **membrane resistance (R_m).** When R_m is high, current does not readily flow across the cell membrane, which makes it difficult to change the membrane potential, thus increasing the time constant. The second factor, **membrane capacitance (C_m),** is the ability of the cell membrane to store charge. When C_m is high, the time constant is increased because injected current first must discharge the membrane capacitor before it can depolarize the membrane. Thus the time constant is greatest (i.e., takes longest) when R_m and C_m are high.

The **length constant (λ)** is the distance from the site of current injection where the potential has fallen by 63% of its original value. The length constant indicates how far a depolarizing current will spread along a nerve. In other words, the longer the length constant, the farther the current spreads down the nerve fiber. Thus

$$\lambda \propto \sqrt{R_m / R_i}$$

where

λ = Length constant
R_m = Membrane resistance
R_i = Internal resistance

Again, R_m represents membrane resistance. Internal resistance, R_i, is inversely related to the ease of current flow in the cytoplasm of the nerve fiber. Therefore the length constant will be greatest (i.e., current will travel the farthest) when the diameter of the nerve is large, when membrane resistance is high, and when internal

resistance is low. In other words, current flows along the path of least resistance.

Changes in Conduction Velocity

There are two mechanisms that *increase* conduction velocity along a nerve: increasing the size of the nerve fiber and myelinating the nerve fiber. These mechanisms can best be understood in terms of the cable properties of time constant and length constant.

- **Increasing nerve diameter.** Increasing the size of a nerve fiber increases conduction velocity, a relationship that can be explained as follows: Internal resistance, R_i, is inversely proportional to the cross-sectional area ($A = \pi r^2$). Therefore the larger the fiber, the lower the internal resistance. The length constant is inversely proportional to the square root of R_i (refer to the equation for length constant). Thus the length constant (λ) will be large when internal resistance (R_i) is small (i.e., fiber size is large). The largest nerves have the longest length constants, and current spreads farthest from the active region to propagate action potentials. Increasing nerve fiber size is certainly an important mechanism for increasing conduction velocity in the nervous system, but anatomic constraints limit how large nerves can become. Therefore a second mechanism, myelination, is invoked to increase conduction velocity.
- **Myelination.** Myelin is a lipid insulator of nerve axons that increases membrane resistance and decreases membrane capacitance. The **increased membrane resistance** forces current to flow along the path of least resistance of the axon interior rather than across the high resistance path of the axonal membrane. The **decreased membrane capacitance** produces a decrease in time constant; thus at breaks in the myelin sheath (see following), the axonal membrane depolarizes faster in response to inward current. Together, the effects of increased membrane resistance and decreased membrane capacitance result in **increased conduction velocity** (Box 1.4).

BOX 1.4 Clinical Physiology: Multiple Sclerosis

DESCRIPTION OF CASE. A 32-year-old woman had her first episode of blurred vision 5 years ago. She had trouble reading the newspaper and the fine print on labels. Her vision returned to normal on its own, but 10 months later, the blurred vision recurred, this time with other symptoms including double vision, and a "pins and needles" feeling and severe weakness in her legs. She was too weak to walk even a single flight of stairs. She was referred to a neurologist, who ordered a series of tests. Magnetic Resonance Imaging (MRI) of the brain showed lesions typical of multiple sclerosis. Visual evoked potentials had a prolonged latency that was consistent with decreased nerve conduction velocity. Since the diagnosis, she has had two relapses and she is currently being treated with interferon beta.

EXPLANATION OF CASE. Action potentials are propagated along nerve fibers by spread of local currents as follows: When an action potential occurs, the inward current of the upstroke of the action potential depolarizes the membrane at that site and reverses the polarity (i.e., that site briefly becomes inside positive). The depolarization then spreads to adjacent sites along the nerve fiber by local current flow. Importantly, if these local currents depolarize an adjacent region to threshold, it will fire an action potential (i.e., the action potential will be propagated). The speed of propagation of the action potential is called conduction velocity. The further local currents can spread without decay (expressed as the length constant), the faster the conduction velocity. There are two main factors that increase length constant and therefore increase conduction velocity in nerves: increased nerve diameter and myelination.

Myelin is an insulator of axons that increases membrane resistance and decreases membrane capacitance. By increasing membrane resistance, current is forced to flow down the axon interior and less current is lost across the cell membrane (increasing length constant); because more current flows down the axon, conduction velocity is increased. By decreasing membrane capacitance, local currents depolarize the membrane more rapidly, which also increases conduction velocity. In order for action potentials to be conducted in myelinated nerves, there must be periodic breaks in the myelin sheath (at the nodes of Ranvier), where there is a concentration of Na^+ and K^+ channels. Thus at the nodes, the ionic currents necessary for the action potential can flow across the membrane (e.g., the inward Na^+ current necessary for the upstroke of the action potential). Between nodes, membrane resistance is very high and current is forced to flow rapidly down the nerve axon to the next node, where the next action potential can be generated. Thus the action potential appears to "jump" from one node of Ranvier to the next. This is called saltatory conduction.

Multiple sclerosis is the most common demyelinating disease of the central nervous system. Loss of the myelin sheath around nerves causes a decrease in membrane resistance, which means that current "leaks out" across the membrane during conduction of local currents. For this reason, local currents decay more rapidly as they flow down the axon (decreased length constant) and, because of this decay, may be insufficient to generate an action potential when they reach the next node of Ranvier.

If the entire nerve were coated with the lipid myelin sheath, however, no action potentials could occur because there would be no low resistance breaks in the membrane across which depolarizing current could flow. Therefore it is important to note that at intervals of 1 to 2 mm, there are breaks in the myelin sheath, at the **nodes of Ranvier.** At the nodes, membrane resistance is low, current can flow across the membrane, and action potentials can occur. Thus conduction of action potentials is faster in myelinated nerves than in unmyelinated nerves because action potentials "jump" long distances from one node to the next, a process called **saltatory conduction.**

SYNAPTIC AND NEUROMUSCULAR TRANSMISSION

A **synapse** is a site where information is transmitted from one cell to another. The information can be transmitted either electrically (electrical synapse) or via a chemical transmitter (chemical synapse).

Types of Synapses

Electrical Synapses

Electrical synapses allow current to flow from one excitable cell to the next via low resistance pathways between the cells called **gap junctions.** Gap junctions are found in cardiac muscle and in some types of smooth muscle and account for the very fast conduction in these tissues. For example, rapid cell-to-cell conduction occurs in cardiac ventricular muscle, in the uterus, and in the bladder, allowing cells in these tissues to be activated simultaneously and ensuring that contraction occurs in a coordinated manner.

Chemical Synapses

In chemical synapses, there is a gap between the presynaptic cell membrane and the postsynaptic cell membrane, known as the **synaptic cleft.** Information is transmitted across the synaptic cleft via a neurotransmitter, a substance that is released from the presynaptic terminal and binds to receptors on the postsynaptic terminal.

The following sequence of events occurs at chemical synapses: An action potential in the presynaptic cell causes Ca^{2+} channels to open. An influx of Ca^{2+} into the presynaptic terminal causes the neurotransmitter, which is stored in synaptic vesicles, to be released by exocytosis. The neurotransmitter diffuses across the synaptic cleft, binds to receptors on the postsynaptic membrane, and produces a change in membrane potential on the postsynaptic cell.

The change in membrane potential on the postsynaptic cell membrane can be either excitatory or inhibitory, depending on the nature of the neurotransmitter released from the presynaptic nerve terminal. If the neurotransmitter is excitatory, it causes depolarization of the postsynaptic cell; if the neurotransmitter is inhibitory, it causes hyperpolarization of the postsynaptic cell.

In contrast to electrical synapses, neurotransmission across chemical synapses is **unidirectional** (from presynaptic cell to postsynaptic cell). The **synaptic delay** is the time required for the multiple steps in chemical neurotransmission to occur.

Neuromuscular Junction—Example of a Chemical Synapse

Motor Units

Motoneurons are the nerves that innervate muscle fibers. A **motor unit** comprises a single motoneuron and the muscle fibers it innervates. Motor units vary considerably in size: A single motoneuron may activate a few muscle fibers or thousands of muscle fibers. Predictably, small motor units are involved in fine motor activities (e.g., facial expressions), and large motor units are involved in gross muscular activities (e.g., quadriceps muscles used in running).

Sequence of Events at the Neuromuscular Junction

The synapse between a motoneuron and a muscle fiber is called the **neuromuscular junction** (Fig. 1.16). An action potential in the motoneuron produces an action potential in the muscle fibers it innervates by the following sequence of events: The numbered steps correlate with the circled numbers in Figure 1.16.

1. Action potentials are propagated down the motoneuron, as described previously. Local currents depolarize each adjacent region to threshold. Finally, the presynaptic terminal is depolarized, and this depolarization causes voltage-gated **Ca^{2+} channels** in the presynaptic membrane to open.
2. When these Ca^{2+} channels open, the Ca^{2+} permeability of the presynaptic terminal increases, and Ca^{2+} flows into the terminal down its electrochemical gradient.
3. Ca^{2+} uptake into the terminal causes release of the neurotransmitter **acetylcholine (ACh),** which has been previously synthesized and stored in synaptic vesicles. To release ACh, the synaptic vesicles fuse with the plasma membrane and empty their contents into the synaptic cleft by exocytosis.

   ACh is formed from acetyl coenzyme A (acetyl CoA) and choline by the action of the enzyme **choline acetyltransferase** (Fig. 1.17). ACh is stored in vesicles with ATP and proteoglycan for subsequent

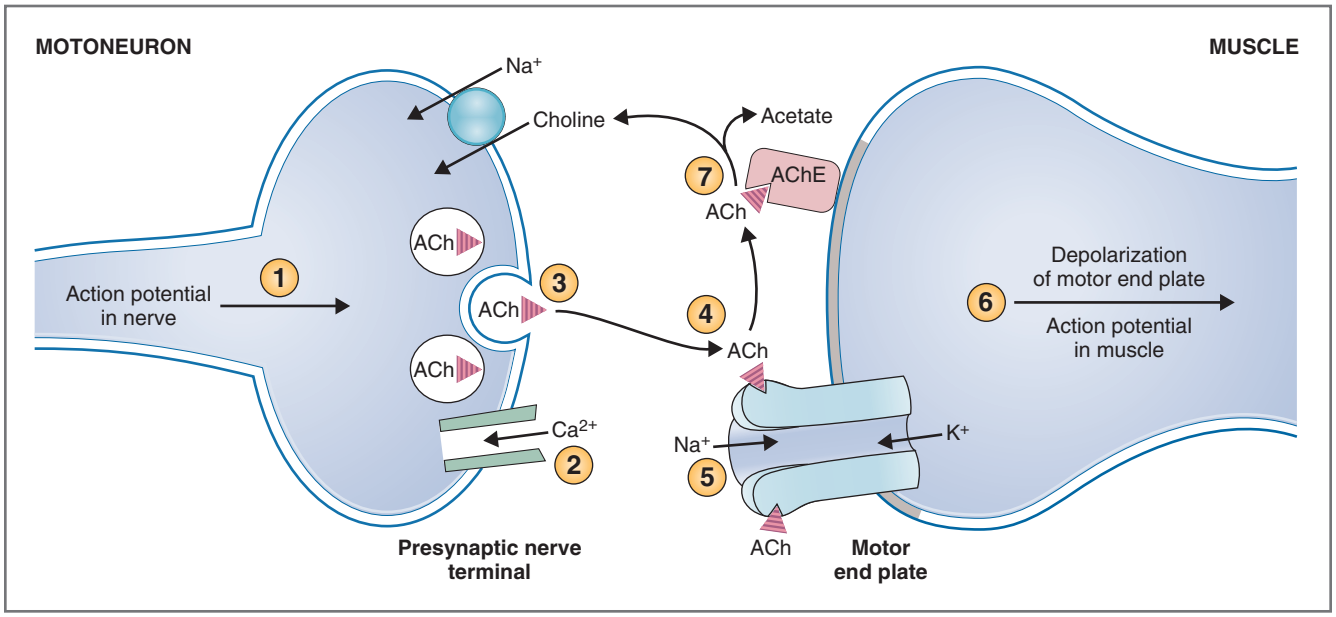


Fig. 1.16 Sequence of events in neuromuscular transmission. *1,* Action potential travels down the motoneuron to the presynaptic terminal. *2,* Depolarization of the presynaptic terminal opens Ca^{2+} channels, and Ca^{2+} flows into the terminal. *3,* Acetylcholine *(ACh)* is extruded into the synapse by exocytosis. *4,* ACh binds to its receptor on the motor end plate. *5,* Channels for Na^+ and K^+ are opened in the motor end plate. *6,* Depolarization of the motor end plate causes action potentials to be generated in the adjacent muscle tissue. *7,* ACh is degraded to choline and acetate by acetylcholinesterase *(AChE);* choline is taken back into the presynaptic terminal on an Na^+-choline cotransporter.

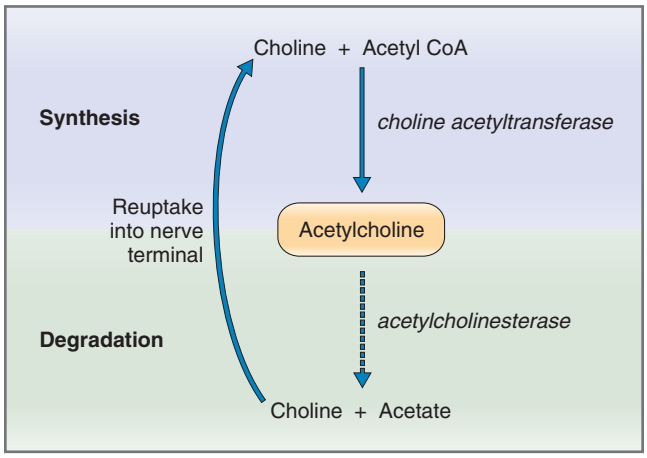


Fig. 1.17 Synthesis and degradation of acetylcholine. *Acetyl CoA,* Acetyl coenzyme A.

release. On stimulation, the entire content of a synaptic vesicle is released into the synaptic cleft. The smallest possible amount of ACh that can be released is the content of one synaptic vesicle (one quantum), and for this reason, the release of ACh is said to be **quantal.**

4. ACh diffuses across the synaptic cleft to the postsynaptic membrane. This specialized region of the muscle fiber is called the **motor end plate,** which contains **nicotinic receptors** for ACh. ACh binds to the α subunits of the nicotinic receptor and causes a conformational change. It is important to note that the nicotinic receptor for ACh is an example of a ligand-gated ion channel: It *also* is an Na^+ and K^+ channel. When the conformational change occurs, the central core of the channel opens, and the permeability of the motor end plate to both Na^+ and K^+ increases.

5. When these channels open, both Na^+ and K^+ flow down their respective electrochemical gradients, Na^+ moving into the end plate and K^+ moving out, each ion attempting to drive the motor **end plate potential (EPP)** to its equilibrium potential. Indeed, if there were no other ion channels in the motor end plate, the end plate would depolarize to a value about halfway between the equilibrium potentials for Na^+ and K^+, or approximately 0 mV. (In this case, zero is not a "magic number"—it simply happens to be the value about halfway between the two equilibrium potentials.) In practice, however, because other ion channels that influence membrane potential are present in the end plate, the motor end plate only depolarizes to about –50 mV, which is the EPP. The EPP is not an action potential but is simply a local depolarization of the specialized motor end plate.

The content of a single synaptic vesicle produces the smallest possible change in membrane potential of the motor end plate, the **miniature end plate**

potential (MEPP). MEPPs summate to produce the full-fledged EPP. The spontaneous appearance of MEPPs proves the quantal nature of ACh release at the neuromuscular junction.

Each MEPP, which represents the content of one synaptic vesicle, depolarizes the motor end plate by about 0.4 mV. An EPP is a multiple of these 0.4-mV units of depolarization. *How many such quanta are required to depolarize the motor end plate to the EPP?* Because the motor end plate must be depolarized from its resting potential of –90 mV to the threshold potential of –50 mV, it must therefore depolarize by 40 mV. Depolarization by 40 mV requires 100 quanta (because each quantum or vesicle depolarizes the motor end plate by 0.4 mV).

6. Depolarization of the motor end plate (the EPP) then spreads by local currents to adjacent muscle fibers, which are depolarized to threshold and fire action potentials. Although the motor end plate itself cannot fire action potentials, it depolarizes sufficiently to initiate the process in the neighboring "regular" muscle cell membranes. Action potentials are propagated down the muscle fiber by a continuation of this process.
7. The EPP at the motor end plate is terminated when ACh is degraded to choline and acetate by **acetylcholinesterase (AChE)** on the motor end plate. Approximately 50% of the choline is returned to the presynaptic terminal by **Na^+-choline cotransport,** to be used again in the synthesis of new ACh.

Agents That Alter Neuromuscular Function

Several agents interfere with normal activity at the neuromuscular junction, and their mechanisms of action can be readily understood by considering the steps involved in neuromuscular transmission (Table 1.3; see Fig. 1.16).

- **Botulinus toxin** blocks the release of ACh from presynaptic terminals, causing total blockade of neuromuscular transmission, paralysis of skeletal muscle, and, eventually, death from respiratory failure.
- **Curare** competes with ACh for the nicotinic receptors on the motor end plate, decreasing the size of the EPP. When administered in maximal doses, curare causes paralysis and death. **D-Tubocurarine,** a form of curare, is used therapeutically to cause relaxation of skeletal muscle during anesthesia. A related substance, **α-bungarotoxin,** binds irreversibly to ACh receptors. Binding of radioactive α-bungarotoxin has provided an experimental tool for measuring the density of ACh receptors on the motor end plate.
- **AChE inhibitors** (anticholinesterases) such as **neostigmine** prevent degradation of ACh in the synaptic cleft, and they prolong and enhance the action of ACh at the motor end plate. AChE inhibitors can be used in the treatment of **myasthenia gravis,** a disease characterized by skeletal muscle weakness and fatigability, in which ACh receptors are blocked by antibodies (Box 1.5).
- **Hemicholinium** blocks choline reuptake into presynaptic terminals, thus depleting choline stores from the motoneuron terminal and decreasing the synthesis of ACh.

Types of Synaptic Arrangements

There are several types of relationships between the input to a synapse (the presynaptic element) and the output (the postsynaptic element): one-to-one, one-to-many, or many-to-one.

- **One-to-one synapses.** The one-to-one synapse is illustrated by the **neuromuscular junction** (see Fig. 1.16). A single action potential in the presynaptic cell, the motoneuron, causes a single action potential in the postsynaptic cell, the muscle fiber.
- **One-to-many synapses.** The one-to-many synapse is uncommon, but it is found, for example, at the

TABLE 1.3 Agents Affecting Neuromuscular Transmission

Example	Action	Effect on Neuromuscular Transmission
Botulinus toxin	Blocks ACh release from presynaptic terminals	Total blockade, paralysis of respiratory muscles, and death
Curare	Competes with ACh for receptors on motor end plate	Decreases size of EPP; in maximal doses produces paralysis of respiratory muscles and death
Neostigmine	AChE inhibitor (anticholinesterase)	Prolongs and enhances action of ACh at motor end plate
Hemicholinium	Blocks reuptake of choline into presynaptic terminal	Depletes ACh stores from presynaptic terminal

ACh, Acetylcholine; *AChE,* acetylcholinesterase; *EPP,* end plate potential.

BOX 1.5 Clinical Physiology: Myasthenia Gravis

DESCRIPTION OF CASE. An 18-year-old college woman comes to the student health service complaining of progressive weakness. She reports that occasionally her eyelids "droop" and that she tires easily, even when completing ordinary daily tasks such as brushing her hair. She has fallen several times while climbing a flight of stairs. These symptoms improve with rest. The physician orders blood studies, which reveal elevated levels of antibodies to ACh receptors. Nerve stimulation studies show decreased responsiveness of skeletal muscle on repeated stimulation of motoneurons. The woman is diagnosed with myasthenia gravis and is treated with the drug pyridostigmine. After treatment, she reports a return of muscle strength.

EXPLANATION OF CASE. This young woman has classic myasthenia gravis. In the autoimmune form of the disease, antibodies are produced to ACh receptors on the motor end plates of skeletal muscle. Her symptoms of severe muscle weakness (eye muscles; arms and legs) are explainable by the presence of antibodies that block ACh receptors. Although ACh is released in normal amounts from the terminals of motoneurons, binding of ACh to its receptors on the motor end plates is impaired. Because ACh cannot bind, depolarization of the motor end plate (EPP) will not occur and normal action potentials cannot be generated in the skeletal muscle. Muscle weakness and fatigability ensue.

TREATMENT. Treatment of the patient with myasthenia gravis depends on a clear understanding of the physiology of the neuromuscular junction. Because this patient's condition improved with the administration of pyridostigmine (a long-acting AChE inhibitor), the success of the treatment confirmed the diagnosis of myasthenia gravis. AChE on the motor end plate normally degrades ACh (i.e., AChE terminates the action of ACh). By inhibiting the ACh-degradative enzyme with pyridostigmine, ACh levels in the neuromuscular junction are maintained at a high level, prolonging the time available for ACh to activate its receptors on the motor end plate. Thus a more normal EPP in the muscle fiber can be produced even though many of the ACh receptors are blocked by antibodies.

synapses of motoneurons on Renshaw cells of the spinal cord. An action potential in the presynaptic cell, the motoneuron, causes a burst of action potentials in the postsynaptic cells. This arrangement causes amplification of activity.

- **Many-to-one synapses.** The many-to-one synapse is a very common arrangement in the nervous system. In these synapses, an action potential in the presynaptic cell is insufficient to produce an action potential in the postsynaptic cell. Instead, many presynaptic cells converge on the postsynaptic cell, these inputs summate, and the sum of the inputs determines whether the postsynaptic cell will fire an action potential.

Synaptic Input—Excitatory and Inhibitory Postsynaptic Potentials

The many-to-one synaptic arrangement is a common configuration in which many presynaptic cells converge on a single postsynaptic cell, with the inputs being either **excitatory** or **inhibitory.** The postsynaptic cell integrates all the converging information, and if the sum of the inputs is sufficient to bring the postsynaptic cell to threshold, it will then fire an action potential.

Excitatory Postsynaptic Potentials

Excitatory postsynaptic potentials (EPSPs) are synaptic inputs that **depolarize** the postsynaptic cell, bringing the membrane potential closer to threshold and closer to firing an action potential. EPSPs are produced by **opening Na^+ and K^+ channels,** similar to the nicotinic ACh receptor. The membrane potential is driven to a value approximately halfway between the equilibrium potentials for Na^+ and K^+, or 0 mV, which is a depolarized state. Excitatory neurotransmitters include ACh, norepinephrine, epinephrine, dopamine, glutamate, and serotonin.

Inhibitory Postsynaptic Potentials

Inhibitory postsynaptic potentials (IPSPs) are synaptic inputs that **hyperpolarize** the postsynaptic cell, taking the membrane potential away from threshold and farther from firing an action potential. IPSPs are produced by **opening Cl^- channels.** The membrane potential is driven toward the Cl^- equilibrium potential (approximately –90 mV), which is a hyperpolarized state. Inhibitory neurotransmitters are γ-aminobutyric acid (GABA) and glycine.

Integration of Synaptic Information

The presynaptic information that arrives at the synapse may be integrated in one of two ways, spatially or temporally.

Spatial Summation

Spatial summation occurs when two or more presynaptic inputs arrive at a postsynaptic cell simultaneously. If both inputs are excitatory, they will combine to produce greater depolarization than either input would produce separately. If one input is excitatory and the other is inhibitory, they will cancel each other out. Spatial summation may occur, even if the inputs are far

apart on the nerve cell body, because EPSPs and IPSPs are conducted so rapidly over the cell membrane.

Temporal Summation

Temporal summation occurs when two presynaptic inputs arrive at the postsynaptic cell in rapid succession. Because the inputs overlap in time, they summate.

Other Phenomena That Alter Synaptic Activity

Facilitation, augmentation, and **post-tetanic potentiation** are phenomena that may occur at synapses. In each instance, repeated stimulation causes the response of the postsynaptic cell to be greater than expected. The common underlying mechanism is believed to be an increased release of neurotransmitter into the synapse, possibly caused by accumulation of Ca^{2+} in the presynaptic terminal. **Long-term potentiation** occurs in storage of memories and involves both increased release of neurotransmitter from presynaptic terminals and increased sensitivity of postsynaptic membranes to the transmitter.

Synaptic fatigue may occur where repeated stimulation produces a smaller than expected response in the postsynaptic cell, possibly resulting from the depletion of neurotransmitter stores from the presynaptic terminal.

Neurotransmitters

The transmission of information at chemical synapses involves the release of a neurotransmitter from a presynaptic cell, diffusion across the synaptic cleft, and binding of the neurotransmitter to specific receptors on the postsynaptic membrane to produce a change in membrane potential.

The following criteria are used to formally designate a substance as a neurotransmitter: The substance must be synthesized in the presynaptic cell; the substance must be released by the presynaptic cell on stimulation; and, if the substance is applied exogenously to the postsynaptic membrane at physiologic concentration, the response of the postsynaptic cell must mimic the in vivo response.

Neurotransmitter substances can be grouped into the following categories: ACh, biogenic amines, amino acids, and neuropeptides (Table 1.4).

Acetylcholine

The role of ACh as a neurotransmitter is vitally important for several reasons. ACh is the *only* neurotransmitter that is utilized at the neuromuscular junction. It is the neurotransmitter released from *all* preganglionic and most postganglionic neurons in the parasympathetic

TABLE 1.4 Classification of Neurotransmitter Substances

Choline Esters	Biogenic Amines	Amino Acids	Neuropeptides
Acetylcholine (ACh)	Dopamine	γ-Aminobutyric acid (GABA)	Adrenocorticotropin (ACTH)
	Epinephrine	Glutamate	Cholecystokinin
	Histamine	Glycine	Dynorphin
	Norepinephrine		Endorphins
	Serotonin		Enkephalins
			Gastrin-releasing peptide (GRP)
			Glucose-dependent insulinotropic peptide (GIP)
			Glucagon
			Neurophysin II
			Neurotensin
			Oxytocin
			Secretin
			Somatostatin
			Substance P
			Thyrotropin-releasing hormone (TRH)
			Vasopressin, or antidiuretic hormone (ADH)
			Vasoactive intestinal peptide (VIP)

nervous system and from *all* preganglionic neurons in the sympathetic nervous system. It is also the neurotransmitter that is released from presynaptic neurons of the adrenal medulla.

Figure 1.17 illustrates the synthetic and degradative pathways for ACh. In the presynaptic terminal, choline and acetyl CoA combine to form ACh, catalyzed by choline acetyltransferase. When ACh is released from the presynaptic nerve terminal, it diffuses to the postsynaptic membrane, where it binds to and activates nicotinic ACh receptors. AChE is present on the postsynaptic membrane, where it degrades ACh to choline and acetate. This degradation terminates the action of ACh at the postsynaptic membrane. Approximately one-half of the choline that is released from the degradation of ACh is taken back into the presynaptic terminal to be reutilized for synthesis of new ACh.

Norepinephrine, Epinephrine, and Dopamine

Norepinephrine, epinephrine, and dopamine are members of the same family of biogenic amines: They share a common precursor, tyrosine, and a common biosynthetic pathway (Fig. 1.18). Tyrosine is converted to **L-dopa** by tyrosine hydroxylase, and L-dopa is converted to **dopamine** by dopa decarboxylase. If dopamine β-hydroxylase is present in small dense-core vesicles of

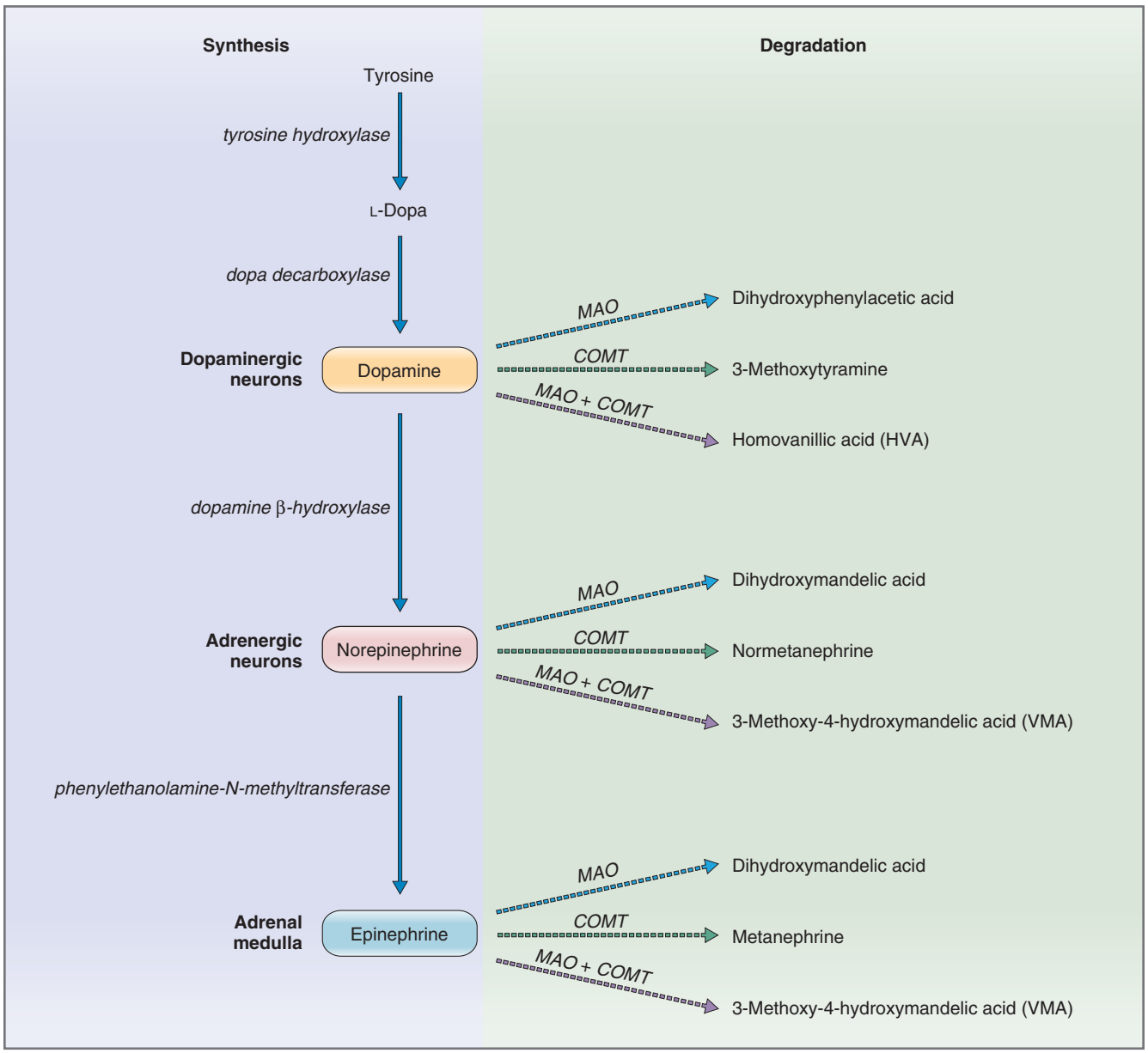


Fig. 1.18 Synthesis and degradation of dopamine, norepinephrine, and epinephrine.
COMT, Catechol-*O*-methyltransferase; *MAO,* monoamine oxidase.

the nerve terminal, dopamine is converted to **norepinephrine.** If phenylethanolamine-*N*-methyl transferase (PNMT) is present (with *S*-adenosylmethionine as the methyl donor), then norepinephrine is methylated to form **epinephrine.**

The specific neurotransmitter secreted depends on which portion, or portions, of the enzymatic pathway are present in a particular type of nerve or gland. Thus **dopaminergic neurons** secrete dopamine because the presynaptic nerve terminal contains tyrosine hydroxylase and dopa decarboxylase but not the other enzymes. **Adrenergic neurons** secrete norepinephrine because they contain dopamine β-hydroxylase, in addition to tyrosine hydroxylase and dopa decarboxylase, but not PNMT. The **adrenal medulla** contains the complete enzymatic pathway; therefore it secretes primarily epinephrine.

The degradation of dopamine, norepinephrine, and epinephrine to inactive substances occurs via two enzymes: catechol-*O*-methyltransferase (COMT) and monoamine oxidase (MAO). **COMT,** a methylating enzyme, is *not found in nerve terminals,* but it is distributed widely in other tissues including the liver. **MAO** is located in presynaptic nerve terminals and catalyzes oxidative deamination. If a neurotransmitter is to be degraded by MAO, there must be reuptake of the neurotransmitter from the synapse.

Each of the biogenic amines can be degraded by MAO alone, by COMT alone, or by both MAO and COMT (in any order). Thus there are three possible degradative products from each neurotransmitter, and typically these products are excreted in the urine (see Fig. 1.8). The major metabolite of norepinephrine is **normetanephrine.** The major metabolite of epinephrine is **metanephrine.** Both norepinephrine and epinephrine are degraded to **3-methoxy-4-hydroxymandelic acid (VMA).**

Serotonin

Serotonin, another biogenic amine, is produced from tryptophan in serotonergic neurons in the brain and in the gastrointestinal tract (Fig. 1.19). Following its release from presynaptic neurons, serotonin may be returned intact to the nerve terminal, or it may be degraded in the presynaptic terminal by MAO to 5-hydroxyindoleacetic acid. Additionally, serotonin serves as the precursor to melatonin in the pineal gland.

Histamine

Histamine, a biogenic amine, is synthesized from histidine, catalyzed by histidine decarboxylase. It is present in neurons of the hypothalamus, as well as in nonneural tissue such as **mast cells** of the gastrointestinal tract.

Glutamate

Glutamate, an amino acid, is the major **excitatory** neurotransmitter in the central nervous system. It plays a significant role in the spinal cord and cerebellum. There are four subtypes of glutamate receptors. Three of the subtypes are **ionotropic receptors,** or ligand-gated ion channels including the **NMDA** (*N*-methyl-D-aspartate) receptor that is widely distributed throughout the central nervous system. A fourth subtype comprises **metabotropic receptors,** which are coupled via heterotrimeric guanosine triphosphate (GTP)–binding proteins (G proteins) to ion channels.

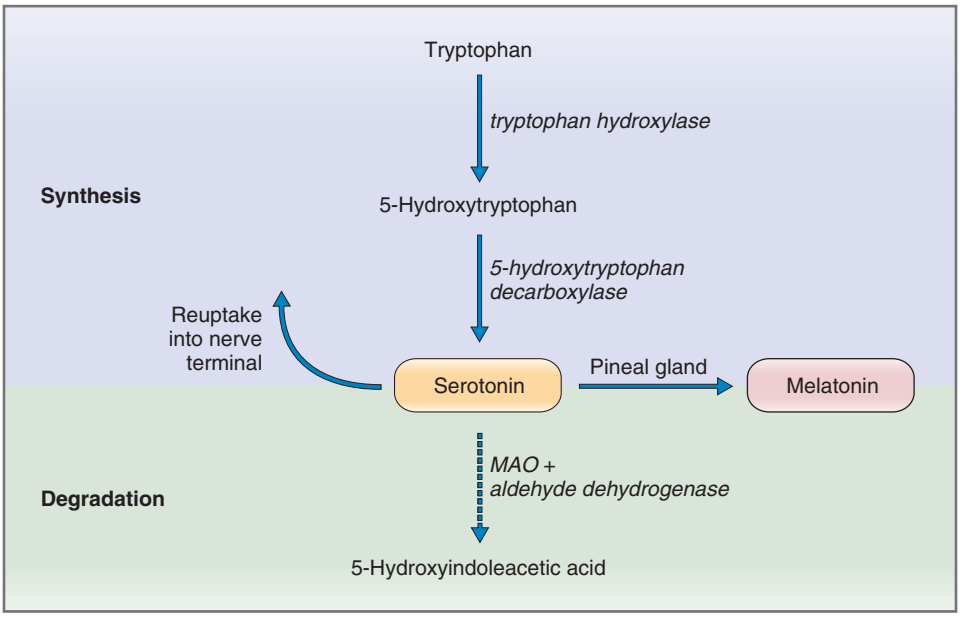


Fig. 1.19 Synthesis and degradation of serotonin. *MAO,* Monoamine oxidase.

Glycine

Glycine, an amino acid, is an **inhibitory** neurotransmitter that is found in the spinal cord and brain stem. Its mechanism of action is to **increase Cl^- conductance** of the postsynaptic cell membrane. By increasing Cl^- conductance, the membrane potential is driven closer to the Cl^- equilibrium potential. Thus the postsynaptic cell membrane is hyperpolarized or inhibited.

γ-Aminobutyric Acid (GABA)

GABA is an amino acid and an **inhibitory** neurotransmitter that is distributed widely in the central nervous system in GABAergic neurons. GABA is synthesized from glutamic acid, catalyzed by glutamic acid decarboxylase, an enzyme that is unique to GABAergic neurons (Fig. 1.20). Following its release from presynaptic nerves and its action at the postsynaptic cell membrane, GABA can be either recycled back to the presynaptic terminal or degraded by GABA transaminase to enter the citric acid cycle. Unlike the other amino acids that serve as neurotransmitters (e.g., glutamate and glycine), GABA does not have any metabolic functions (i.e., it is not incorporated into proteins).

The two types of GABA receptors on postsynaptic membranes are the $GABA_A$ and the $GABA_B$ receptors. The **$GABA_A$ receptor** is directly linked to a Cl^- channel and thus is **ionotropic.** When stimulated, it increases Cl^- conductance and thus hyperpolarizes (inhibits) the postsynaptic cell. The $GABA_A$ receptor is the site of action of **benzodiazepines** and **barbiturates** in the central nervous system. The **$GABA_B$ receptor** is coupled via a G protein to a K^+ channel and thus is **metabotropic.** When stimulated, it increases K^+ conductance and hyperpolarizes the postsynaptic cell.

Huntington disease is associated with GABA deficiency. The disease is characterized by hyperkinetic choreiform movements related to a deficiency of GABA in the projections from the striatum to the globus pallidus. The characteristic uncontrolled movements are, in part, attributed to lack of GABA-dependent inhibition of neural pathways.

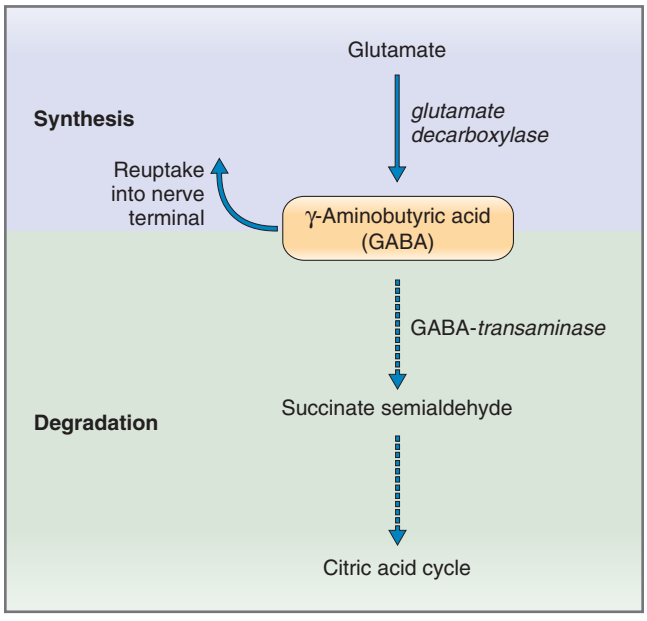


Fig. 1.20 Synthesis and degradation of γ-aminobutyric acid *(GABA).*

Nitric Oxide

Nitric oxide (NO) is a short-acting inhibitory neurotransmitter in the gastrointestinal tract and the central nervous system. In presynaptic nerve terminals, the enzyme **NO synthase** converts arginine to citrulline and NO. Then, NO, a permeant gas, simply diffuses from the presynaptic terminal to its target cell (instead of the usual packaging of neurotransmitter in synaptic vesicles and release by exocytosis). In addition to serving as a neurotransmitter, NO also functions in signal transduction of guanylyl cyclase in a variety of tissues including vascular smooth muscle (see Chapter 4).

Neuropeptides

There is a long and growing list of neuropeptides that function as neuromodulators, neurohormones, and neurotransmitters (see Table 1.4 for a partial list).

- **Neuromodulators** are substances that act on the presynaptic cell to alter the amount of neurotransmitter released in response to stimulation. Alternatively, a neuromodulator may be cosecreted with a neurotransmitter and alter the response of the postsynaptic cell to the neurotransmitter.
- **Neurohormones,** like other hormones, are released from secretory cells (in these cases, neurons) into the blood to act at a distant site.
- In several instances, **neuropeptides** are copackaged and cosecreted from presynaptic vesicles along with the classical neurotransmitters. For example, vasoactive intestinal peptide (VIP) is stored and secreted with ACh, particularly in neurons of the gastrointestinal tract. Somatostatin, enkephalin, and neurotensin are secreted with norepinephrine. Substance P is secreted with serotonin.

In contrast to classical neurotransmitters, which are synthesized in presynaptic nerve terminals, neuropeptides are synthesized in the nerve cell body. As occurs in all protein synthesis, the cell's DNA is transcribed into specific messenger RNA, which is translated into polypeptides on the ribosomes. Typically, a preliminary polypeptide containing a signal peptide sequence is synthesized first. The signal peptide is removed in the endoplasmic reticulum, and the final peptide is

delivered to secretory vesicles. The secretory vesicles are then moved rapidly down the nerve by **axonal transport** to the presynaptic terminal, where they become the synaptic vesicles.

Purines

ATP and adenosine function as neuromodulators in the autonomic and central nervous systems. For example, **ATP** is synthesized in the sympathetic neurons that innervate vascular smooth muscle. It is costored and cosecreted with the "regular" neurotransmitter of these neurons, norepinephrine. When stimulated, the neuron releases both ATP and norepinephrine and both transmitters cause contraction of the smooth muscle; in fact, the ATP-induced contraction precedes the norepinephrine-induced contraction.

SKELETAL MUSCLE

Contraction of skeletal muscle is under voluntary or reflex control. Each skeletal muscle cell is innervated by a branch of a motoneuron. Action potentials are propagated along the motoneurons, leading to release of ACh at the neuromuscular junction, depolarization of the motor end plate, and initiation of action potentials in the muscle fiber.

What events, then, elicit contraction of the muscle fiber? These events, occurring between the action potential in the muscle fiber and contraction of the muscle fiber, are called **excitation-contraction coupling.** The mechanisms of excitation-contraction coupling in skeletal muscle and smooth muscle are discussed in this chapter, and the mechanisms of excitation-contraction coupling in cardiac muscle are discussed in Chapter 4.

Muscle Filaments

Each muscle fiber behaves as a single unit, is multinucleate, and contains myofibrils. The myofibrils are surrounded by SR and are invaginated by transverse tubules (T tubules). Each myofibril contains interdigitating thick and thin filaments, which are arranged longitudinally and cross-sectionally in sarcomeres (Fig. 1.21). The repeating units of sarcomeres account for the unique banding pattern seen in striated muscle (which includes both skeletal and cardiac muscle).

Thick Filaments

The thick filaments comprise a large molecular weight protein called **myosin,** which has six polypeptide chains including one pair of **heavy chains** and two pairs of **light chains** (see Fig. 1.21A). Most of the heavy-chain myosin has an α-helical structure, in which the two chains coil around each other to form the "**tail**" of the myosin molecule. The four light chains and the N terminus of each heavy chain form two globular "**heads**" on the myosin molecule. These globular heads have an actin-binding site, which is necessary for cross-bridge formation, and a site that binds and hydrolyzes ATP (myosin ATPase).

Thin Filaments

The thin filaments are composed of three proteins: actin, tropomyosin, and troponin (see Fig. 1.21B).

Actin is a globular protein and, in this globular form, is called G-actin. In the thin filaments, G-actin is polymerized into two strands that are twisted into an α-helical structure to form filamentous actin, called F-actin. Actin has myosin-binding sites. When the muscle is at rest, the myosin-binding sites are covered by tropomyosin so that actin and myosin cannot interact.

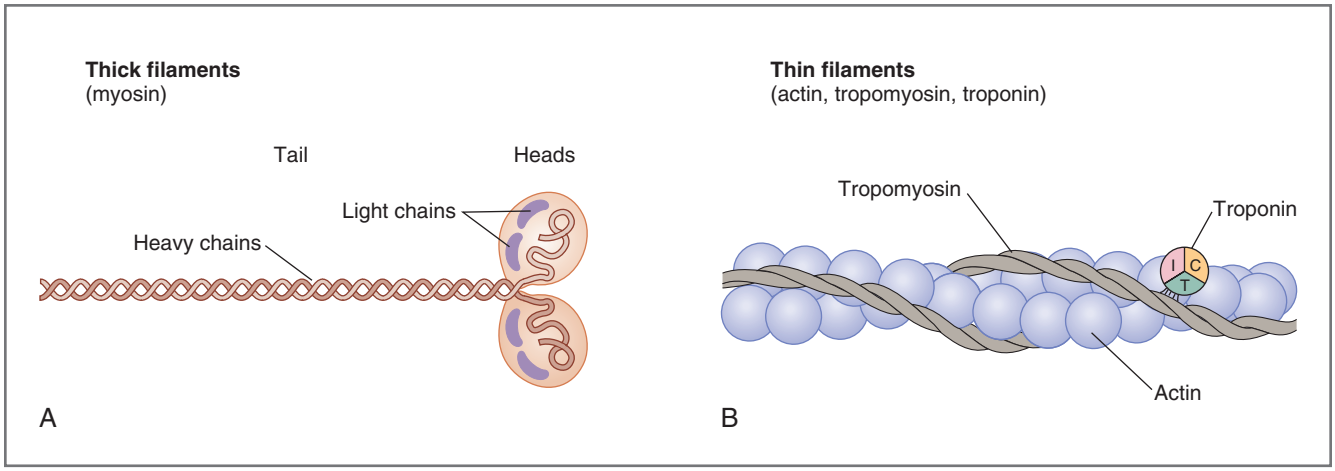


Fig. 1.21 Structure of thick (A) and thin (B) filaments of skeletal muscle. Troponin is a complex of three proteins: *I,* Troponin I; *T,* troponin T; and *C,* troponin C.

Tropomyosin is a filamentous protein that runs along the groove of each twisted actin filament. At rest, its function is to block the myosin-binding sites on actin. If contraction is to occur, tropomyosin must be moved out of the way so that actin and myosin can interact.

Troponin is a complex of three globular proteins (troponin T, troponin I, and troponin C) located at regular intervals along the tropomyosin filaments. **Troponin T** (T for tropomyosin) attaches the troponin complex to tropomyosin. **Troponin I** (I for inhibition), along with tropomyosin, inhibits the interaction of actin and myosin by covering the myosin-binding site on actin. **Troponin C** (C for Ca^{2+}) is a Ca^{2+}-binding protein that plays a central role in the initiation of contraction. When the intracellular Ca^{2+} concentration increases, Ca^{2+} binds to troponin C, producing a conformational change in the troponin complex. This conformational change moves tropomyosin out of the way, permitting the binding of actin to the myosin heads.

Arrangement of Thick and Thin Filaments in Sarcomeres

The **sarcomere** is the basic contractile unit, and it is delineated by the Z disks. Each sarcomere contains a full A band in the center and one-half of two I bands on either side of the A band (Fig. 1.22).

The **A bands** are located in the center of the sarcomere and contain the thick (myosin) filaments, which appear dark when viewed under polarized light. Thick and thin filaments may overlap in the A band; these areas of overlap are potential sites of cross-bridge formation.

The **I bands** are located on either side of the A band and appear light when viewed under polarized light. They contain the thin (actin) filaments, intermediate filamentous proteins, and Z disks. They have no thick filaments.

The **Z disks** are darkly staining structures that run down the middle of each I band, delineating the ends of each sarcomere.

The **bare zone** is located in the center of each sarcomere. There are no thin filaments in the bare zone; thus there can be no overlap of thick and thin filaments or cross-bridge formation in this region.

The **M line** bisects the bare zone and contains darkly staining proteins that link the central portions of the thick filaments together.

Cytoskeletal Proteins

Cytoskeletal proteins establish the architecture of the myofibrils, ensuring that the thick and thin filaments are aligned correctly and at proper distances with respect to each other.

Transverse cytoskeletal proteins link thick and thin filaments, forming a "scaffold" for the myofibrils and linking sarcomeres of adjacent myofibrils. A system of intermediate filaments holds the myofibrils together, side by side. The entire myofibrillar array is anchored to the cell membrane by an actin-binding protein called **dystrophin.** (In patients with muscular dystrophy, dystrophin is defective or absent.)

Longitudinal cytoskeletal proteins include two large proteins called titin and nebulin. **Titin,** which is associated with thick filaments, is a large molecular weight protein that extends from the M lines to the Z disks. Part of the titin molecule passes through the thick

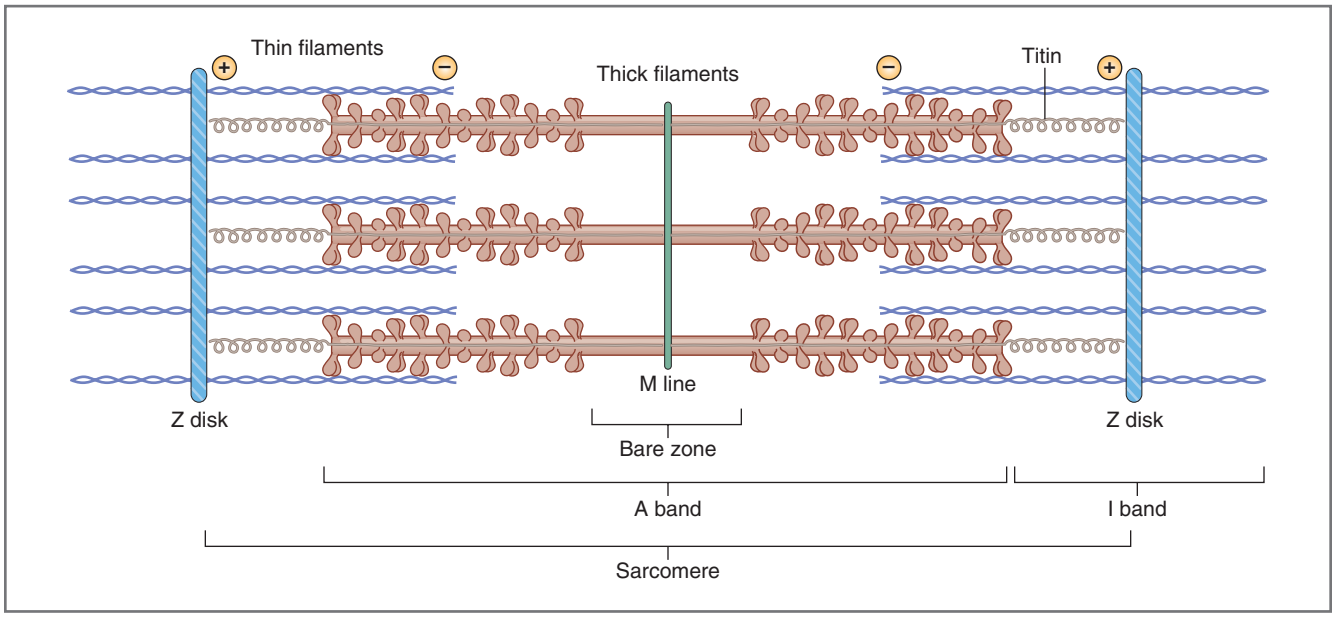


Fig. 1.22 **Arrangement of thick and thin filaments of skeletal muscle in sarcomeres.**

filament; the rest of the molecule, which is elastic or springlike, is anchored to the Z disk. As the length of the sarcomere changes, so does the elastic portion of the titin molecule. Titin also helps center the thick filaments in the sarcomere. **Nebulin** is associated with thin filaments. A single nebulin molecule extends from one end of the thin filament to the other. Nebulin serves as a "molecular ruler," setting the length of thin filaments during their assembly. **α-Actinin** anchors the thin filaments to the Z disk.

Transverse Tubules and the Sarcoplasmic Reticulum

The **transverse (T) tubules** are an extensive network of muscle cell membrane (sarcolemmal membrane) that invaginates deep into the muscle fiber. The T tubules are responsible for carrying depolarization from action potentials at the muscle cell surface to the interior of the fiber. The T tubules make contact with the terminal cisternae of the SR and contain a voltage-sensitive protein called the **dihydropyridine receptor,** named for the drug that inhibits it (Fig. 1.23).

The **sarcoplasmic reticulum (SR)** is an internal tubular structure, which is the site of storage and release of Ca^{2+} for excitation-contraction coupling. As previously noted, the terminal cisternae of the SR make contact with the T tubules in a triad arrangement. The SR contains a Ca^{2+}-release channel called the **ryanodine receptor** (named for the plant alkaloid that opens this release channel). The significance of the physical relationship between the T tubules (and their dihydropyridine receptor) and the SR (and its ryanodine receptor) is described in the section on excitation-contraction coupling.

Ca^{2+} is accumulated in the SR by the action of **Ca^{2+} ATPase (SERCA)** in the SR membrane. The Ca^{2+} ATPase pumps Ca^{2+} from the ICF of the muscle fiber into the interior of the SR, keeping the intracellular Ca^{2+} concentration low when the muscle fiber is at rest. Within the SR, Ca^{2+} is bound to **calsequestrin,** a low-affinity, high-capacity Ca^{2+}-binding protein. Calsequestrin, by binding Ca^{2+}, helps to maintain a low free Ca^{2+} concentration inside the SR, thereby reducing the work of the Ca^{2+} ATPase pump. Thus a large quantity of Ca^{2+} can be stored inside the SR in *bound* form, while the intrasarcoplasmic reticulum *free* Ca^{2+} concentration remains extremely low.

Excitation-Contraction Coupling in Skeletal Muscle

The mechanism that translates the muscle action potential into the production of tension is excitation-contraction coupling. Figure 1.24 shows the temporal relationships between an action potential in the skeletal muscle fiber, the subsequent increase in intracellular free Ca^{2+} concentration (which is released from the SR), and contraction of the muscle fiber. These temporal relationships are critical in that the action potential always *precedes* the rise in intracellular Ca^{2+} concentration, which always *precedes* contraction.

The steps involved in excitation-contraction coupling are described as follows and illustrated in Figure 1.25 (Step 6 is illustrated in Fig. 1.26):

1. **Action potentials** in the muscle cell membrane are propagated to the **T tubules** by the spread of local currents. Thus the T tubules are continuous with the sarcolemmal membrane and carry the depolarization from the surface to the interior of the muscle fiber.

2a. and b. **Depolarization of the T tubules** causes a critical conformational change in their voltage-sensitive **dihydropyridine receptors.** This conformational

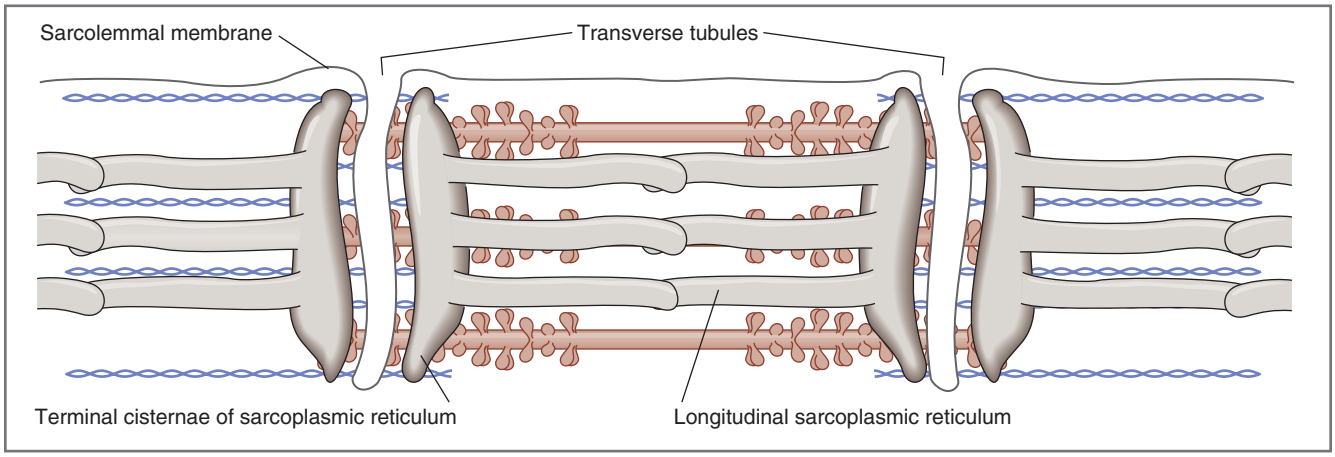


Fig. 1.23 Transverse tubules and sarcoplasmic reticulum (SR) of skeletal muscle. The transverse tubules are continuous with the sarcolemmal membrane and invaginate deep into the muscle fiber, making contact with terminal cisternae of the SR.

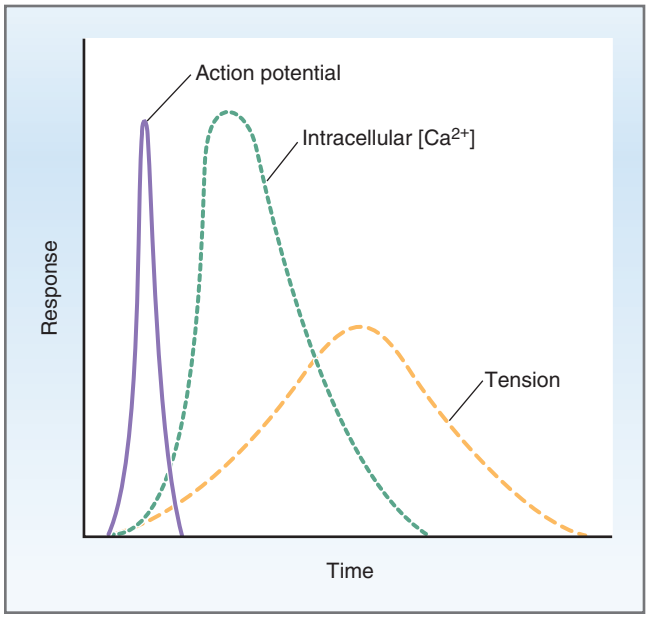


Fig. 1.24 Temporal sequence of events in excitation-contraction coupling in skeletal muscle. The muscle action potential precedes a rise in intracellular [Ca^{2+}], which precedes contraction.

change opens Ca^{2+}-release channels **(ryanodine receptors)** on the nearby SR. (As an aside, although the T tubules' dihydropyridine receptors are L-type voltage-gated Ca^{2+} channels, Ca^{2+} influx into the cell through these channels is *not* required for excitation-contraction coupling in skeletal muscle.)

3. When these Ca^{2+}-release channels open, Ca^{2+} is released from its storage site in the SR into the ICF of the muscle fiber, resulting in an **increase in intracellular Ca^{2+} concentration.** At rest, the intracellular free Ca^{2+} concentration is less than 10^{-7} M. After its release from the SR, intracellular free Ca^{2+} concentration increases to levels between 10^{-7} M and 10^{-6} M.
4. **Ca^{2+} binds to troponin C** on the thin filaments, causing a conformational change in the troponin complex. Troponin C can bind as many as four Ca^{2+} ions per molecule of protein. Because this binding is cooperative, each molecule of bound Ca^{2+} increases the affinity of troponin C for the next Ca^{2+}. Thus even a small increase in Ca^{2+} concentration increases the likelihood that all of the binding sites will be occupied to produce the necessary conformational change in the troponin complex.
5. The **conformational change in troponin** causes tropomyosin (which was previously blocking the interaction of actin and myosin) to be moved out of the way so that cross-bridge cycling can begin. When tropomyosin is moved away, the myosin-binding sites on actin, previously covered, are exposed.

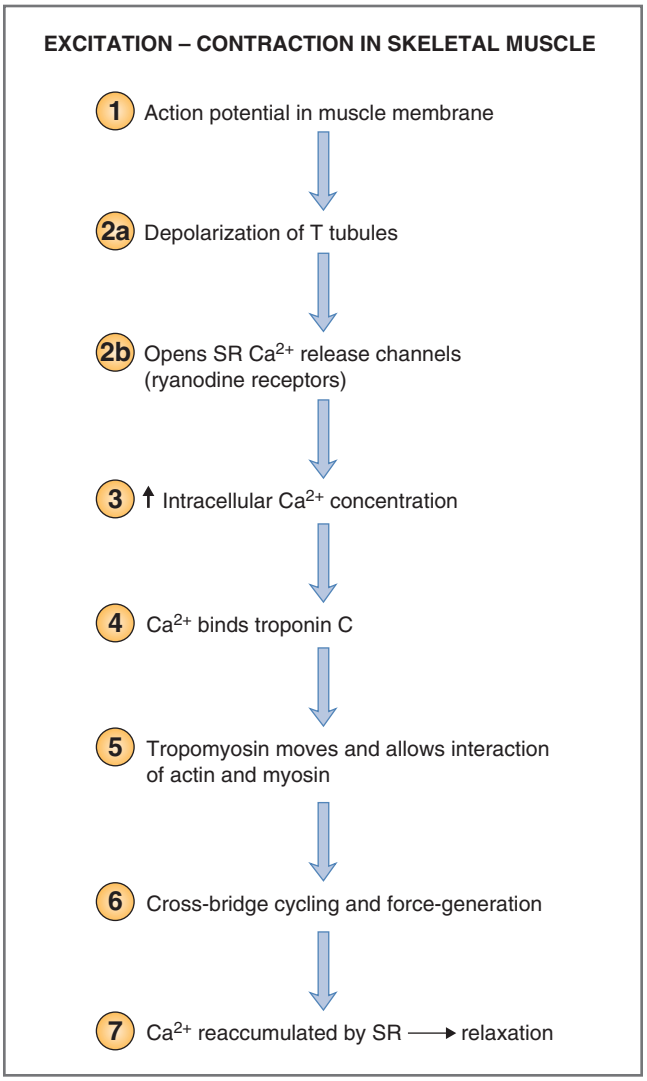


Fig. 1.25 Steps in excitation-contraction in skeletal muscle. *SR,* Sarcoplasmic reticulum; *T tubules,* transverse tubules. See text for explanation of the circled numbers.

6. **Cross-bridge cycling.** With Ca^{2+} bound to troponin C and tropomyosin moved out of the way, myosin heads can now bind to actin and form so-called **cross-bridges.** Formation of cross-bridges is associated with hydrolysis of ATP and generation of force.

The sequence of events in the cross-bridge cycle is shown in Figure 1.26. *A,* At the beginning of the cycle, no ATP is bound to myosin, and myosin is tightly attached to actin in a "rigor" position. In rapidly contracting muscle, this state is brief. However, in the absence of ATP, this state is permanent (i.e., rigor mortis). *B,* The binding of ATP to a cleft on the back of the myosin head produces a conformational change in myosin that decreases its affinity for actin; thus myosin is released from the original actin-binding site. *C,* The cleft closes around the bound ATP molecule, producing a further

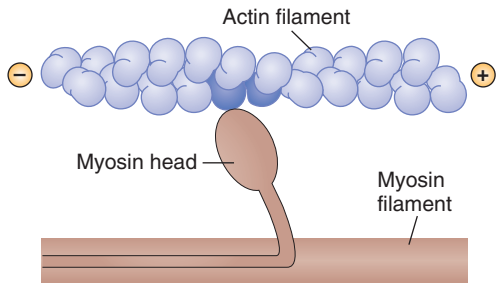
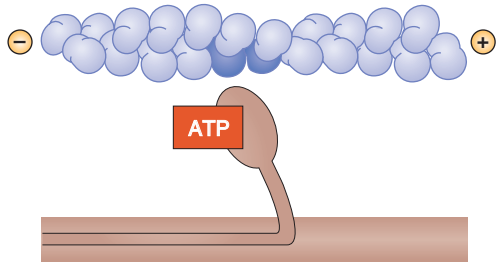
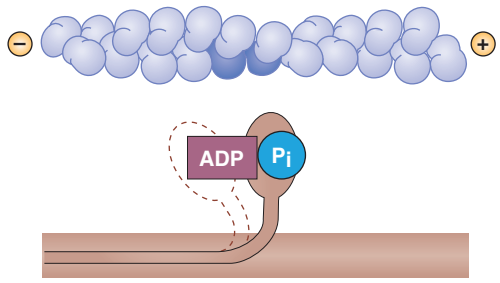
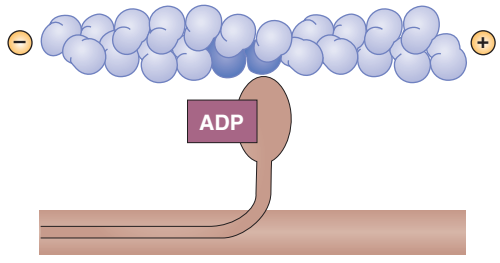
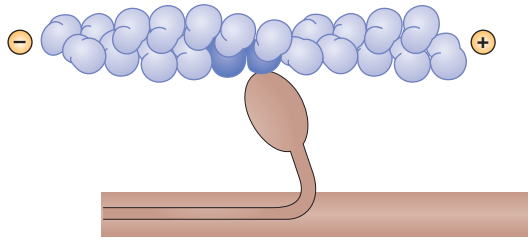
Position of Actin and Myosin During Cross-Bridge Cycling	Events	ATP/ADP
	Rigor	No nucleotides bound
	ATP binds to cleft on myosin head Conformational change in myosin Decreased affinity of myosin for actin Myosin released	ATP bound
	Cleft closes around ATP Conformational change Myosin head displaced toward ⊕ end of actin ATP hydrolysis	ATP → ADP + P_i ADP + P_i bound
	Myosin head binds new site on actin Power stroke = force	ADP bound
	ADP released Rigor	No nucleotides bound

Fig. 1.26 Cross-bridge cycle in skeletal muscle. Mechanism by which myosin "walks" toward the plus end of the actin filament. **A–E,** See the discussion in the text. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; P_i, inorganic phosphate.

conformational change that causes myosin to be displaced toward the plus end of actin. ATP is hydrolyzed to ADP and P_i, which remain attached to myosin. *D,* Myosin binds to a new site on actin (toward the plus end), constituting the force-generating, or power, stroke. Each cross-bridge cycle "walks" the myosin head 10 nanometers (10^{-8} meters) along the actin filament. *E,* ADP is released, and myosin is returned to its original state with no nucleotides bound *(A)*. Cross-bridge cycling continues, with myosin "walking" toward the plus end of the actin filament, as long as Ca^{2+} is bound to troponin C.

7. **Relaxation** occurs when Ca^{2+} is reaccumulated in the SR by the Ca^{2+} ATPase of the SR membrane **(SERCA).** When the intracellular Ca^{2+} concentration decreases to less than 10^{-7} M, there is insufficient Ca^{2+} for binding to troponin C. When Ca^{2+} is released from troponin C, tropomyosin returns to its resting position, where it blocks the myosin-binding site on actin. As long as the intracellular Ca^{2+} is low, cross-bridge cycling cannot occur and the muscle will relax.

The cross-bridge cycle produces force (tension) at the level of the contractile elements. In order for this force to be transmitted to the muscle surface, the series elastic elements (e.g., titin) must first be stretched out. As a result, there is a delay in transmission of force from the cross-bridges to the muscle surface (see Fig. 1.24). Once cross-bridge cycling has concluded, there is also a delay in the fall of muscle tension; the series elastic elements remain stretched out and thus force transmission to the muscle surface continues after intracellular Ca^{2+} has fallen and cross-bridge cycling has ceased.

Mechanism of Tetanus

A single action potential results in the release of a fixed amount of Ca^{2+} from the SR, which produces a single twitch. The twitch is terminated (relaxation occurs) when the SR reaccumulates this Ca^{2+}. However, if the muscle is stimulated repeatedly, there is insufficient time for the SR to reaccumulate Ca^{2+}, and the intracellular Ca^{2+} concentration never returns to the low levels that exist during relaxation. Instead, the level of intracellular Ca^{2+} concentration remains high, resulting in continued binding of Ca^{2+} to troponin C and continued cross-bridge cycling. In this state, there is a sustained contraction called **tetanus,** rather than just a single twitch.

Length-Tension Relationship

The length-tension relationship in muscle refers to the effect of muscle fiber length on the amount of tension the fiber can develop (Fig. 1.27). The amount of tension is determined for a muscle undergoing an **isometric contraction,** in which the muscle is allowed to develop tension at a preset length (called **preload**) but is not allowed to shorten. (Imagine trying to lift a 500-lb barbell. The tension developed would be great, but no shortening or movement of muscle would occur!) The

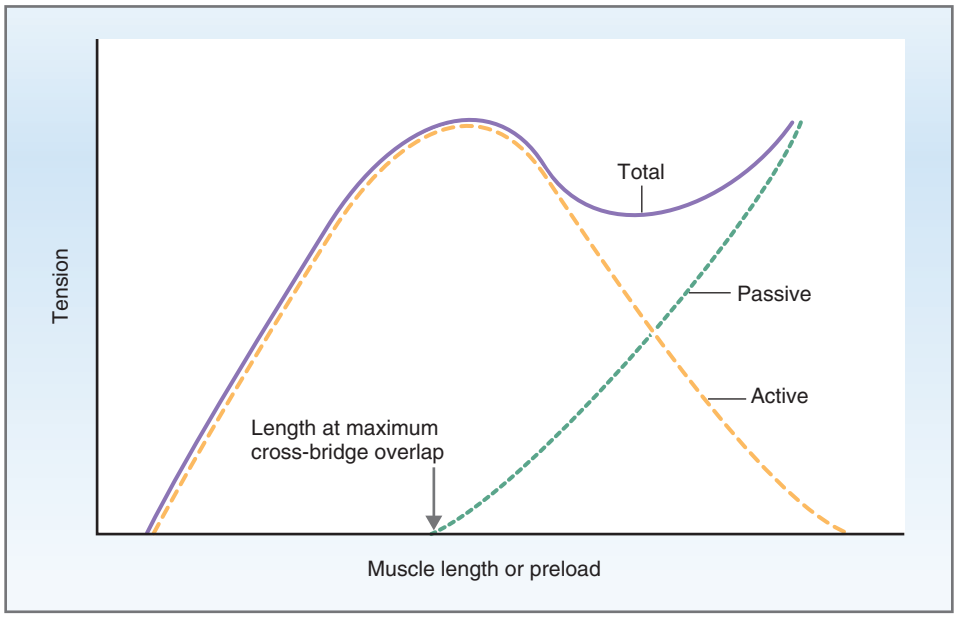


Fig. 1.27 Length-tension relationship in skeletal muscle. Maximal active tension occurs at muscle lengths where there is maximal overlap of thick and thin filaments.

following measurements of tension can be made as a function of preset length (or preload):

- **Passive tension** is the tension developed by simply stretching a muscle to different lengths. (Think of the tension produced in a rubber band as it is progressively stretched to longer lengths.)
- **Total tension** is the tension developed when a muscle is stimulated to contract at different preloads. It is the sum of the active tension developed by the cross-bridge cycling in the sarcomeres and the passive tension caused by stretching the muscle.
- **Active tension** is determined by subtracting the passive tension from the total tension. It represents the active force developed during cross-bridge cycling.

  The unusual relationship between active tension and muscle length is the **length-tension relationship** and can be explained by the mechanisms involved in the cross-bridge cycle (see Fig. 1.27). The *active tension developed is proportional to the number of cross-bridges that cycle.* Therefore the active tension is maximal when there is maximal overlap of thick and thin filaments and maximal possible cross-bridges. When the muscle is stretched to longer lengths, the number of possible cross-bridges is reduced and active tension is reduced. Likewise, when muscle length is decreased, the thin filaments collide with each other in the center of the sarcomere, reducing the number of possible cross-bridges and reducing active tension.

Force-Velocity Relationship

The force-velocity relationship, shown in Figure 1.28, describes the velocity of shortening when the force against which the muscle contracts, the **afterload,** is varied (see Fig. 1.28, *left*). In contrast to the length-tension relationship, the force-velocity relationship is determined by allowing the muscle to shorten. The force, rather than the length, is fixed, and therefore it is called an **isotonic contraction.** The velocity of shortening reflects the **speed of cross-bridge cycling.** As is intuitively obvious, the velocity of shortening will be maximal (V_{max}) when the afterload on the muscle is zero. As the afterload on the muscle increases, the velocity will be decreased because cross-bridges can cycle less rapidly against the higher resistance. As the afterload increases to even higher levels, the velocity of shortening is reduced to zero. (Imagine how quickly you can lift a feather as opposed to a ton of bricks!)

The effect of afterload on the velocity of shortening can be further demonstrated by setting the muscle to a preset length (preload) and then measuring the velocity of shortening at various levels of afterload (see Fig. 1.28, *right*). A "family" of curves is generated, each one representing a different fixed preload. The curves always intersect at V_{max}, the point where afterload is zero and where velocity of shortening is maximal.

SMOOTH MUSCLE

Smooth muscle lacks striations, which distinguishes it from skeletal and cardiac muscle. The striations found in skeletal and cardiac muscle are created by the banding patterns of thick and thin filaments in the sarcomeres. In smooth muscle, there are no striations because the thick and thin filaments, while present, are not organized in sarcomeres.

Smooth muscle is found in the walls of hollow organs such as the gastrointestinal tract, the bladder, and the uterus, as well as in the vasculature, the ureters, the bronchioles, and the muscles of the eye. The functions of smooth muscle are twofold: to produce motility

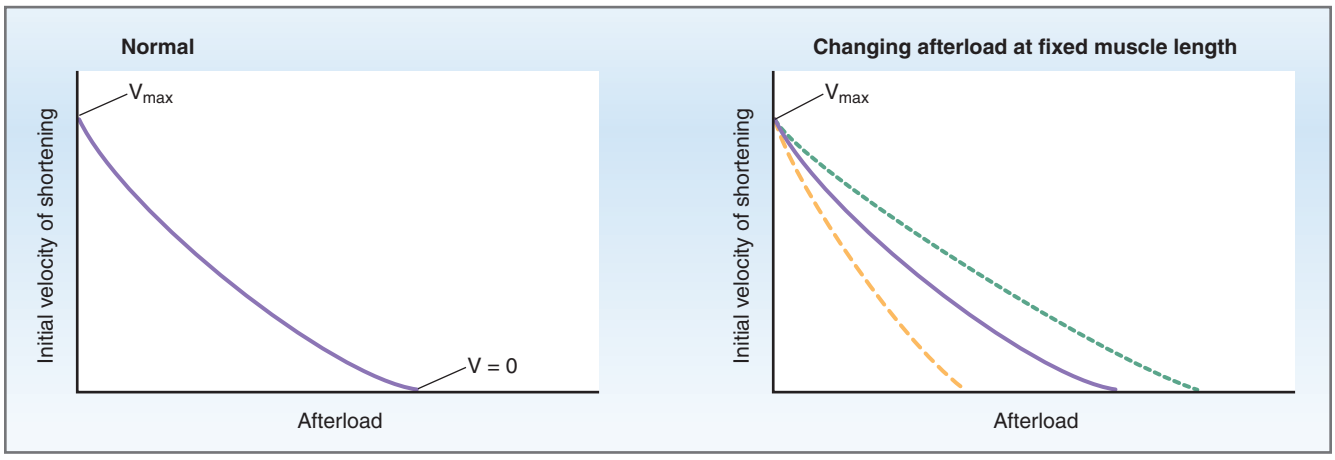


Fig. 1.28 **Initial velocity of shortening as a function of afterload in skeletal muscle.**

(e.g., to propel chyme along the gastrointestinal tract or to propel urine along the ureter) and to maintain tension (e.g., smooth muscle in the walls of blood vessels).

Types of Smooth Muscle

Smooth muscles are classified as multiunit or unitary, depending on whether the cells are electrically coupled. Unitary smooth muscle has gap junctions between cells, which allow for the fast spread of electrical activity throughout the organ, followed by a coordinated contraction. Multiunit smooth muscle has little or no coupling between cells. A third type, a combination of unitary and multiunit smooth muscle, is found in vascular smooth muscle.

Unitary Smooth Muscle

Unitary (single unit) smooth muscle is present in the gastrointestinal tract, bladder, uterus, and ureter. The smooth muscle in these organs contracts in a coordinated fashion because the cells are linked by **gap junctions.** Gap junctions are low-resistance pathways for current flow, which permit electrical coupling between cells. For example, action potentials occur simultaneously in the smooth muscle cells of the bladder so that contraction (and emptying) of the entire organ can occur at once.

Unitary smooth muscle is also characterized by spontaneous pacemaker activity, or **slow waves.** The frequency of slow waves sets a characteristic pattern of action potentials within an organ, which then determines the frequency of contractions.

Multiunit Smooth Muscle

Multiunit smooth muscle is present in the iris, in the ciliary muscles of the lens, and in the vas deferens. Each muscle fiber behaves as a separate motor unit (similar to skeletal muscle), and there is little or no coupling between cells. Multiunit smooth muscle cells are densely innervated by postganglionic fibers of the parasympathetic and sympathetic nervous systems, and it is these innervations that regulate function.

Excitation-Contraction Coupling in Smooth Muscle

The mechanism of excitation-contraction coupling in smooth muscle differs from that of skeletal muscle. Recall that in skeletal muscle, binding of actin and myosin is permitted when Ca^{2+} binds troponin C. In smooth muscle, however, there is no troponin. Rather, the interaction of actin and myosin is controlled by the binding of Ca^{2+} to another protein, **calmodulin.** In turn, Ca^{2+}-calmodulin regulates myosin-light-chain kinase, which regulates cross-bridge cycling.

Steps in Excitation-Contraction Coupling in Smooth Muscle

The steps involved in excitation-contraction coupling in smooth muscle are illustrated in Figure 1.29 and occur as follows:

1. **Depolarization of smooth muscle** opens voltage-gated Ca^{2+} channels in the sarcolemmal membrane. With these Ca^{2+} channels open, Ca^{2+} flows into the cell down its electrochemical gradient. This influx of

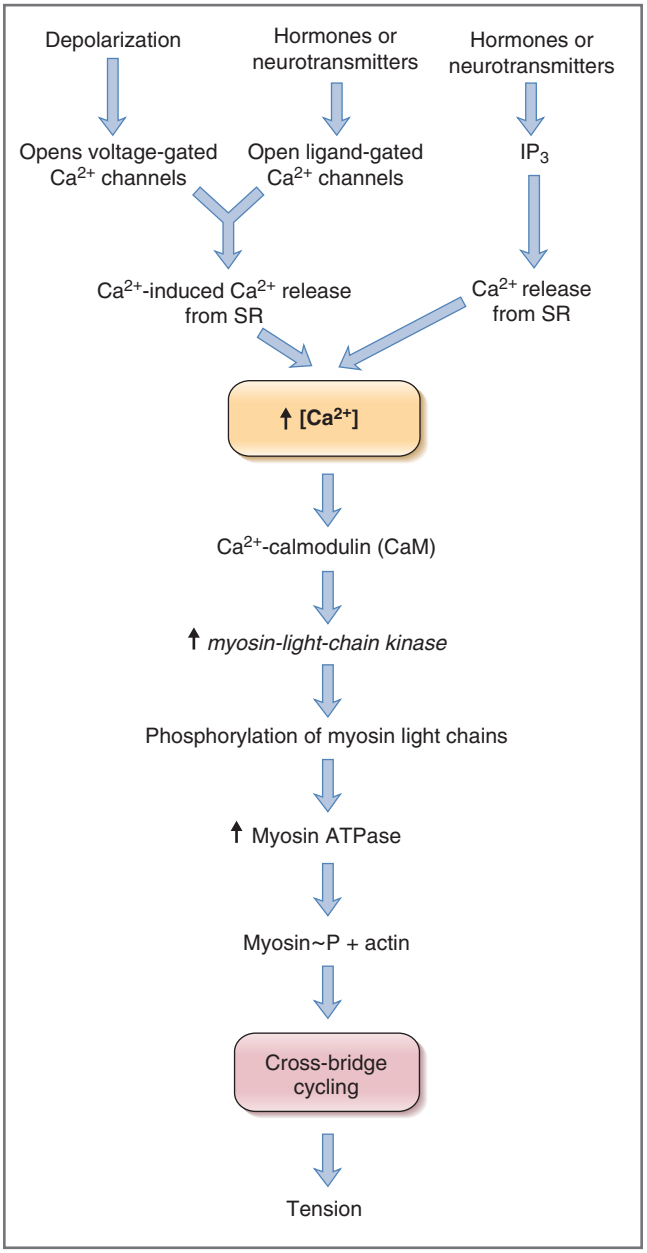


Fig. 1.29 The sequence of molecular events in contraction of smooth muscle. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; *ATPase,* adenosine triphosphatase; IP_3, inositol 1,4,5-triphosphate; *Myosin~P,* phosphorylated myosin; P_i, inorganic phosphate; *SR,* sarcoplasmic reticulum.

Ca^{2+} from the ECF causes an **increase in intracellular Ca^{2+} concentration.** In contrast to skeletal muscle, where action potentials are required to produce contraction, in smooth muscle, subthreshold depolarization (which does not lead to an action potential) can open these voltage-gated Ca^{2+} channels and cause an increase in intracellular Ca^{2+} concentration. *If* the depolarization of the smooth muscle membrane reaches threshold, then **action potentials** *can occur,* causing even greater depolarization and even greater opening of voltage-gated Ca^{2+} channels.

Ca^{2+} that enters the smooth muscle cells through voltage-gated Ca^{2+} channels releases additional Ca^{2+} from the SR (called **Ca^{2+}-induced Ca^{2+} release**). Thus the rise in intracellular Ca^{2+} is partly due to Ca^{2+} entry across the sarcolemmal membrane and partly due to Ca^{2+} release from intracellular SR stores.

2. Two additional mechanisms may contribute to the increase in intracellular Ca^{2+} concentration: ligand-gated Ca^{2+} channels and IP_3-gated Ca^{2+} release channels. **Ligand-gated Ca^{2+} channels** in the sarcolemmal membrane may be opened by various hormones and neurotransmitters, permitting the entry of *additional* Ca^{2+} from the ECF. **IP_3-gated Ca^{2+} release channels** in the membrane of the SR may be opened by hormones and neurotransmitters. Either of these mechanisms may augment the rise in intracellular Ca^{2+} concentration caused by depolarization.

3. The rise in intracellular Ca^{2+} concentration causes Ca^{2+} to bind to **calmodulin.** Like troponin C in skeletal muscle, calmodulin binds four ions of Ca^{2+} in a cooperative fashion. The Ca^{2+}-calmodulin complex binds to and activates **myosin-light-chain kinase.**

4. When activated, myosin-light-chain kinase **phosphorylates myosin light chain.** When myosin light chain is phosphorylated, the conformation of the myosin head is altered, greatly increasing its ATPase activity. (In contrast, skeletal muscle myosin ATPase activity is always high.) The increase in myosin ATPase activity allows myosin to bind actin, thus initiating cross-bridge cycling and production of tension. The amount of tension is proportional to the intracellular Ca^{2+} concentration.

5. Ca^{2+}-calmodulin, in addition to the effects on myosin described earlier, also has effects on two thin filament proteins, **calponin** and **caldesmon.** At low levels of intracellular Ca^{2+}, calponin and caldesmon bind actin, inhibiting myosin ATPase and preventing the interaction of actin and myosin. When the intracellular Ca^{2+} increases, the Ca^{2+}-calmodulin complex leads to phosphorylation of calponin and caldesmon, releasing their inhibition of myosin ATPase and facilitating the formation of cross-bridges between actin and myosin.

6. **Relaxation** of smooth muscle occurs when the intracellular Ca^{2+} concentration falls below the level needed to form Ca^{2+}-calmodulin complexes. A fall in intracellular Ca^{2+} concentration can occur by a variety of mechanisms including hyperpolarization (which closes voltage-gated Ca^{2+} channels); direct inhibition of Ca^{2+} channels by ligands such as cAMP and cyclic guanosine monophosphate (cGMP); inhibition of IP_3 production and decreased release of Ca^{2+} from SR; and increased Ca^{2+} ATPase activity in SR. Additionally, relaxation of smooth muscle can involve activation of myosin-light-chain phosphatase, which dephosphorylates myosin light chain, leading to inhibition of myosin ATPase.

Mechanisms That Increase Intracellular Ca^{2+} Concentration in Smooth Muscle

Depolarization of smooth muscle opens sarcolemmal voltage-gated Ca^{2+} channels and Ca^{2+} enters the cell from ECF. As already noted, this is only *one* source of Ca^{2+} for contraction. Ca^{2+} also can enter the cell through ligand-gated channels in the sarcolemmal membrane, or it can be released from the SR by second messenger (IP_3)-gated mechanisms (Fig. 1.30). (In contrast, recall that in skeletal muscle the rise in intracellular Ca^{2+} concentration is caused exclusively by release from the SR—Ca^{2+} does not enter the cell from the ECF.) The three mechanisms involved in Ca^{2+} entry in smooth muscle are described as follows:

- **Voltage-gated Ca^{2+} channels** are sarcolemmal Ca^{2+} channels that open when the cell membrane potential depolarizes. Thus action potentials in the smooth muscle cell membrane cause voltage-gated Ca^{2+} channels to open, allowing Ca^{2+} to flow into the cell down its electrochemical potential gradient.
- **Ligand-gated Ca^{2+} channels** also are present in the sarcolemmal membrane. They are not regulated by changes in membrane potential but by receptor-mediated events. Various hormones and neurotransmitters interact with specific receptors in the sarcolemmal membrane, which are coupled via a GTP-binding protein (G protein) to the Ca^{2+} channels. When the channel is open, Ca^{2+} flows into the cell down its electrochemical gradient. (See Chapters 2 and 9 for further discussion of G proteins.)
- **IP_3-gated Ca^{2+} channels** are present in the SR membrane. The process begins at the cell membrane, but the source of the Ca^{2+} is the SR rather than the ECF. Hormones or neurotransmitters interact with specific receptors on the sarcolemmal membrane

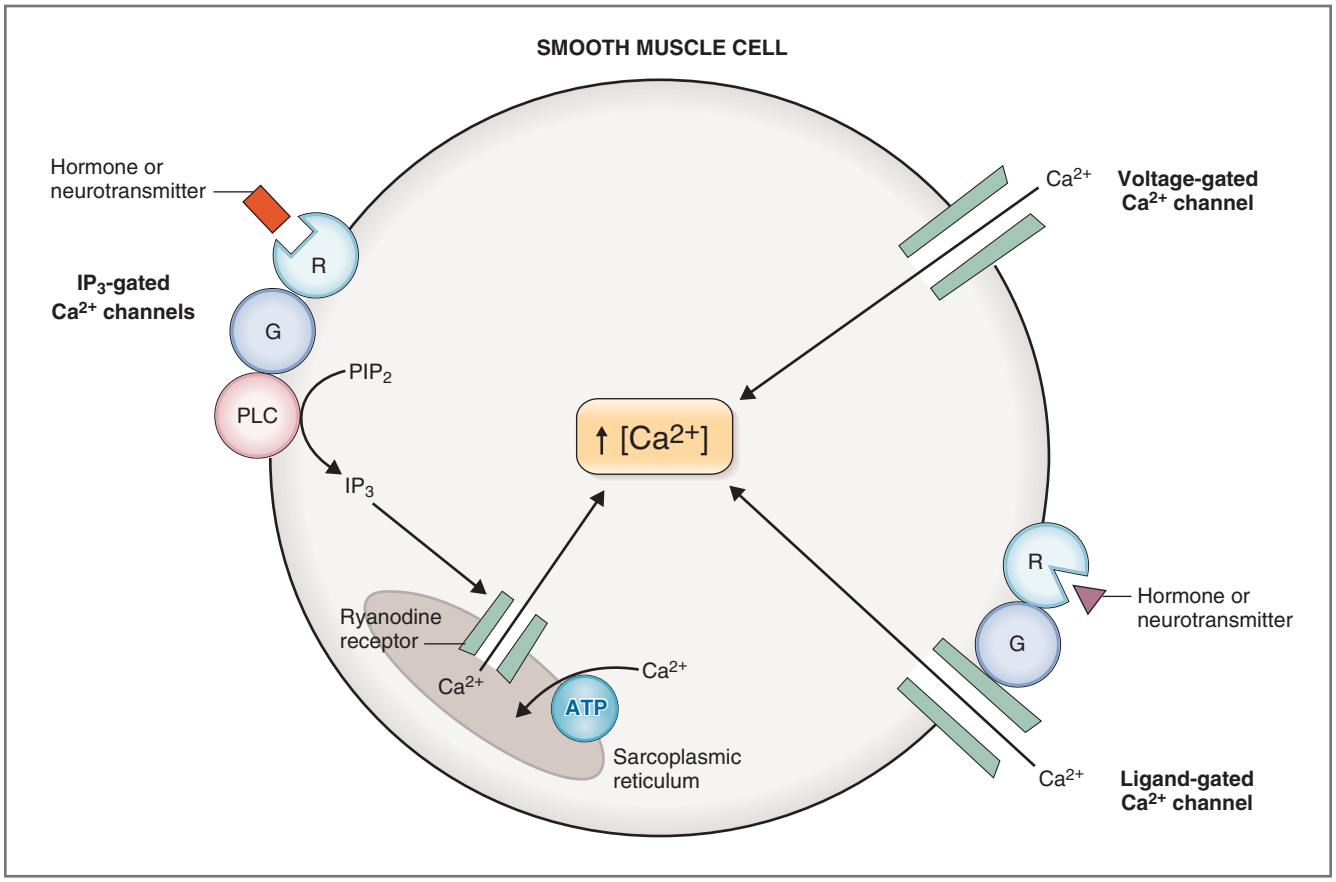


Fig. 1.30 Mechanisms for increasing intracellular [Ca^{2+}] in smooth muscle. *ATP,* Adenosine triphosphate; *G,* GTP-binding protein (G protein); *IP_3,* inositol 1,4,5-triphosphate; *PIP_2,* phosphatidylinositol 4,5-diphosphate; *PLC,* phospholipase C; *R,* receptor for hormone or neurotransmitter.

(e.g., norepinephrine with α_1 receptors). These receptors are coupled, via a G protein, to **phospholipase C (PLC).** PLC catalyzes the hydrolysis of phosphatidylinositol 4,5-diphosphate (PIP_2) to IP_3 and diacylglycerol (DAG). **IP_3** then diffuses to the SR, where it opens Ca^{2+} release channels (similar to the mechanism of the ryanodine receptor in skeletal muscle). When these Ca^{2+} channels are open, Ca^{2+} flows from its storage site in the SR into the ICF. (See Chapter 9 for discussion of IP_3-mediated hormone action.)

Ca^{2+}-Independent Changes in Smooth Muscle Contraction

In addition to the contractile mechanisms in smooth muscle that depend on changes in intracellular Ca^{2+} concentration, the degree of contraction also can be regulated by Ca^{2+}-independent mechanisms. For example, in the presence of a constant level of intracellular Ca^{2+}, if there is activation of myosin-light-chain kinase, more cross-bridges will cycle and more tension will be produced **(Ca^{2+}-sensitization);** conversely, if there is activation of myosin-light-chain phosphatase, fewer cross-bridges will cycle and less tension will be produced **(Ca^{2+}-desensitization).**

SUMMARY

- Water, a major component of the body, is distributed among two major compartments, ICF and ECF. ECF is further distributed among the plasma and the interstitial fluid. The differences in composition of ICF and ECF are created and maintained by transport proteins in the cell membranes.
- Transport may be either passive or active. If transport occurs down an electrochemical gradient, it is passive and does not consume energy. If transport occurs against an electrochemical gradient, it is active. The energy for active transport may be primary (using ATP) or secondary (using energy from the Na^+ gradient). Osmosis occurs when an impermeable solute creates an osmotic pressure difference across a membrane, which drives water flow.