Physiology

SIXTH 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, SIXTH 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 2014, 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

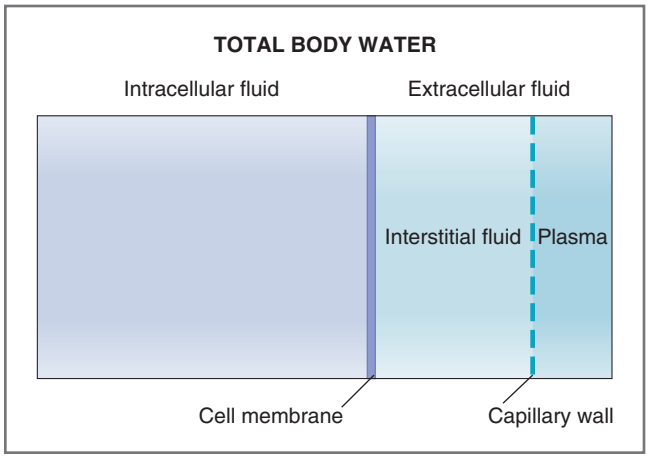

Fig. 1.1 **Body fluid compartments.**

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.

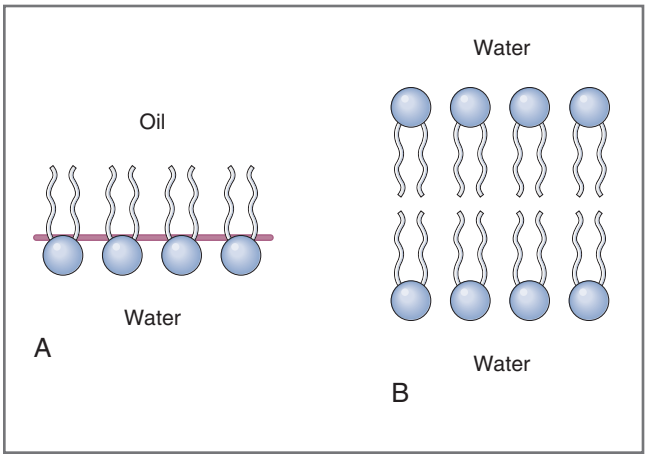

Fig. 1.2 Orientation of phospholipid molecules at oil and water interfaces. Depicted are the orientation of phospholipid at an oil-water interface **(A)** and the orientation of phospholipid in a bilayer, as occurs in the cell membrane **(B).**

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

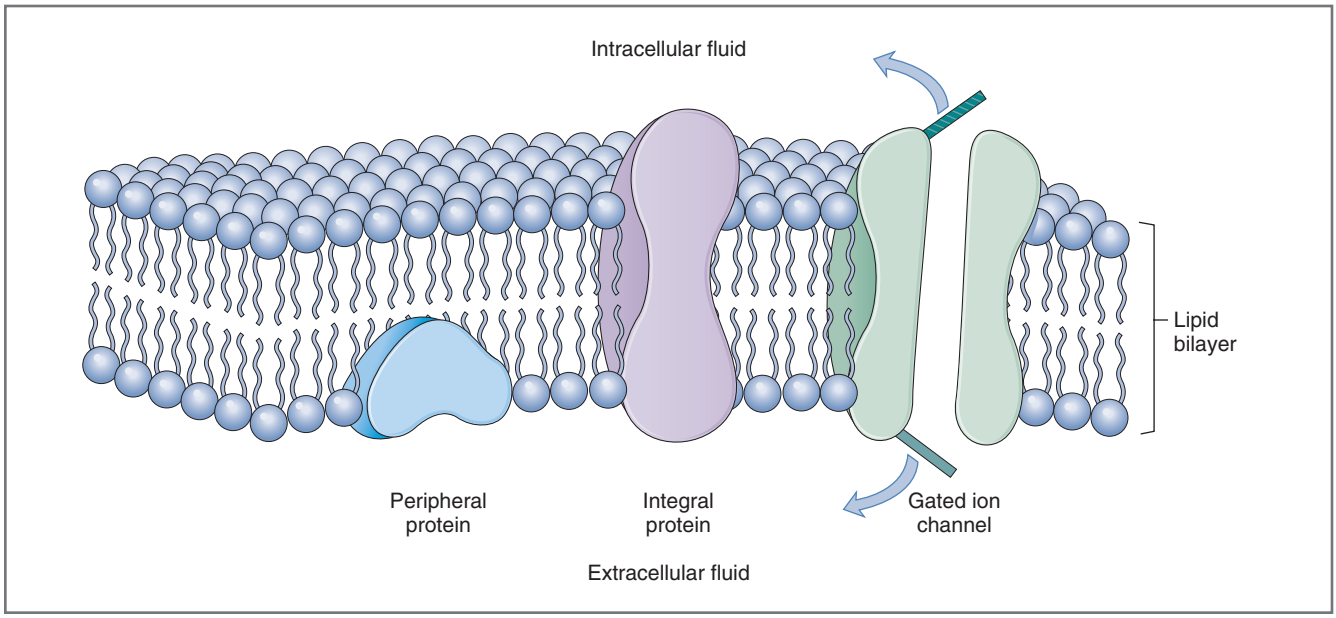

Fig. 1.3 **Fluid mosaic model for cell membranes.**

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

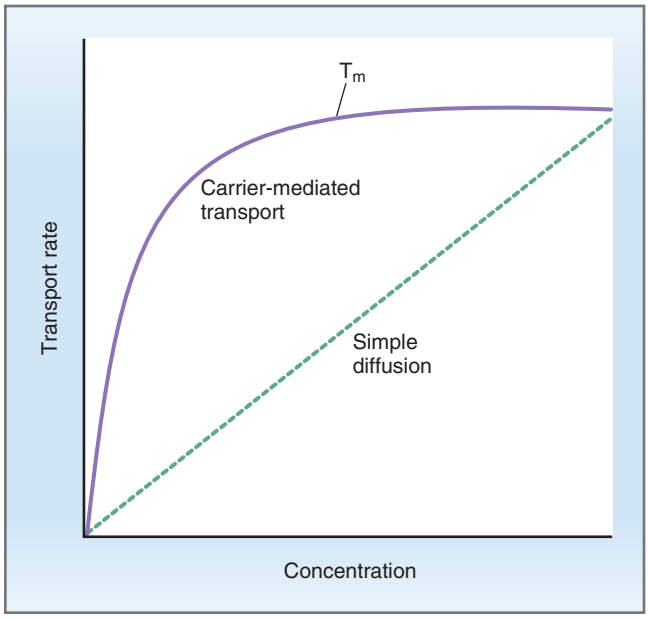

Fig. 1.4 Kinetics of carrier-mediated transport. T_m, Transport maximum.

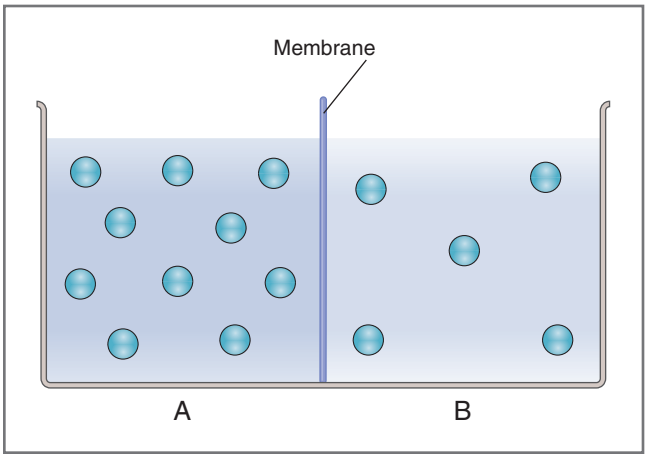

Fig. 1.5 Simple diffusion. The two solutions, **A** and **B,** are separated by a membrane, which is permeable to the solute *(circles).* Solution A initially contains a higher concentration of the solute than does Solution B.

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

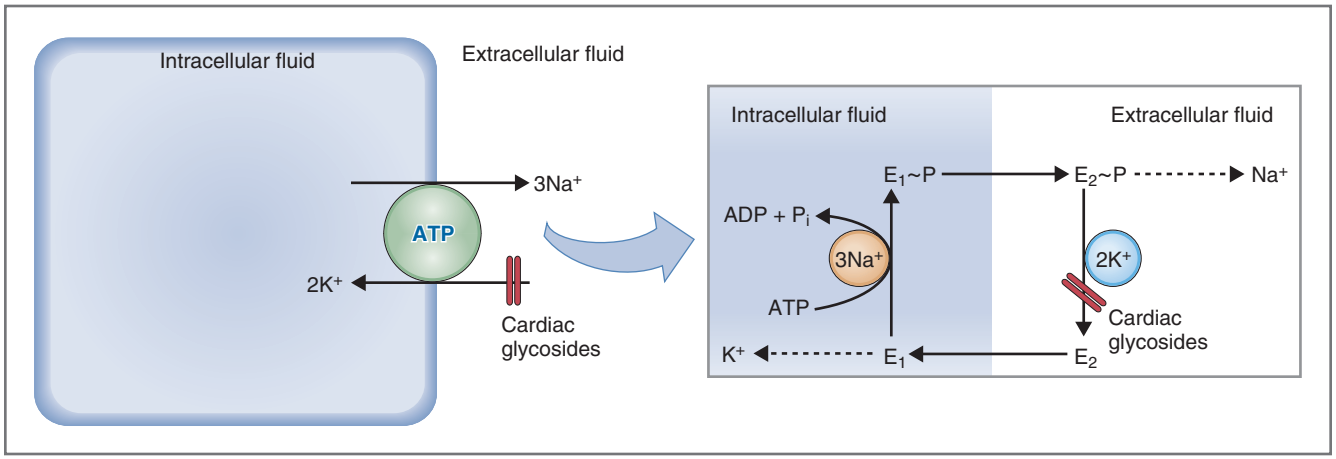

Fig. 1.6 Na^+-K^+ pump of cell membranes. *ADP,* Adenosine diphosphate; *ATP,* adenosine triphosphate; *E,* Na^+-K^+ ATPase; *E~P,* phosphorylated Na^+-K^+ ATPase; P_i, inorganic phosphate.

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

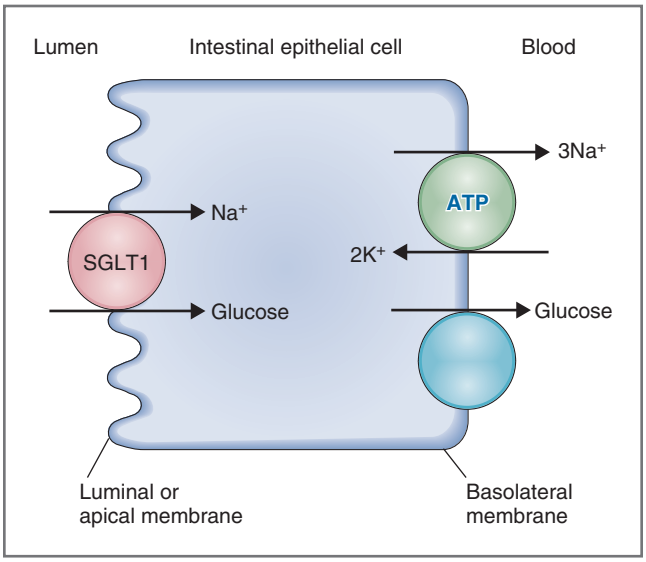

Fig. 1.7 Na^+-glucose cotransport in an intestinal epithelial cell. *ATP,* Adenosine triphosphate; *SGLT1,* Na^+-glucose transport protein 1.

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.

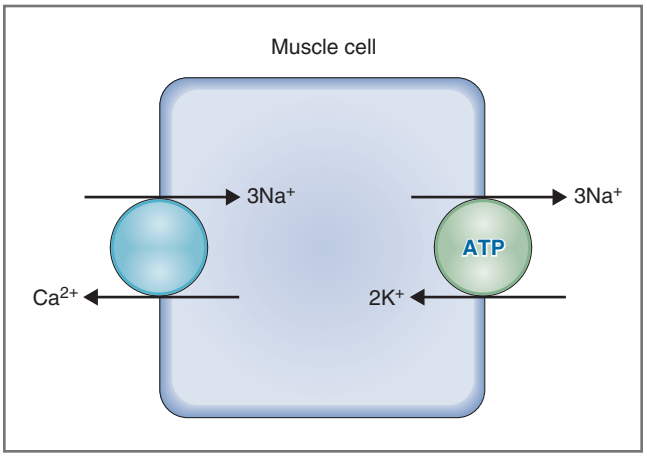

Fig. 1.8 Ca^{2+}-Na^+ countertransport (exchange) in a muscle cell. *ATP,* Adenosine triphosphate.

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

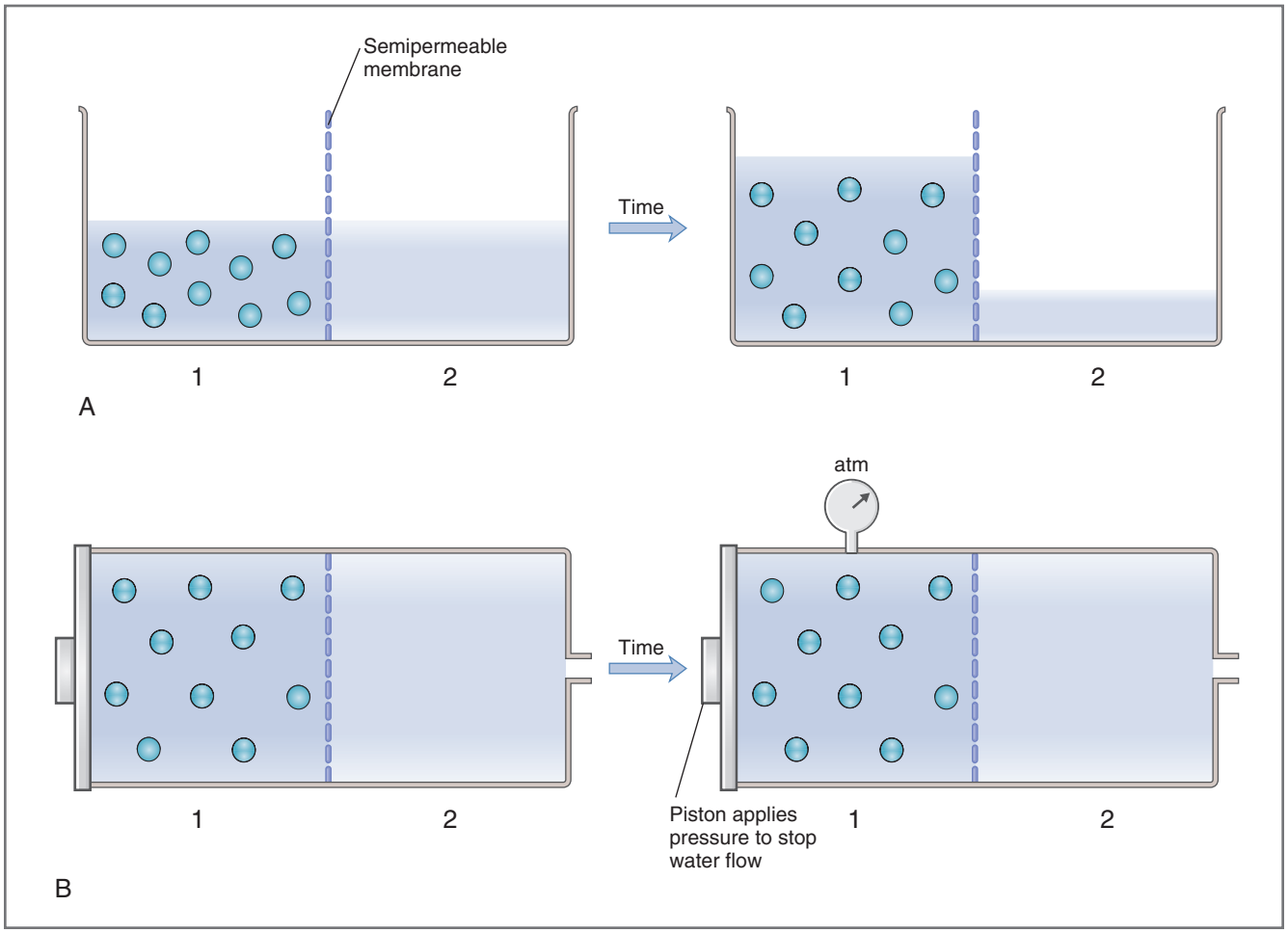

Fig. 1.9 Osmosis across a semipermeable membrane. A, Solute *(circles)* is present on one side of a semipermeable membrane; with time, the osmotic pressure created by the solute causes water to flow from Solution 2 to Solution 1. The resulting volume changes are shown. **B,** The solutions are closed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

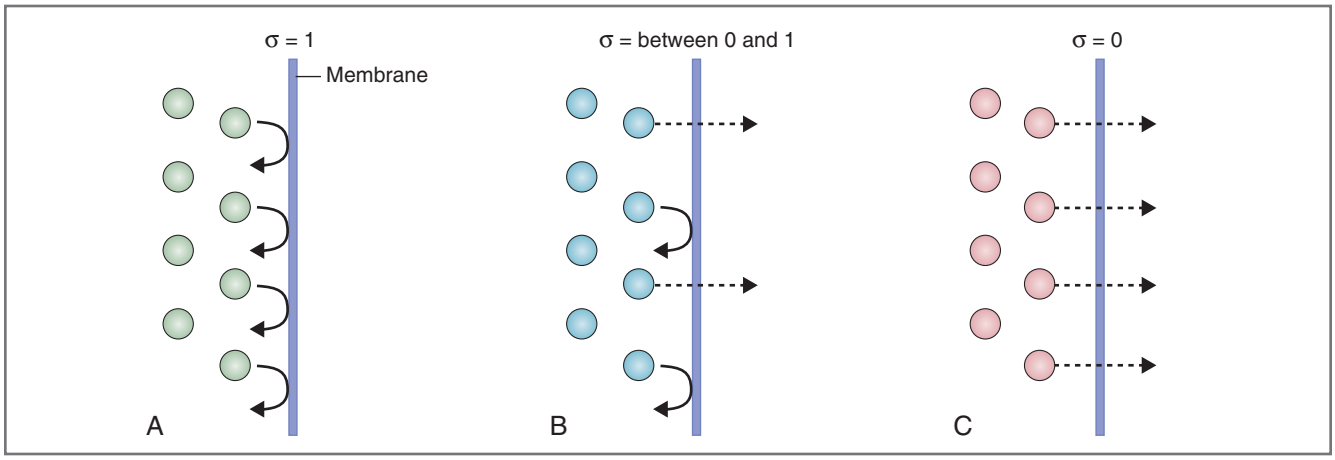

Fig. 1.10 **Reflection coefficient (σ).**

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

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

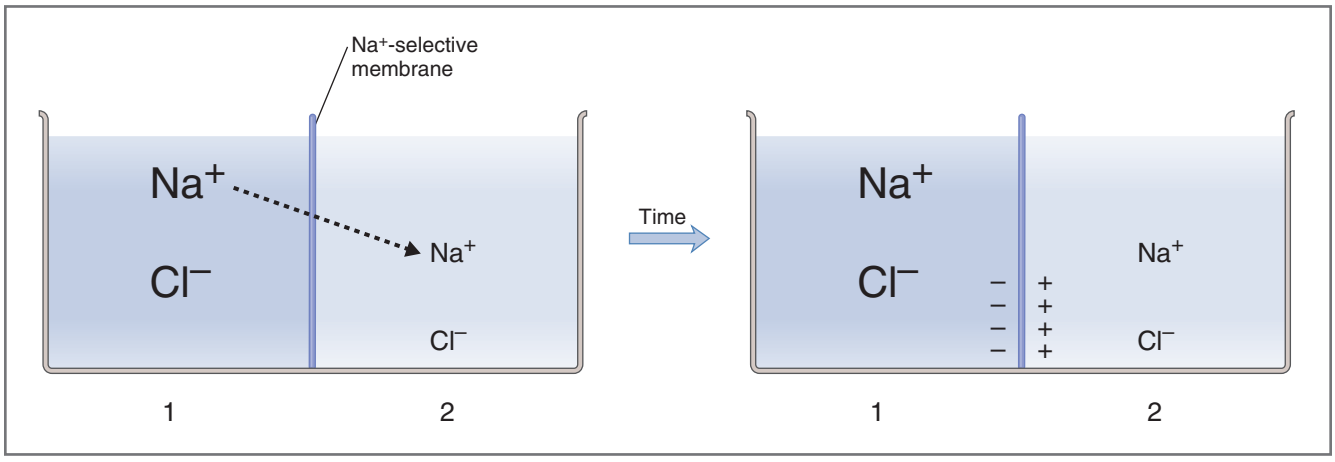

Fig. 1.11 Generation of an Na^+ diffusion potential.

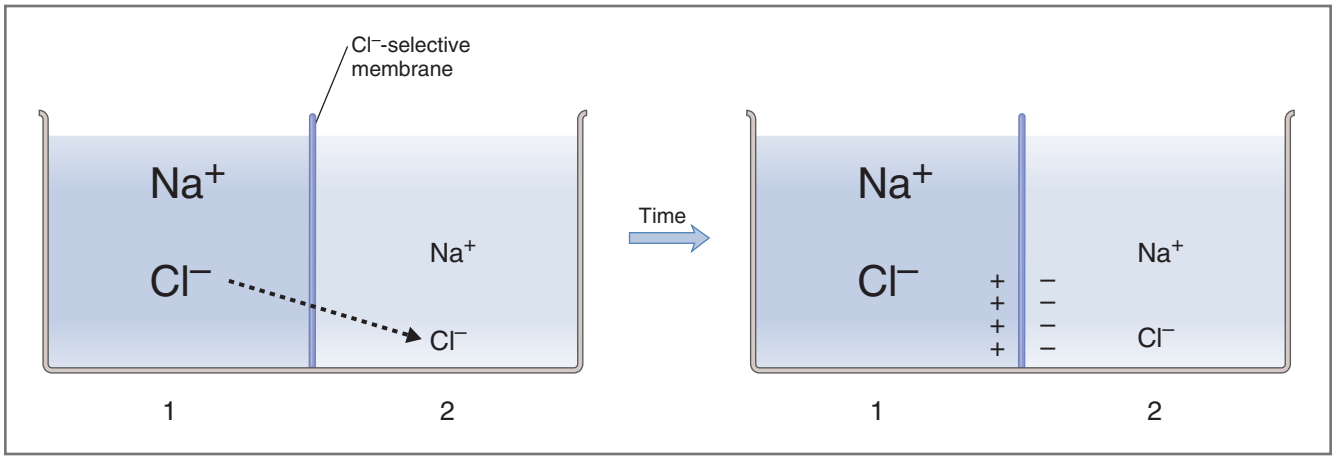

Fig. 1.12 **Generation of a Cl^- diffusion potential.**

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*