CRC PRESS
PHARMACY EDUCATION SERIES

Essentials of Human Physiology for Pharmacy

Second Edition

LAURIE KELLY McCORRY

CRC Press
Taylor & Francis Group

Essentials *of* Human Physiology *for* Pharmacy

Second Edition

CRC PRESS
PHARMACY EDUCATION SERIES

Essentials of Human Physiology for Pharmacy, Second Edition
Laurie Kelly McCorry

Pharmaceutical Dosage Forms and Drug Delivery
Ram I. Mahato

Basic Pharmacokinetics
Mohsen A. Hedaya

Essentials of Law and Ethics for Pharmacy Technicians, Second Edition
Kenneth M. Strandberg

Pharmacy: What It Is and How It Works, Second Edition
William N. Kelly

Basic Pharmacology: Understanding Drug Actions and Reactions
Maria A. Hernandez and Appu Rathinavelu

Managing Pharmacy Practice: Principles, Strategies, and Systems
Andrew M. Peterson

Essential Math and Calculations for Pharmacy Technicians
Indra K. Reddy and Mansoor A. Khan

Pharmacoethics: A Problem-Based Approach
David A. Gettman and Dean Arneson

Pharmaceutical Care: Insights from Community Pharmacists
William N. Tindall and Marsha K. Millonig

Essentials of Pathophysiology for Pharmacy
Martin M. Zdanowicz

Quick Reference to Cardiovascular Pharmacotherapy
Judy W. M. Cheng

Essentials of Pharmacy Law
Douglas J. Pisano

Pharmacokinetic Principles of Dosing Adjustments: Understanding the Basics
Ronald D. Schoenwald

Pharmaceutical and Clinical Calculations, Second Edition
Mansoor A. Khan and Indra K. Reddy

Strauss's Federal Drug Laws and Examination Review, Fifth Edition Revised
Steven Strauss

Inside Pharmacy: Anatomy of a Profession
Raymond A. Gosselin and Jack Robbins

Understanding Medical Terms: A Guide for Pharmacy Practice, Second Edition
Walter F. Stanaszek, Mary J. Stanaszek, Robert J. Holt, and Steven Strauss

Essentials *of* Human Physiology *for* Pharmacy

Second Edition

LAURIE KELLY McCORRY

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2008 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Version Date: 20110715

International Standard Book Number-13: 978-1-4200-4391-4 (eBook - PDF)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

To TKM

$\infty \times 9$

Contents

Preface

This book has been written for students of pharmacy who are studying human physiology. Specifically, it is meant to prepare these students for their subsequent study of pathophysiology, pharmacology, and pharmacotherapeutics. As such, an important goal of this book is to present concepts in physiology in a way that students will clearly understand the relevance of these concepts to the practice of pharmacy.

In this second edition of the textbook each of the original 19 chapters has been updated and expanded to provide a more thorough presentation of these topics. Additional figures and tables have been added to present and summarize the material more clearly and concisely. New sections with pronunciations and definitions of relevant medical terms are found at the end of each chapter. The number of subsections of *pharmacy applications* has been increased from 30 to 43 to better illustrate the application of physiological principles to the practice of pharmacy. Finally, this edition includes two new chapters, "The Reproductive System" and "The Immune System," so that the discussion of human physiology is more comprehensive.

As with the first edition, this book begins with an overview of the fundamental aspects of cell membrane physiology with particular emphasis on nerve cell function. This is followed by a detailed discussion of the two major regulatory systems in the body: the nervous system, including the brain, spinal cord, pain, and autonomic nervous system; as well as the endocrine system. The book then continues with in-depth presentations of the reproductive, muscular, cardiovascular, immune, respiratory, digestive, and renal systems. An important focus throughout the text is how tissue and organ function are regulated in order to maintain homeostasis.

The intent of this book is to present the material in a manner as clearly and concisely as possible. It has been written with simple, straightforward language aimed toward the undergraduate student studying human physiology for the first time. Study objectives are provided at the beginning of each chapter to help students focus on important principles and mechanisms.

Whenever possible, information is provided in the form of bulleted lists, tables, figures, or flowcharts. Finally, subsections of *pharmacy applications* separate from the text serve to relate a given concept in physiology to the student's future career in the practice of pharmacy.

Laurie Kelly McCorry
Bay State College
Boston, Massachusetts

Acknowledgment

I am extremely grateful to my friend and former colleague, Lynne Sylvia, Pharm.D., who reviewed all of the pharmacy-related material for accuracy.

About the author

Laurie Kelly McCorry, Ph.D., was an associate professor of physiology for 16 years at the Massachusetts College of Pharmacy and Health Sciences, where she taught all aspects of physiology, from introductory and advanced courses to exercise physiology and pathophysiology. She is currently the chair of the Department of Allied Health at Bay State College in Boston, Massachusetts. Her other publications include *Physiology Case Studies in Pharmacy* (American Pharmacists Association, Washington, D.C., 2006).

chapter one

Physiology and the concept of homeostasis

Study objectives

- Define the *internal environment*
- Understand the importance of homeostasis
- Describe the overall function of each of the three major components of the nervous system
- Compare the general functions of the nervous system and the endocrine system
- Distinguish between negative feedback and positive feedback
- Describe the potential role of medications in the maintenance of homeostasis

1.1 Introduction

Physiology is the study of the functions of the human body. In other words, if anatomy describes the structure or morphology of the parts of the body, physiology describes how these parts work. This discipline considers the mechanisms by which each of the various tissues and organs carries out their specific activities. Emphasis is placed on the processes that control and regulate the physiological activities in the body.

1.2 Homeostasis

In order for the body to function optimally, conditions within the body, referred to as the *internal environment*, must be very carefully regulated. Therefore, many important variables, such as body temperature, blood pressure, blood glucose, oxygen and carbon dioxide content of the blood, as well as electrolyte balance are actively maintained within narrow physiological limits.

This maintenance of relatively constant or steady-state internal conditions is referred to as *homeostasis*. It is important because the cells and tissues of the body will survive and function efficiently only when these internal conditions are properly maintained. This is not to say that the internal environment is fixed or unchanging. The body is constantly faced with a changing external environment as well as with events and activities occurring within

Table 1.1 Contribution of Organ Systems to the Maintenance of Homeostasis

Organ System	Function
Nervous System	Regulates muscular activity and glandular secretion, responsible for all activities associated with the mind
Endocrine System	Regulates metabolic processes through secretion of hormones
Muscular System	Allows for body movement, contributes to thermoregulation
Circulatory System	Transports nutrients, oxygen, waste, carbon dioxide, electrolytes, and hormones throughout the body
Respiratory System	Obtains oxygen and eliminates carbon dioxide, regulates acid–base balance (pH)
Gastrointestinal Tract	Digests and absorbs food substances to provide nutrients to the body
Renal System	Eliminates waste products from the body, regulates blood volume and blood pressure, regulates acid–base balance (pH)

the body that may alter the balance of important variables. For example, most metabolic reactions within our cells consume oxygen and glucose. These substances must then be replaced. In addition, these reactions produce metabolic wastes including carbon dioxide and urea, which must then be eliminated. Therefore, it is more accurate to say that the internal environment is in a *dynamic steady state*, one that is constantly changing, but where optimal conditions are physiologically maintained.

All of the organ systems in the body, except the reproductive system, contribute to the maintenance of homeostasis (see Table 1.1). For example, the gastrointestinal tract digests foods to provide nutrients to the body. The respiratory system obtains oxygen and eliminates carbon dioxide. The circulatory system transports all of these materials and others from one part of the body to another. The renal system eliminates wastes and plays a role in regulating blood volume and blood pressure.

The study of physiology includes not only the study of how each of these systems carries out its functions, but also the mechanisms involved which regulate these activities in order to maintain homeostasis under a variety of conditions. For example, the body's needs are very different during a resting state compared to that of exercise. How do organ systems adjust their activities in response to varied levels of physical exertion or when confronted with altered internal and external environments? In order to maintain homeostasis, the body must be able to monitor and sense changes in the internal environment. Second, it must be able to compensate, or make adjustments, for these changes.

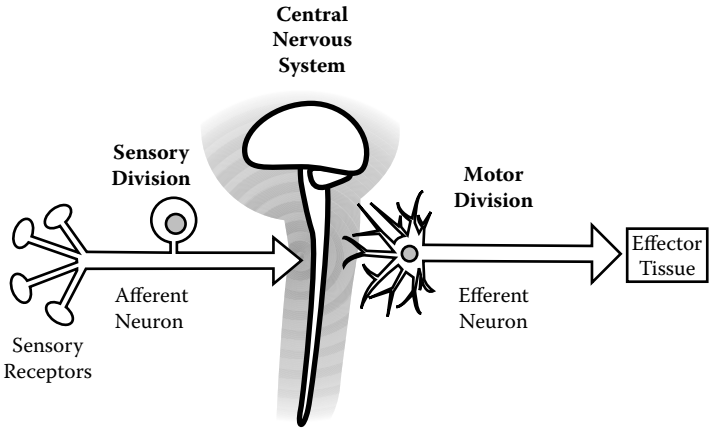


Figure 1.1 Functional components of the nervous system. The sensory division of the peripheral nervous system is sensitive to changes in the internal and external environment. The information gathered by this component of the nervous system is transmitted to the central nervous system (CNS) where it is processed, integrated, and interpreted. The CNS then determines the appropriate response to this input. This response is carried out by the transmission of nerve impulses in the motor division of the peripheral nervous system to the effector tissues.

There are two regulatory systems in the body that influence the activity of all the other organ systems so that homeostasis is ultimately maintained. These are the nervous system and the endocrine system.

There are three functional components of the *nervous system* (see Figure 1.1):

- Sensory division of the peripheral nervous system
- Central nervous system
- Motor division of the peripheral nervous system

Many different types of sensory receptors are located throughout the body. These receptors monitor the status of the internal environment or that of the surroundings. *Sensory receptors* are sensitive to specific types of stimuli and measure the value of a physiological variable. For example, *arterial baroreceptors* measure arterial blood pressure, and *chemoreceptors* measure the amount of oxygen and carbon dioxide in the arterial blood. The information detected by these sensors then travels by way of *afferent* neuronal pathways to the *central nervous system* (CNS).

The CNS is the *integrative portion* of the nervous system and consists of the *brain* and the *spinal cord*. The brain receives, processes, and stores sensory input; generates thoughts; and determines the reactions that the body should perform in response to this input. The spinal cord is important in processing reflexes. It is within this integration area of the nervous system that the actual

value of a physiological variable as measured by a sensory receptor is compared to its set point or optimal value. One or more compensatory responses to the sensory input are then determined.

The third component of the nervous system is the *motor division*. Appropriate signals are transmitted from the CNS to various body parts or *effector tissues* by way of *efferent* neuronal pathways. These effector tissues include a variety of the body's tissues and organs, specifically, the muscles and glands within the tissues and organs. The effector tissues carry out the appropriate physiological responses to bring the variable back to within its normal limits.

The other regulatory system in the body contributing to the maintenance of homeostasis is the *endocrine system* that carries out its effects by secreting *hormones*. These hormones are transported in the blood to the specific tissues upon which they exert their effects.

The following generalizations regarding the two regulatory systems may be made. The nervous system primarily regulates muscular activity and glandular secretion. The endocrine system primarily regulates metabolic activity in the body's cells. However, these two systems may not only work together in the regulation of many organs, they may also influence each other's activity.

1.3 Negative feedback

Most of the body's *compensatory homeostatic mechanisms* function by way of *negative feedback*. This is a response that causes the level of a variable to change in a direction opposite to that of the initial change. Because the response returns the variable back to its baseline level, it has a *stabilizing* effect on the body. For example, when blood pressure increases, the arterial baroreceptors are stimulated, and an increased number of nerve impulses are transmitted to the CNS through afferent pathways. The region of the brain regulating the cardiovascular system responds to this sensory input by altering efferent nerve activity to the heart. The result is a decrease in heart rate and, therefore, a decrease in blood pressure back to its baseline value (see Figure 1.2). In general, when some physiological variable becomes too high or too low, a control system elicits a negative feedback response consisting of one or a series of changes that returns the variable to within its normal physiological range. These compensatory mechanisms operating via negative feedback allow the body to effectively maintain homeostasis.

Interestingly, one of the greatest stressors on the body and, therefore, challenges to the maintenance of homeostasis is increased physical activity or exercise. During intense exercise, glucose utilization can be increased up to twentyfold, skeletal muscle pH drops dramatically, several liters of water can be lost in the form of sweat, and core body temperature can increase to as high as 106°F. These profound disturbances must be compensated for in order to ensure cell survival. An important focus throughout this textbook

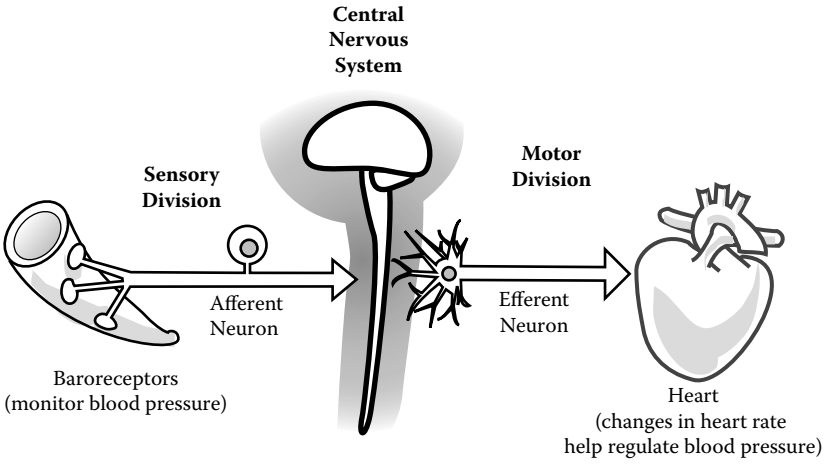


Figure 1.2 Negative feedback. Negative feedback responses are employed throughout the body in order to maintain homeostasis. In this example, any change in blood pressure, which is monitored within the circulatory system and processed or interpreted within the CNS, will cause reflex changes in heart rate. The change in heart rate will be in the opposite direction of the change in blood pressure. If blood pressure increases, then heart rate decreases. If blood pressure decreases, then heart rate increases. In this way, blood pressure is adjusted back to its normal value.

will be how tissue and organ system functions are regulated under various normal physiological conditions and, where appropriate, under abnormal pathophysiological conditions. Furthermore, discussions of how basic physiological principles may be applied to the practice of pharmacy are included.

1.4 Positive feedback

Although there are fewer physiological examples, there are also processes in the body that utilize positive feedback responses. These responses *amplify* the initial change in a variable. In other words, if the level of a variable increases, positive feedback responses cause the level of the variable to increase further. In essence, this has a *destabilizing* effect on the body. An interesting example involves parturition, or childbirth. As labor begins and the uterus contracts, it pushes the baby downward and stretches the cervix. Consequently, signals are transmitted to the hypothalamus of the brain which stimulates the secretion of the hormone oxytocin. This hormone acts on the uterus to cause powerful muscular contractions, again pushing downward on the baby and resulting in further stretching of the cervix. This stretching causes the secretion of more oxytocin and so on until the baby has been expelled from the uterus. Other examples of positive feedback involve voltage-gated sodium channels and the generation of electrical signals in nerve cells (discussed in Chapter 4) as well as the mechanism for blood clotting (discussed in Chapter 17).

PHARMACY APPLICATION:
HOMEOSTATIC FUNCTIONS OF DRUGS

Diseases are generally divided into two categories: those in which the pathophysiology involves the internal failure of some normal physiological process and those that originate from some external source such as bacterial or viral infection. In either case, the individual is unable to maintain homeostasis, and one or more variables in the internal environment will be disrupted. As a result, tissue or organ function is impaired. Therefore, many of the medications currently in use are designed to assist the body in the maintenance of homeostasis when its own regulatory mechanisms fail to do so. For example, angiotensin-converting enzyme (ACE) inhibitors, such as enalopril, and beta-blockers, such as propranolol, lower blood pressure in patients with idiopathic (unexplained) hypertension (elevated blood pressure). A common complication in patients with type 1 diabetes mellitus is hyperglycemia (excess glucose in the blood). Insulin injections allow the cells of these patients to take up and store glucose, which effectively lowers the blood glucose to the normal range. Diuretics, such as furosemide, decrease blood volume and, therefore, reduce cardiac workload in patients with congestive heart failure. In each of these disorders, pharmacological intervention is necessary for the given organ system to function efficiently and effectively in order to maintain the health of the patient.

Medical terminology

Afferent (ăf′ĕr-ĕnt): Carrying or transporting toward a central location.

Efferent (e˘f′e˘r-e˘nt): Carrying or transporting away from a central location.

Homeostasis (hō″mē-ō-stā′sĭs): The maintenance of an internal equilibrium or a dynamic steady state in an individual by altering appropriate physiological processes in the body's tissues and organs.

Internal environment: Conditions within the body.

Negative feedback: A response that opposes the original change in the system.

Positive feedback: A response that amplifies the original change in the system.

Bibliography

AHFS Drug Information 2000, American Society of Health-System Pharmacists, Bethesda, MD, 2000.

Papanek, P. E., Exercise Physiology and the Bioenergetics of Muscle Contraction, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 8.

Porterfield, S., *Endocrine Physiology,* Mosby, St. Louis, MO, 2001.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Thomson Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E., Raff, H., and Strang, K., *Vander's Human Physiology, The Mechanisms of Body Function,* 10th ed., McGraw-Hill, New York, 2006.

chapter two

Plasma membrane

Study objectives

- Describe the function of each of the components of the plasma membrane
- Understand the physiological importance of the permeability barrier created by the plasma membrane
- Describe the factors that affect diffusion
- Explain how osmosis takes place
- Understand the clinical significance of the osmotic pressures of solutions
- Describe the factors that affect mediated transport
- Compare and contrast facilitated diffusion and active transport

2.1 Introduction

Each cell is surrounded by a plasma membrane that separates the cytoplasmic contents of the cell, or the intracellular fluid, from the fluid outside of the cell, the extracellular fluid. An important homeostatic function of this plasma membrane is to serve as a *permeability barrier* that insulates or protects the cytoplasm from immediate changes in the surrounding environment. Furthermore, it allows the cell to maintain a cytoplasmic composition that is very different from that of the extracellular fluid. The functions of neurons (nerve cells) and muscle cells depend on this difference. The plasma membrane also contains many enzymes and other components, such as antigens and receptors. These structures allow cells to interact with other cells, neurotransmitters, blood-borne substances such as hormones, and various other chemical substances such as drugs.

2.2 Structure and function of the plasma membrane

The major components of the plasma membrane include the following:

- Phospholipids
- Cholesterol
- Proteins
- Carbohydrates

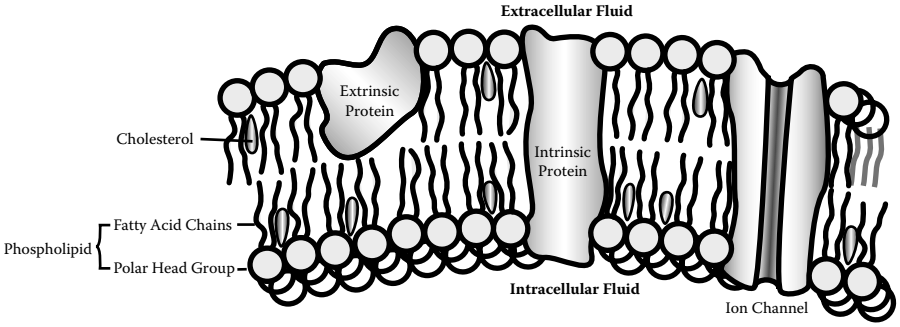


Figure 2.1 Structure of the plasma membrane. The plasma membrane is composed of a bilayer of phospholipid molecules. Associated with this bilayer are intrinsic proteins, which are embedded within and span the membrane, and extrinsic proteins, which are found on the external or internal surface of the membrane. Molecules of cholesterol are found in the inner, nonpolar region of the membrane.

The basic structure of the plasma membrane is formed by *phospholipids* (Figure 2.1). These molecules are one of the more abundant of the membrane components. Phospholipids are *amphipathic* molecules that have both polar (water-soluble) and nonpolar (water-insoluble) regions. They are composed of a phosphorylated glycerol backbone that forms a polar head group that is hydrophilic; and a nonpolar region that contains two hydrophobic fatty acid chains. In an aqueous environment such as the body, these molecules are arranged in a formation referred to as the *lipid bilayer,* consisting of two layers of phospholipids. The polar region of the molecule is oriented toward the outer surface of the membrane where it can interact with water; and the nonpolar, hydrophobic fatty acids are in the center of the membrane away from the water. The functional significance of this lipid bilayer is that it creates a *semipermeable barrier.* Lipophilic, or nonwater-soluble, substances can readily cross the membrane by simply passing through its lipid core. Important examples of these substances include gases, such as oxygen and carbon dioxide, and fatty acid molecules, which are used to form energy within muscle cells.

Most hydrophilic, or water-soluble, substances are repelled by this hydrophobic interior and cannot simply diffuse through the membrane. Instead, these substances must cross the membrane using specialized transport mechanisms. Examples of lipid-insoluble substances that require such mechanisms include proteins, nutrient molecules, such as glucose and amino acids, and all species of ions (Na^+, Ca^{++}, H^+, Cl^-, and HCO_3^-). Therefore, the plasma membrane plays a very important role in determining the composition of the intracellular fluid by selectively permitting substances to move in and out of the cell.

PHARMACY APPLICATION:
LIPID SOLUBILITY AND DRUG ELIMINATION

The lipid solubility of many substances can change when physiological conditions vary. For example, the surrounding pH can determine whether a molecule is in a protonated form (positively charged, lipid insoluble) or in an unprotonated form (uncharged, lipid soluble). As discussed, charged substances do not readily cross the membrane, as do uncharged substances. This principle regarding lipid solubility is used in the treatment of an overdose of phenobarbital, a barbiturate used for sedation and seizure disorders. Phenobarbital is normally 30% removed by urinary excretion. In the case of an overdose, it would be advantageous to enhance urinary excretion. Alkalinization of the urine to a pH of 7.5 to 8 helps to promote excretion. In fact, by alkalinizing the urine, the amount of phenobarbital that is excreted increases five- to tenfold. After alkalinization, more phenobarbital would be ionized in the urine, and, therefore, would be lipid insoluble. As a result, this drug would not be reabsorbed from the kidney, and instead, it would be eliminated in the urine.

Another important aspect of the lipid bilayer is that the phospholipids are not held together by chemical bonds. This enables the molecules to move about freely within the membrane, resulting in a structure that is not rigid in nature but instead is very fluid and pliable. Also contributing to membrane fluidity is the presence of *cholesterol*. Cholesterol has a steroid nucleus that is lipid soluble. Therefore, these molecules are found in the interior of the membrane lying parallel to the fatty acid chains of the phospholipids (see Figure 2.1). As such, they prevent the fatty acid chains from packing together and crystallizing which would decrease membrane fluidity.

Membrane fluidity is very important in terms of function in many cell types. For example, skeletal muscle activity involves the shortening and lengthening of the muscle fibers. Furthermore, as white blood cells leave the blood vessels and enter the tissue spaces to fight infection, they must squeeze through tiny pores in the wall of the capillary, requiring significant deformation of the cell and its membrane. Finally, in all cells, many processes that transport substances across the plasma membrane require the embedded proteins to change their conformation and move about within the bilayer. In each case, in order for the cell membrane, or the entire cell, to change its shape, the membrane must be very fluid and flexible.

Proteins are also associated with the lipid bilayer and essentially float within it. Intrinsic, or transmembrane, proteins are embedded within and span the membrane. Similar to the phospholipids, these proteins are amphipathic,

with the polar regions of the molecule extending beyond the lipid bilayer and the nonpolar region embedded within it. Extrinsic, or peripheral, proteins are found on either the internal or external surface of the membrane (see Figure 2.1). These proteins are not amphipathic and do not associate with the internal region of the membrane. The membrane proteins provide a variety of important cellular functions by forming the following structures:

- Channels
- Carrier molecules
- Enzymes
- Chemical receptors
- Antigens

Some proteins may form *channels* through the cell membrane, which allow small water-soluble substances, such as ions, to enter or leave the cell. These channels are quite specific and allow only one type of ion to pass through it (e.g., sodium channels, calcium channels). Other proteins may serve as *carrier molecules* that selectively transport larger water-soluble molecules, such as glucose or cellular products, across the membrane. *Enzymes*, which regulate specific chemical reactions, are extrinsic proteins and are found on either the internal (e.g., adenylate cyclase) or the external (e.g., acetylcholinesterase) surfaces of the membrane. *Chemical receptors* are found on the outer surface of the cell membrane and selectively bind with various endogenous molecules such as neurotransmitters and hormones as well as drugs. It is through receptor activation that many substances that are unable to enter the cell and cause a direct intracellular effect may indirectly influence intracellular activity without actually crossing the membrane. Other proteins found on the external surface of the plasma membrane are *antigens*. These molecules serve as cell "markers" that allow the body's immune system to distinguish between our own cells and foreign cells or organisms, such as bacteria and viruses.

The plasma membrane contains a small amount of *carbohydrate* (2% to 10% of the mass of the membrane) found predominantly on the outer surface. This carbohydrate is found attached to most of the protein molecules, forming glycoproteins, and to some of the phospholipid molecules (<10%), forming glycolipids. Consequently, the external surface of the cell has a carbohydrate coat, or glycocalyx.

These carbohydrate moieties have several important functions including the following:

- Repel negatively charged substances; many of the carbohydrates are negatively charged, creating an overall negative charge on the surface

of the cell which repels negatively charged extracellular molecules and also helps to keep red blood cells apart from each other.
- Cell-to-cell attachment; the glycocalyx of one cell may attach to the glycocalyx of another cell which causes the cells to become attached.
- Receptors; carbohydrates may also serve as specific membrane receptors for extracellular substances, such as hormones.
- Immune reactions; carbohydrates play a role in the ability of cells to distinguish between "self" cells and foreign cells.

PHARMACY APPLICATION:
HYDROPHILIC DRUGS BIND TO RECEPTORS

Many substances within the body, including neurotransmitters and hormones, are hydrophilic and, therefore, are incapable of entering the cells to carry out their effects directly. Instead, they bind to their specific receptors on the cell surface. This receptor binding then elicits a series of intracellular events that alter cell function and cell metabolism. As such, there are many instances where it would be clinically advantageous to either enhance or inhibit these activities. Therefore, drugs may be designed to bind to these specific receptors. A drug that binds to and stimulates a receptor, mimicking the action of the endogenous chemical substance, is referred to as a receptor *agonist*. An example is albuterol sulfate, a selective β_2-adrenergic receptor agonist. Stimulation of β_2-adrenergic receptors on airway smooth muscle causes dilation of the airways in a patient experiencing an asthmatic attack and relieves the patient's wheezing. Conversely, a drug that binds to and blocks a receptor, preventing the action of the endogenous substance, is referred to as a receptor *antagonist*. An example in this case is cimetidine hydrochloride, which inhibits histamine H_2 receptors on parietal cells in the stomach. Because histamine H_2 receptor stimulation leads to gastric acid secretion, blockade of these receptors with an antagonist reduces acid secretion. This drug, which is the active ingredient in Pepcid® and similar medications, may be used to treat patients with a peptic ulcer or gastroesophageal reflux disease (GERD).

2.3 Membrane transport

The lipid bilayer arrangement of the plasma membrane renders it semipermeable. Uncharged or nonpolar molecules, such as oxygen, carbon dioxide, and fatty acids, are lipid soluble and may permeate through the membrane quite readily. Charged or polar molecules, such as glucose, proteins, and ions, are water soluble and are impermeable, unable to cross the membrane

unassisted. These substances require protein channels or carrier molecules to enter or leave the cell.

The following are mechanisms by which substances may cross the plasma membrane:

- Passive diffusion
- Osmosis
- Mediated transport

2.4 Passive diffusion through the membrane

Molecules and ions are in constant motion, and the velocity of their motion is proportional to their temperature. This passive movement of molecules and ions from one place to another is referred to as *diffusion*. When a molecule is unevenly distributed across a permeable membrane with a higher concentration on one side and a lower concentration on the opposite side, there is said to be a *concentration gradient* or a concentration difference. Although all of the molecules are in motion, there will be a tendency for a greater number of molecules to move from the area of high concentration toward the area of low concentration. This uneven movement of molecules is referred to as *net diffusion*. The net diffusion of molecules continues until the concentrations of the substance on both sides of the membrane are equal and the subsequent movement of molecules through the membrane is in a *dynamic equilibrium*. In other words, the number of molecules moving in one direction across the membrane is equal to the number of molecules moving in the opposite direction. At this point, although the diffusion of molecules continues, there is no further *net* diffusion.

Many molecules diffuse across membranes in our body moving from one compartment to another. For example, oxygen and glucose diffuse from the extracellular fluid (higher concentration) into the intracellular fluid (lower concentration) where the cell uses these substances to make adenosine triphosphate (ATP) (energy). The movement of a substance across a membrane is also referred to as *flux*. Specifically, the movement of a substance *into* the cell is referred to as *influx*, and the movement of a substance *out of* the cell is referred to as *efflux*.

The rate of diffusion of a substance is influenced by several factors (see Table 2.1). It is proportional to the concentration gradient, the permeability of the membrane, and the surface area of the membrane. For example, as the permeability of the membrane increases, the rate of diffusion of a given substance across the membrane increases. Diffusion is inversely proportional to the molecular weight of the substance and the thickness of the membrane. Larger molecules diffuse more slowly.

The movement of ions, in particular, depends not only on a concentration gradient but also on an *electrical gradient*. Opposite charges are attracted

Table 2.1 Factors That Influence the Rate of Diffusion of a Substance

Factor	Rate of Diffusion
↑ Concentration gradient	↑
↑ Permeability of the membrane	↑
↑ Surface area of the membrane	↑
↑ Molecular weight of the substance	↓
↑ Thickness of the membrane	↓

to each other. Therefore, positively charged ions (cations) are attracted to a negatively charged area, and negatively charged ions (anions) are attracted to a positively charged area. This electrical attraction enhances diffusion of the ions. On the other hand, ions of a similar charge tend to repel each other and oppose diffusion. These two forces, the concentration gradient and the electrical gradient, may work together to cause greater diffusion of ions. On the other hand, these forces may oppose each other and, thereby, limit the net diffusion of the ion (see Figure 2.2).

2.5 Osmosis

Water is a small, polar molecule that can easily diffuse across plasma membranes through small intermolecular spaces. This movement of water across cell membranes in some tissues (e.g., kidney) may be significantly facilitated by the presence of a group of membrane proteins that form channels referred to as *aquaporins.*

Osmosis is the net movement of water through a semipermeable membrane down its own concentration gradient from an area of high water concentration to an area of low water concentration. In other words, water moves toward an area of higher *solute* concentration. The solute particles may be thought of as "drawing" the water toward them. Therefore, the *osmotic pressure* of a solution is the pressure or force by which water is drawn into the solution through a semipermeable membrane. The magnitude of this pressure depends on the number of solute particles present. An increase in the number of particles in the solution results in an increase in the osmotic pressure and, therefore, an increase in the movement of water toward it.

The plasma membrane is *semipermeable,* as it is not permeable to all solute particles present. As a result, it maintains a concentration difference for many ions and molecules across itself. Water, however, crosses the membrane freely in either direction. The movement of water in and out of the cell will occur whenever there is a difference in osmotic pressure between the intracellular fluid and the extracellular fluid. For example, an increase in the osmotic pressure of the extracellular fluid (more solute, lower water

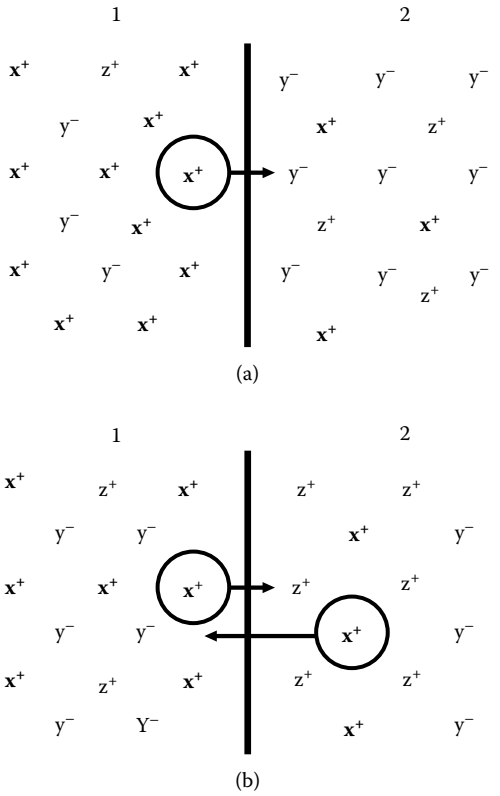


Figure 2.2 Diffusion of ions. The solutions on both side 1 and side 2 contain X^+, Y^-, and Z^+. Assume that the membrane is permeable only to X^+. (a) The concentration of X^+ is greater on side 1 than on side 2. The number of negative charges is greater on side 2 than on side 1. Therefore, there will be net diffusion of X^+ from side 1 to side 2 due to both the concentration gradient and the electrical gradient. Both gradients work together promoting the diffusion of X^+ across the membrane. (b) The concentration of X^+ is greater on side 1 than on side 2. The number of negative charges is greater on side 1 than on side 2. Therefore, X^+ will tend to diffuse from side 1 to side 2 due to the concentration gradient. However, X^+ will also tend to diffuse from side 2 to side 1 due to the electrical gradient. In this case, the two forces promoting diffusion oppose each other and limit the net diffusion of X^+ across the membrane.

concentration) will cause water to leave the cell by osmosis. The resulting decrease in the fluid volume of the cell may lead to cellular dehydration. On the other hand, a decrease in the osmotic pressure in the extracellular fluid (less solute, higher water concentration) will cause water to enter the cell by osmosis. The increase in the fluid volume of the cell may lead to an increase in *hydrostatic pressure* within the cell. An excessive increase in hydrostatic pressure may cause the cell to burst.

PHARMACY APPLICATION: INTRAVENOUS SOLUTIONS

Intravenous (i.v.) solutions are commonly administered to patients in hospitals, long-term care facilities, and ambulances. They are used primarily to replace body fluids and to serve as a vehicle for injecting drugs into the body. The advantages of this pharmaceutical dosage form include the rapid onset of action, the ability to treat patients unable to take medication orally, and the ability to administer a medication unavailable in any other dosage form.

Intravenous solutions must be isotonic (same osmotic pressure) with red blood cells. If red blood cells were to be exposed to an i.v. solution that was hypotonic (lower osmotic pressure), water from this solution would move into the cells causing them to swell and possibly lyse. If red blood cells were to be exposed to an i.v. solution that was hypertonic (higher osmotic pressure), water would move out of the cells causing them to dehydrate and shrink. Both of these conditions would damage the red blood cells and disrupt function. Intravenous injections are often prepared with 0.9% sodium chloride or 5% dextrose, both of which are approximately isotonic with red blood cells. These solutions prevent any unwanted osmosis, or movement of water, into or out of the red blood cells.

Patient discomfort is another important consideration. The stinging caused by a hypotonic or hypertonic i.v. solution is not experienced with one that is isotonic.

2.6 *Mediated transport*

In the process of *mediated transport*, carrier proteins embedded within the plasma membrane assist in the transport of larger, polar molecules into or out of the cell. When a given substance attaches to a specific binding site on the carrier protein, the protein undergoes a conformational change such that this site with the bound substance moves from one side of the plasma membrane to the other. The substance is then released.

Mediated transport displays three important characteristics influencing its function:

- Specificity
- Competition
- Saturation

Carrier proteins display a high degree of *specificity*. In other words, each of these proteins may bind only with select substances that "fit" into its binding site. Another characteristic is *competition*. Different substances that have similar chemical structures may be able to bind to the same carrier protein

and, therefore, compete for transport across the membrane. The third characteristic displayed by mediated transport is *saturation*. The greater the number of carrier proteins being utilized at any given time, the greater the rate of transport. Therefore, initially as the concentration of a substance increases, the number of active carrier molecules increases, and the rate of transport increases. However, there are a finite number of carrier proteins in a given cell membrane. Once all of these proteins are being utilized in the transport process, any further increase in the concentration of the substance no longer increases the rate of transport as it has reached its maximum. At this point the process is saturated.

There are two forms of mediated transport:

1. Facilitated diffusion
2. Active transport

With *facilitated diffusion*, carrier proteins move across the membrane in either direction and will transport a substance down its concentration gradient. In other words, substances are moved from an area of high concentration to an area of low concentration. This process is passive and requires no energy. An example of a substance transported by facilitated diffusion is glucose, which is a large, polar molecule. Because cells are constantly utilizing glucose to form ATP, there is a persistent concentration gradient for diffusion into the cell.

With *active transport*, energy is expended to move a substance against its concentration gradient from an area of low concentration to an area of high concentration. This process is used to accumulate a substance on one side of the plasma membrane or the other. The most common example of active transport is the sodium-potassium pump that involves the activity of Na^+-K^+ ATPase, an intrinsic membrane protein. For each ATP molecule hydrolyzed by the Na^+-K^+ ATPase, this pump moves three Na^+ ions out of the cell and two K^+ ions into the cell. As will be discussed further in the next chapter, the activity of this pump contributes to the difference in the composition of the extracellular fluid and the intracellular fluid which is necessary for nerve cells and muscle cells to function.

Medical terminology

Agonist (ăg′ŏn-ĭst): A chemical substance or drug that binds to a receptor and elicits the same reaction as an endogenously produced substance or neurotransmitter.

Amphipathic (ăm-fē-păth′ĭk): A molecule with a polar (hydrophilic) region and a nonpolar (hydrophobic) region.

Antagonist (ăn-tăg′ō-nĭst): A chemical substance or drug that binds to a receptor and blocks or prevents the effects of an endogenously produced substance or neurotransmitter.

Antigen (ăn′tĭ-jĕn): A protein marker on the surface of a cell that identifies the cell as "self" or "nonself" (foreign).

Diffusion (dĭ-fū′zhŭn): Movement of molecules or ions from a region of high concentration to a region of low concentration.

Efflux (ē′flŭx): Outward movement of a substance.

Equilibrium (ē-kwĭl-ĭ′brē-ŭm): A state of balance.

Hydrophilic (hī-drō-fĭl′ĭk): Attracted to water.

Hydrophobic (hī-drō-fō′bĭk): Repelled by water.

Hydrostatic pressure (hī″drō-stăt′ĭk prĕsh′ŭr): Pressure exerted by fluid.

Influx (ĭn′flŭx): Inward movement of a substance.

Intravenous (ĭn-tră-vē′nŭs): Within a vein.

Lipid bilayer (lĭp′ĭd bī′lā-er): The cell membrane consisting of two layers of phospholipids oriented such that the hydrophilic portion of the molecules faces outward and the hydrophobic portion of the molecules is in the center.

Lipophilic (lī-pō-fĭl′-ĭk): Attracted to fats or lipids.

Osmosis (ŏz-mō′sĭs): Movement of water through a semipermeable membrane from a region of high water concentration to a region of low water concentration.

Osmotic pressure (ŏz-mŏt′ĭk prĕsh′ŭr): A measure of the tendency for a solution to gain water by osmosis when separated from another solution of lower osmolarity by a semipermeable membrane.

Bibliography

AHFS Drug Information 2000, American Society of Health-System Pharmacists, Bethesda, MD, 2000.

Bell, D. R., *Core Concepts in Physiology*, Lippincott-Raven, Philadelphia, PA, 1998.

Costanzo, L. *Physiology*, 3rd ed., W.B. Saunders, Philadelphia, PA, 2006.

Fox, S. *Human Physiology,* 9th ed., McGraw-Hill, New York, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hunt, M. L., Jr., *Training Manual for Intravenous Admixture Personnel*, 5th ed., Baxter Healthcare, Deerfield, IL, 1995.

Lombard, J. H., and Rusch, N. J., Cells, Nerves and Muscles, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA., 1999, chap. 1.

Rhoades, R., and Pflanzer, R., *Human Physiology*, 4th ed., Brooks/Cole, Pacific Grove, CA, 2003.

Sherwood, L. *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E., Raff, H., and Strang, K., *Vander's Human Physiology, The Mechanisms of Body Function,* 10th ed., McGraw-Hill, New York, 2006.

chapter three

Membrane potential

Study objectives

- Define *membrane potential*
- Compare the distribution and permeability differences of ions across the cell membrane
- Describe how differences in ion distribution and permeability contribute to the resting membrane potential
- Describe how the resting membrane potential is developed and maintained
- Explain the role of the Na^+-K^+ ATPase pump in this process

3.1 Introduction

Both the intracellular fluid and the extracellular fluid are electrically neutral solutions, in that each has an equal number of positively charged ions and negatively charged ions. A simple but important concept is that these opposite charges are attracted to each other, and ions of the same charge repel each other. In an unstimulated or resting cell, there is a slight accumulation of negative charges (–) on the internal surface of the plasma membrane which are attracted to an equal number of positive charges (+) that have accumulated on the external surface of the membrane. Therefore, at rest, all cells are electrically *polarized* — that is, the inside of the cell is slightly negative relative to the outside. This separation of charge across the plasma membrane is referred to as the *membrane potential*. The magnitude of the membrane potential depends primarily on the number of opposite charges that are separated by the membrane. The greater the separation of charge then, the greater is the membrane potential. Because the actual number of charges involved is quite small, the potential is measured in millivolts (mV). Furthermore, the sign (+ or –) of the potential is defined by the predominant charge on the internal surface of the cell membrane. Therefore, the membrane potential under resting conditions is negative. As will be discussed, nerve cells and muscle cells rely on changes in this membrane potential for their functions. In other words, changes in the membrane potential convey information to these types of cells.

Table 3.1 Concentration and Permeability of Ions Responsible for the Membrane Potential in a Resting Nerve Cell

	Concentration (millimoles/liter)		
Ion	**Extracellular Fluid**	**Intracellular Fluid**	**Relative Permeability**
Na^+	150	15	1
K^+	5	150	50–75
A^-	0	65	0

3.2 *Development of the resting membrane potential*

In a typical unstimulated neuron, the *resting membrane potential* is approximately –70 mV. The development of this potential depends on the *distribution* and *permeability* of three ions: sodium (Na^+), potassium (K^+), and anions (A^-) (see Table 3.1 and Figure 3.1). These ions are unevenly distributed between the intracellular fluid (ICF) and the extracellular fluid (ECF), and they each have a different degree of permeability across the plasma membrane. Sodium ions are found in a greater concentration in the ECF, K^+ ions are found in a greater concentration in the ICF, and A^- refers to large anionic proteins found only within the cell. Under resting conditions, most mammalian plasma membranes are approximately 50 to 75 times more permeable to K^+ ions than they are to Na^+ ions. The anions are impermeable at all times. It is due to these underlying conditions that the resting membrane potential is generated and maintained.

When permeable, the movement of Na^+ and K^+ ions in and out of the cell depends on two factors:

1. Concentration gradient
2. Electrical gradient

Consider the condition where the membrane is permeable only to potassium. Because potassium is in a greater concentration inside the cell, the K^+ ions initially diffuse out of the cell down their *concentration gradient*. As a result, an excess of these positively charged ions would accumulate in the ECF along the external surface of the plasma membrane. The impermeable A^- ions, which are attracted to these positive charges, would remain inside the cell along the internal surface of the plasma membrane. This outward movement of positive charges creates a membrane potential that is negative because the inside of the cell is now negative relative to the outside. However, as the positively charged K^+ ions continue to diffuse outward, an electrical gradient begins to develop which also influences the diffusion of K^+ ions. The K^+ ions that had moved out of the cell down their concentration gradient have caused an excess of (+) charges to accumulate on the external surface of the membrane. Because like charges repel each other, these initial (+) charges would begin to repel any additional K^+ ions and oppose the further

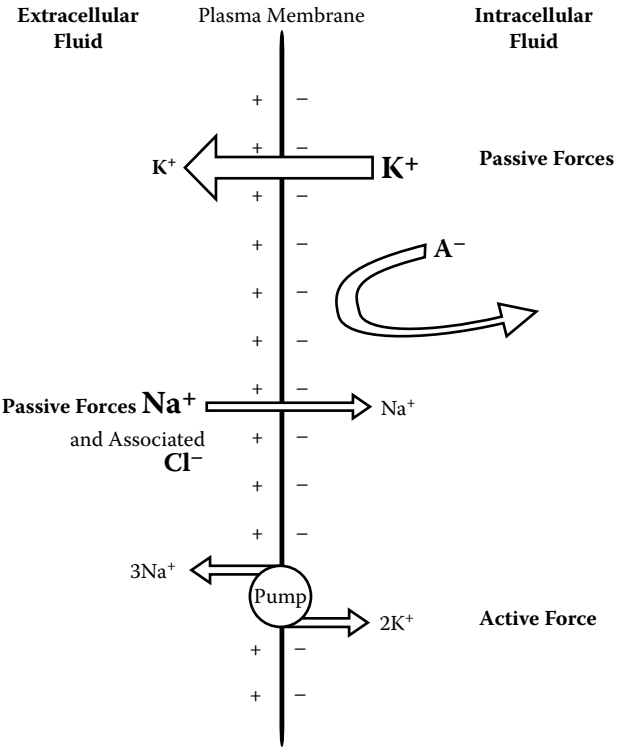


Figure 3.1 Generation of the resting membrane potential. Under resting conditions, potassium (K^+) is significantly more permeable than sodium (Na^+), and the negatively charged intracellular anions (A^-) are impermeable. Therefore, the abundant outward movement of K^+ ions down their concentration gradient exerts a powerful effect, driving the membrane potential toward the equilibrium potential for potassium (–90 mV). However, the slight inward movement of Na^+ ions, which would tend to drive the membrane potential toward the equilibrium potential for sodium (60 mV), renders the membrane potential somewhat less negative than –90 mV. The balance of these two opposing effects results in a resting membrane potential of –70 mV in a typical neuron. The maintenance of the concentration differences for sodium and potassium is due to the continuous activity of the Na^+-K^+ pump.

movement of (+) charges outward. Instead, the positively charged K^+ ions are now electrically attracted to the negatively charged A^- ions remaining inside the cell. At this point, K^+ ions not only diffuse outward down their concentration gradient, K^+ ions also diffuse into the cell down their *electrical gradient*. Eventually, the subsequent force that moves K^+ ions inward exactly balances the initial force that moved K^+ ions outward, and there is no further net diffusion of potassium. The membrane potential at this point has reached the *equilibrium potential for* K^+ (E_{K^+}) and is equal to –90 mV. Therefore, when the permeability of the plasma membrane to potassium is high compared to that of sodium, the membrane potential approaches –90 mV.

Next consider the condition where the membrane is permeable only to sodium. Because sodium is in a greater concentration outside of the cell, the Na^+ ions initially diffuse into the cell down their concentration gradient. As a result, an excess of these positively charged ions accumulates in the ICF along the internal surface of the plasma membrane, and an excess of negative charges in the form of the impermeable extracellular anion, chloride (Cl^-), remains outside the cell along the external surface of the plasma membrane. This inward movement of positive charges creates a membrane potential that is positive because the inside of the cell is now positive relative to the outside. However, as the positively charged Na^+ ions continue to diffuse inward, once again an electrical gradient develops. The initial (+) charges that have accumulated in the ICF begin to repel any additional Na^+ ions and oppose the further movement of (+) charges inward. Instead, the positively charged Na^+ ions are now attracted to the negatively charged Cl^- ions remaining outside the cell. Eventually, the initial force moving Na^+ ions inward down their concentration gradient is exactly balanced by the subsequent force moving Na^+ ions outward down their electrical gradient, and there is no further net diffusion of sodium. The membrane potential at this point has reached the *equilibrium potential for* Na^+ (E_{Na^+}) and is equal to 60 mV. Therefore, when the permeability of the plasma membrane to sodium is high compared to that of potassium, the membrane potential approaches 60 mV.

At any given time, the membrane potential is closer to the equilibrium potential of the ion that is more permeable. Under normal resting conditions, both Na^+ ions and K^+ ions are permeable; however, potassium is significantly (50 to 75 times) more permeable than sodium. Therefore, a large number of K^+ ions diffuse outward and a very small number of Na^+ ions diffuse inward down their concentration gradients. As a result, the comparatively copious outward movement of K^+ ions exerts a powerful influence on the value of the resting membrane potential, driving it toward its equilibrium potential of –90 mV. However, the slight inward movement of Na^+ ions which would tend to drive the membrane potential toward its equilibrium potential of 60 mV renders the membrane potential slightly less negative than –90 mV. The balance of these two opposing effects results in a resting membrane potential in a typical neuron of –70 mV (see Figure 3.1).

The Na^+-K^+ pump also plays a vital role in this process. For each molecule of ATP expended, three Na^+ ions are pumped out of the cell into the ECF and two K^+ ions are pumped into the cell into the ICF. The result is the unequal transport of positively charged ions across the membrane such that the outside of the cell becomes more positive compared to the inside of the cell, or in other words, the inside of the cell is more negative compared to the outside of the cell. Therefore, the activity of the pump has a small direct contribution to the generation of the resting membrane potential.

The other even more important effect of the Na^+–K^+ pump is that it maintains the concentration differences for sodium and potassium by accumulating Na^+ ions outside of the cell and K^+ ions inside of the cell. As previously

discussed, it is the passive diffusion of these ions down their concentration gradients that is predominantly responsible for generating the resting membrane potential. Sodium diffuses inward, and potassium diffuses outward. The continuous activity of the pump returns the Na^+ ions to the ECF and the K^+ ions to the ICF. Therefore, it can be said that the pump also has an indirect contribution to the generation of the resting membrane potential due to the maintenance of the concentration differences for the relevant ions.

Medical terminology

Equilibrium potential (ē″kwĭ-lĭb′rē-ŭm pō-tĕn′shăl): State where the concentration and electrical gradients are balanced, and therefore, the efflux and influx of a given ion are also equal such that there is no further net diffusion.

Membrane potential (mĕm′brān pō-tĕn′shăl): The electrical difference between the inside and the outside of the cell.

Polarization (pō″lăr-ī-zā′shŭn): Condition of the resting cell where the inside is negative relative to the outside; separation of charge.

Bibliography

Lombard, J. H., and Rusch, N. J., Cells, Nerves and Muscles, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 1.

Sherwood, L. *Human Physiology from Cells to Systems*, 4th ed., Brooks/Cole, Pacific Grove, CA, 2001.

chapter four

Electrical signals

Study objectives

- Distinguish between depolarization, hyperpolarization, and repolarization
- Compare and contrast graded potentials and action potentials
- Describe the process of local current flow
- Explain the mechanism by which action potentials are generated
- Understand the function of sodium and potassium voltage-gated channels
- Distinguish between the absolute refractory period and the relative refractory period
- Describe the process of saltatory conduction
- Explain the functional significance of myelin
- Explain why the conduction of the action potential is unidirectional

4.1 Introduction

Both nerve cells and muscle cells rely on changes in their membrane potentials in order to carry out their activities. In this chapter, the focus will be on the nerve cell or *neuron*; however, many of the same principles will also apply to muscle. The function of neurons is to convey information to other cells in the form of electrical signals. There are two types of electrical signals transmitted by neurons: *graded potentials* and *action potentials*. These signals occur due to ion flux (movement) across the plasma membrane. A given stimulus will cause its effect by altering the permeability to one or more ions. The involved ions will then diffuse into or out of the cell according to their concentration and electrical gradients, causing a change in the membrane potential.

4.2 Graded potentials

Graded potentials are short-distance signals (see Table 4.1). They are local changes in the membrane potential that occur at *synapses* where one neuron comes into contact with another neuron. The magnitude of these signals varies with the strength of the stimulus. As the intensity of the stimulus increases, the number of ions diffusing across the cell membrane increases, and the magnitude of the change in the membrane potential increases. This change may be in either direction such that the membrane potential may

Table 4.1 Distinguishing Features of Graded Potentials and Action Potentials

Graded Potentials	Action Potentials
Short-distance signals	Long-distance signals
Magnitude is stimulus dependent	Magnitude is constant (all-or-none phenomenon)
Signal travels by local current flow	Signal travels by local current flow or by saltatory conduction
Magnitude of signal dissipates as it moves away from the site of stimulation	Magnitude of signal is maintained along entire length of neuron
Initiated at synapses (where one neuron comes into contact with another neuron)	Initiated at axon hillock
Result in depolarization or hyperpolarization	Depolarization only

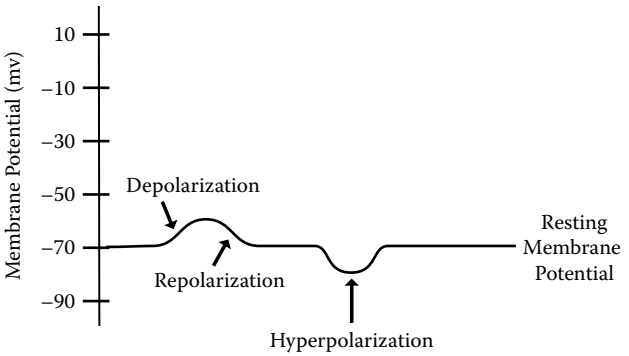


Figure 4.1 Types of changes in membrane potential. The resting membrane potential in a typical neuron is –70 mV. Movement of the membrane potential toward zero (less negative) is referred to as depolarization. The return of the membrane potential to its resting value is referred to as repolarization. Movement of the membrane potential farther away from zero (more negative) is referred to as hyperpolarization.

become more or less negative compared to the resting membrane potential (see Figure 4.1). *Depolarization* occurs when the membrane potential becomes less negative, moving toward zero. As will be discussed, depolarization makes the neuron more excitable. *Hyperpolarization* occurs when the membrane potential becomes more negative, moving farther away from zero. Hyperpolarization tends to make the neuron less excitable. These signals are transient or short-lived. Once the stimulus has been removed, the membrane potential returns to its resting state. Following a depolarization, the membrane is said to undergo *repolarization* and a return to its resting potential.

The mechanism by which the signal is transmitted along the cell membrane is referred to as *local current flow* or the movement of positively charged ions. In the area of a stimulus causing a depolarization, the inside of the cell

becomes positive (less negative) relative to the outside of the cell. Because opposite charges attract, the positive (+) charges in this area are attracted to and move toward the negative (–) charges on the adjacent areas of the internal surface of the cell membrane. As a result, these adjacent areas become depolarized due to the presence of these positive (+) charges. This process continues, and the electrical signal travels along the cell membrane away from the initial site of the stimulus. However, these graded or local potentials travel only short distances. The cell membrane is not well insulated, and the current (positive charges) tends to drift away from the internal surface of the cell membrane. Consequently, as the signal travels along the membrane, the number of positive (+) charges causing the depolarization of the next region of membrane continually decreases; therefore, the magnitude of the depolarization also decreases. The farther away from the initial site of stimulation, the smaller the magnitude of the signal until it eventually dies out.

4.3 *Action potentials*

Action potentials are long-distance electrical signals (see Figure 4.2). These signals travel along the entire neuronal membrane. Unlike graded potentials where the magnitude of the signal steadily dissipates, the magnitude of the action potential is maintained throughout the length of the axon. Furthermore, in contrast to graded potentials whose magnitude is stimulus dependent, action potentials are always the same size. If a stimulus is strong enough to depolarize the membrane to a critical level referred to as *threshold*, then the membrane continues to depolarize on its own, independent of the stimulus. Typically, the threshold potential is approximately 20 mV less negative than the resting membrane potential. Once threshold is reached, the continued depolarization takes place automatically. This is due to the diffusion of ions according to their concentration and electrical gradients and not due to the original stimulus.

Given that action potentials are always of a similar magnitude, how can stimuli of varied strengths be distinguished? A suprathreshold stimulus, one that is larger than necessary to depolarize the membrane simply to threshold, does not produce a larger action potential, but it does increase the *frequency* at which action potentials are generated. In other words, a stronger stimulus will trigger a greater number of action potentials per second.

The generation of an action potential involves changes in permeability to both Na^+ ions and K^+ ions through *voltage-gated ion channels*. However, these permeability changes take place at slightly different times (see Figure 4.2). Voltage-gated ion channels open and close in response to changes in membrane potential. Initially, some stimulus will cause the membrane to depolarize toward threshold. When this occurs, voltage-gated Na^+ channels begin to open. As a result, Na^+ ions enter the cell down their concentration and electrical gradients. (Recall that at this point, Na^+ is not only in a greater concentration outside the cell, but the inside of the cell is negative relative to the

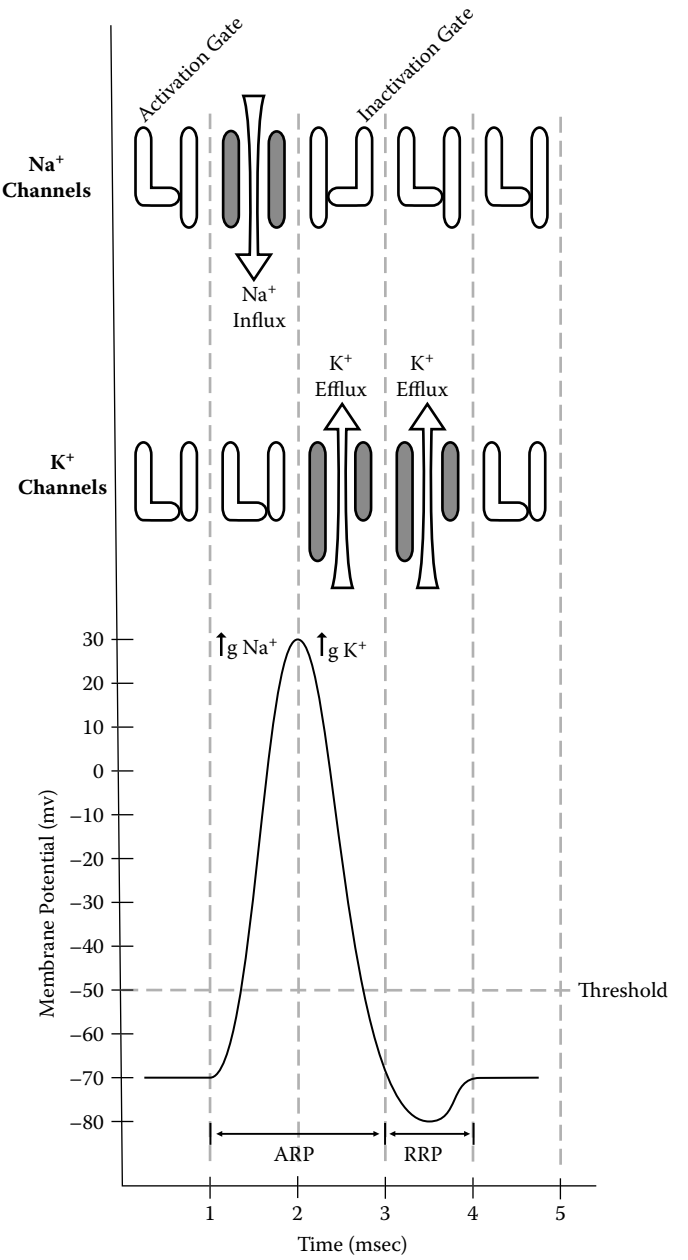


Figure 4.2 The action potential. At the resting membrane potential (–70 mV), most ion channels are in their resting state, closed but capable of opening. When the neuron is stimulated and depolarized, the activation gates of the voltage-gated Na^+ channels open, permitting the influx of Na^+ ions and further depolarization toward threshold. At the threshold potential, all voltage-gated Na^+ channels are open, resulting in the

outside as well.) The influx of Na^+ ions causes further depolarization, resulting in the opening of more voltage-gated Na^+ channels, the continued influx of Na^+ ions, and so on. As mentioned in Chapter 1, this is an example of a positive-feedback mechanism. This process of depolarization and opening of voltage-gated Na^+ channels continues until the membrane is depolarized to threshold. At this point, all of the Na^+ channels are open, and there is a very rapid and abundant influx of Na^+ ions. At this time, the permeability to Na^+ ions is approximately 600 times greater than normal. This ion flux causes the upward swing or the "spike" of the action potential. During this phase of the action potential, the membrane reverses polarity due to the marked influx of positive (+) charges, and the membrane potential at the peak of the action potential is 30 mV.

Approximately 1 msec after the Na^+ channels open, they close. This prevents any further diffusion of positive (+) charges into the cell. At the same time, voltage-gated K^+ channels open and K^+ ions leave the cell down their concentration and electrical gradients. (At this point, K^+ is not only in a greater concentration inside the cell, the inside of the cell is positive relative to the outside.) During this phase of the action potential, the permeability to K^+ ions is approximately 300 times greater than normal. This efflux of positive (+) charges causes the membrane to repolarize back toward the resting membrane potential.

Sodium channels open more rapidly than the K^+ channels because they are more voltage-sensitive and just a small depolarization is sufficient to open them. Larger changes in membrane potential associated with further cell excitation are required to open the less-voltage-sensitive K^+ channels. Therefore, the increase in the permeability to K^+ ions occurs later than that of Na^+ ions. This is functionally significant because if both types of ion channels opened concurrently, then the change in membrane potential that would occur due to Na^+ ion influx would be cancelled out by K^+ ion efflux and the action potential could not be generated.

To more fully understand the mechanism by which the action potential is generated, further explanation concerning the structure and activity of

Figure 4.2 (continued) "spike" of the action potential. Approximately 1 msec after the activation gates open, the inactivation gates of the Na^+ channels close. In addition, the activation gates of the K^+ channels open, resulting in the efflux of K^+ ions and the repolarization of the neuron. The protracted increase in K^+ ion permeability results in the after-hyperpolarization. It is during this time when the membrane potential in the neuron is farther away from threshold that the cell is in its relative refractory period (RRP) and a larger than normal stimulus is needed to generate an action potential. The absolute refractory period (ARP) begins when the voltage-gated Na^+ channels have all become activated and continues through the inactivation phase. During this time, there can be no further Na^+ ion influx, and no new action potentials can be generated. Voltage-gated Na^+ channels return to their resting state (activation gates closed, inactivation gates open) when the membrane potential approaches the resting membrane potential of the neuron.

the voltage-gated ion channels is necessary. A *voltage-gated Na^+ channel* has two different gates: the *activation gate* and the *inactivation gate*. At the resting membrane potential of –70 mV in an unstimulated neuron, the activation gate is closed, and the permeability to Na^+ ions is very low. In this resting state, the channel is closed but capable of opening in response to a stimulus. When stimulated by depolarization to threshold, the activation gates open very rapidly and Na^+ ions diffuse into the cell causing the upward swing of the action potential. Once these activation gates open, the inactivation gates begin to close, although the closure of these gates takes place more slowly. Approximately 1 msec after the gates open, they begin to close. At the peak of the action potential when the inactivation gates are now all closed, these channels are no longer permeable to Na^+ ions, and they are incapable of opening regardless of further stimulation. Therefore, the Na^+ channels cannot reopen, Na^+ ions cannot enter the cell, and another action potential cannot be generated. In fact, these voltage-gated channels cannot return to their resting position and become capable of opening until the neuron has first repolarized to –70 mV from the existing action potential. This period of time beginning when all of the Na^+ channels are open and lasting through the inactivation phase of the Na^+ channels is referred to as the a*bsolute refractory period*. Regardless of the strength of the stimulus, no new action potentials can be generated. The length of this period, approximately 2 msec, limits the number of action potentials that neurons can generate to up to 500 per second.

The *voltage-gated K^+ channel* has only one gate, which is typically closed at the resting membrane potential. This gate also opens in response to depolarization of the membrane toward zero. However, unlike the activation gate of the voltage-gated Na^+ channel which opens very quickly, this gate opens very slowly so that the permeability to K^+ ions is delayed. In fact, they begin to open at approximately the same time that the inactivation gates in the Na^+ channels close. Therefore, there is a simultaneous decrease in Na^+ ion permeability and an increase in K^+ ion permeability resulting in the outward movement of positive (+) charges and a rapid repolarization.

The voltage-gated K^+ channels not only open slowly, they close slowly as well. Therefore, the increase in permeability to K^+ ions is prolonged. As a result, K^+ ions continue to exit the cell, and the membrane potential approaches the equilibrium potential for potassium. This phase of the action potential is referred to as the *after-hyperpolarization*. Because the membrane potential is now farther away from threshold, a larger than normal stimulus is necessary to cause depolarization to threshold. During this phase of hyperpolarization, it is possible but more difficult for the neuron to generate another action potential. This r*elative refractory period* lasts from the end of the absolute refractory period until the voltage-gated K^+ channels have returned to their resting state and the membrane once again returns to its resting potential.

During the course of the action potential, Na^+ ions entered the cell and K^+ ions exited the cell. In order to prevent the eventual dissipation of the concentration gradients for Na^+ ions and K^+ ions across the cell membrane over time, these substances must be returned to their original positions. The slow but continuous activity of the Na^+-K^+ pump is responsible for this function and returns Na^+ ions to the extracellular fluid and K^+ ions to the intracellular fluid.

PHARMACY APPLICATION: SEIZURES

Epilepsies affect approximately 2.5 million patients in the United States. This debilitating disease is characterized by seizures resulting from the sudden discharge of electrical activity in specific populations of brain neurons. This inappropriate electrical activity is thought to arise from the cerebral cortex. Accordingly, the symptoms experienced by a given patient are related to the region of the cortex involved.

Phenytoin (Dilantin®) is useful in the treatment of all types of partial and tonic-clonic seizures. Its mechanism of action involves the slowed rate of recovery of voltage-gated Na^+ channels. As a result, these channels remain in their inactivation state longer, and the absolute refractory period is prolonged. This effect limits the sustained, repetitive firing of the neurons and, therefore, prevents the seizures.

4.4 Conduction of the action potential

A typical neuron consists of four functional regions:

1. Cell body
2. Axon hillock
3. Axon
4. Axon terminal

First is the *cell body* with its *dendrites*, which are projections from the cell body that greatly increase the surface area. This is the site of communication and input from other neurons. These inputs result in the generation of graded potentials that travel a short distance to the *axon hillock*, the region of the cell body from which the axon arises. Following sufficient stimulation of the neuron, an action potential is generated at the axon hillock. This action potential must then be propagated or regenerated along the third portion of the neuron, the axon. The *axon*, or nerve fiber, is an elongated projection that transmits the action potential away from the cell body toward other cells. The final component of the neuron is the *axon terminal* where the neuron communicates, by way of this action potential, with another cell or cells. The conduction of the action potential along the length of the axon is the subject

of this section. Communication between a neuron and another cell is discussed in Chapter 5.

The action potential is initiated at the axon hillock (see Figure 4.3). This region is particularly excitable due to the presence of an abundance of voltage-gated Na^+ channels. As the axon hillock is stimulated by excitatory inputs, there is a marked influx of Na^+ ions, and this region of the cell membrane becomes positive inside resulting in an action potential. The rest of the axon is still at its resting membrane potential and is negative inside. As with graded potentials, this electrical signal also travels by local current flow (see Figure 4.3). The positive charges in the region of the action potential are attracted to the negative charges in the immediately adjacent region of the axonal membrane. This current flow depolarizes the new region, causing an increase in the permeability of the cell membrane to Na^+ ions through the voltage-gated ion channels. The subsequent influx of Na^+ ions further depolarizes the membrane so that it reaches threshold, and a *new action potential* is generated in this region. At the same time, the original site of action potential generation at the axon hillock repolarizes due to the efflux of K^+ ions. This process of generating new action potentials sequentially along the membrane enables the signal to maintain its strength as it travels the distance to the axon terminal.

Another mechanism of conduction of an action potential along the length of a neuron is *saltatory conduction,* and it occurs in *myelinated* axons (see Figure 4.4). Myelin is a lipid sheath wrapped around the axon at regular intervals. The myelin is not actually part of the axon itself, but instead it comes from other cells. In the central nervous system (brain and spinal cord), the myelin-forming cell is the *oligodendrocyte*, one of several types of support cells for centrally located neurons. In the peripheral nervous system (all neurons that lie outside of the central nervous system and communicate with various body parts), myelin is formed by the *Schwann cells*. The lipid of the myelin in each case comes from multiple layers of the plasma membrane of these cells as they wrap around and around the axon. This lipid provides good insulation, preventing the movement of current across the cell membrane. Without ion flux, action potentials cannot be generated in the regions covered with myelin. Instead, action potentials occur only at the breaks in the myelin sheath referred to as the *nodes of Ranvier.* These nodes are located about 1 to 2 mm apart. The flow of current from an active node "skips" down the axon to the adjacent node to cause depolarization and generate a new action potential. This transmission of the impulse from node to node is referred to as saltatory conduction, from the Latin word *saltare*, meaning "to leap."

Saltatory conduction results in a significant increase in the *velocity of conduction* of the nerve impulse down the axon compared to that of local current flow in an unmyelinated axon (see Table 4.2). The speed of conduction is directly correlated to the urgency of the information that is being conveyed by a given neuron. Nerve fibers carrying less important information, such as

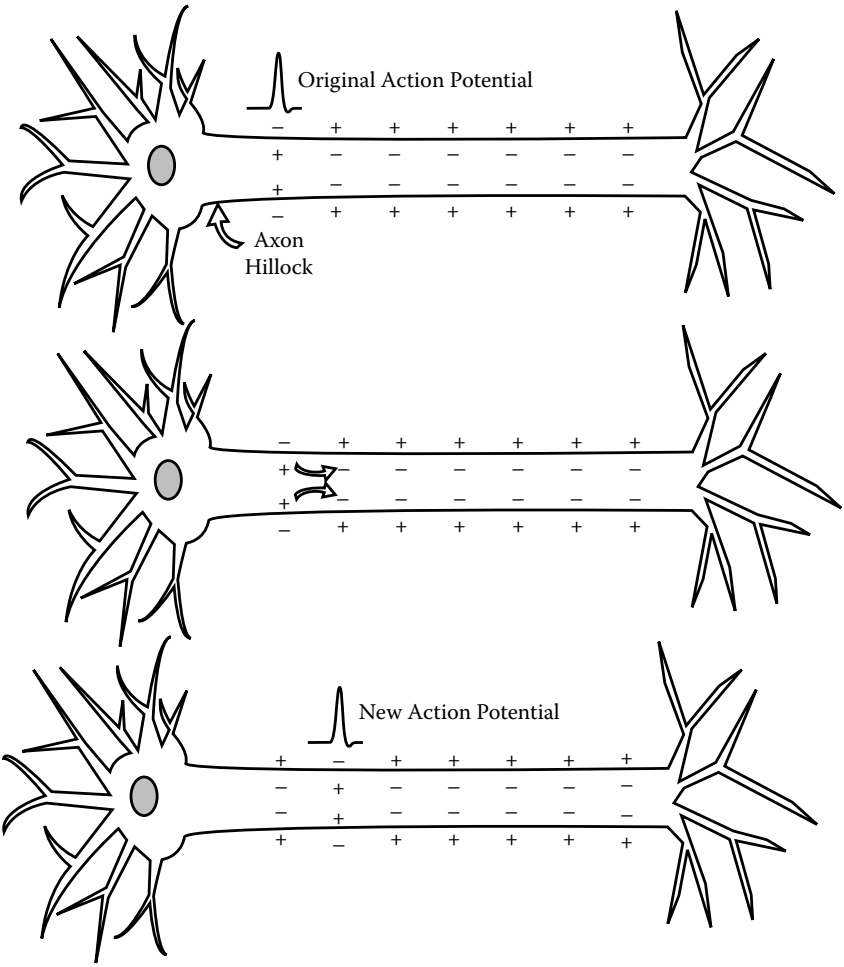


Figure 4.3 Conduction of the action potential along an axon by local current flow. Upper panel: Action potentials are generated at the axon hillock. When stimulated to threshold, this region of the membrane becomes positive (30 mV) inside relative to the outside due to the influx of Na^+ ions. The remainder of the axon is at its resting membrane potential (–70 mV). Middle panel: Because opposite charges attract, the positive (+) charges in the stimulated area are attracted to the negative (–) charges in the adjacent region of the membrane. This movement of positive (+) charges, or local current flow, depolarizes this adjacent region. Lower panel: The depolarization of the adjacent region causes the activation of voltage-gated Na^+ channels, the influx of Na^+ ions, and the generation of a new action potential. The original area of stimulation, meanwhile, has repolarized back to the resting membrane potential. This process, which is unidirectional, continues along the length of the axon.

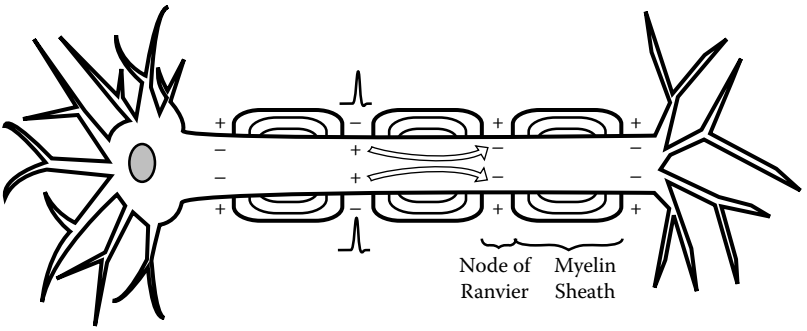


Figure 4.4 Saltatory conduction. Transmission of electrical impulses in a myelinated axon occurs by way of saltatory conduction. The myelin sheath, which is composed primarily of lipid, insulates the axon and prevents the generation of membrane potentials. Membrane potentials occur only at the gaps in the myelin sheath, referred to as the nodes of Ranvier. Therefore, transmission of the impulse, or the generation of action potentials, occurs only at the nodes.

Table 4.2 Factors That Affect the Velocity of Conduction

Factor	Velocity of Conduction
Myelination of axon (saltatory conduction)	↑
↑ Diameter of axon	↑

those regulating slow digestive processes, are unmyelinated. An example of a nerve fiber with myelin is one that innervates skeletal muscle so that movements can be executed rapidly.

The functional significance of the myelin is revealed by the neurological deficits observed in patients with *multiple sclerosis*. This disorder is caused by the demyelination of neurons in the brain, the spinal cord, and the optic nerve. The loss of myelin disrupts the normal conduction of impulses along the axons of these neurons and results in weakness, numbness, loss of bladder control, and visual disturbances.

Another advantage of the presence of myelin along an axon is that the conduction of the impulse is energetically more efficient. Because action potentials occur only at the nodes of Ranvier, fewer Na^+ ions and K^+ ions move in and out of the cell. Therefore, less metabolic energy is required to return these ions to their original positions along the cell membrane and to maintain the proper concentration gradients. In unmyelinated axons, action potentials and, therefore, ion flux occur along the entire length of the axon. These neurons expend more energy returning these ions to their original positions.

A second factor that influences the velocity of conduction of the action potential is the *diameter of the axon*. The greater the diameter of the axon, the lower is the resistance to current flow along the axon. Therefore, the impulse is conducted along large nerve fibers more rapidly. Large myelinated nerve

fibers, such as those innervating skeletal muscle, exhibit the highest conduction velocity. Small unmyelinated fibers, such as those of the autonomic nervous system innervating the heart, smooth muscle of the blood vessels and the gastrointestinal tract, and glands, conduct nerve impulses more slowly.

Conduction of the action potential along the axon is *unidirectional*. In other words, the nerve impulse travels away from the cell body and the axon hillock and toward the axon terminal only. As the current flows from the initial area of activity to the adjacent region of the axon, the new region becomes depolarized and generates an action potential. Simultaneously, the initial area has entered its absolute refractory period due to the inactivation of the voltage-gated Na^+ channels. As a result, as the current flows away from the second active area, it has no effect on the original site of activity. Instead, the current continues forward and depolarizes the next adjacent region of the axon. By the time the original site has recovered from the refractory period and is capable of being restimulated, the action potential has traveled too far along the axon to affect this site by way of local current flow. This unidirectional conduction ensures that the signal reaches the axon terminal where it can influence the activity of the innervated cell as opposed to traveling back and forth along the axon ineffectively.

PHARMACY APPLICATION: LOCAL ANESTHETICS

Pain is a protective mechanism that alerts an individual to the occurrence of tissue damage. Stimulation of nociceptors (pain receptors) alters the membrane permeability to ions, the predominant effect of which is the influx of Na^+ ions down their electrical and chemical gradients. Sufficient Na^+ ion influx results in the generation of an action potential that is then propagated along the afferent neuron to the CNS where the painful stimulus is perceived. Local anesthetics, such as lidocaine and procaine (also known as Novocain) prevent or relieve the perception of pain by interrupting the conduction of the nervous impulse. These drugs bind to a specific receptor site on the voltage-gated Na^+ channels and block ion movement through them. Without Na^+ ion influx, an action potential cannot be generated in the afferent neuron, and the signal fails to reach the CNS. In general, the action of these drugs is restricted to the site of application and becomes less effective upon diffusion of the drug away from the site of action in the nerve.

Medical terminology

Absolute refractory (rē-frăk'tō-rē) period: Condition where the neuron is completely resistant to any stimulus such that there is no change in membrane potential.

Action potential (ăk′shŭn pō-tĕn′shăl): Long-distance electrical signal that is transmitted nondecrementally along the entire length of the axon.

Axon (ăk′sŏn): Elongated process of a neuron that transmits action potentials away from the cell body.

Axon hillock (ăk′sŏn hĭl′ŏk): Small projection from the neuronal cell body from which the axon arises; initial site of action potential generation in a neuron.

Dendrite (dĕn′drīt): Process extending from the neuronal cell body which receives electrical signals from other neurons.

Depolarization (dē-pō″lăr-ĭ-zā′shŭn): When the inside of the neuron becomes less negative; a decrease in the membrane potential or separation of charge.

Graded potential (grād′ĕd pō-tĕn′shăl): Short-distance electrical signal; local change in membrane potential.

Hyperpolarization (hī-pĕr-pō″lăr-ĭ-zā′shŭn): When the inside of the neuron becomes more negative; an increase in the membrane potential or separation of charge.

Myelin (mī′ĕ-lĭn): Plasma membranes of Schwann cells (peripheral nervous system) or oligodendrocytes (central nervous system) that wrap around an axon forming an electrically insulating sheath.

Relative refractory (rē-frăk′tō-rē) period: Condition where the neuron requires a stronger stimulus to depolarize to threshold and generate an action potential.

Repolarization (rē-pō″lăr-ĭ-zā′shŭn): Return to the resting membrane potential following the depolarization of the cell membrane.

Saltatory conduction (săl′tă-tō″rē kŏn-dŭk′shŭn): Transmission of the action potential along a myelinated axon.

Synapse (sĭn′ăps): Point of contact between two neurons, typically including the axon terminal of one neuron and the cell body or dendrites of the second neuron.

Threshold (thrĕsh′ōld): Membrane potential at which an action potential may be generated.

Bibliography

Catterall, W., and Mackie, K., Local Anesthetics, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 14.

Costanzo, L., *Physiology*, 3rd ed., W.B. Saunders, Philadelphia, PA, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Lombard, J. H., and Rusch, N. J., Cells, Nerves and Muscles, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 1.

McNamara, J. O., Pharmacotherapy of the Epilepsies, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 19.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.
Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.
Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

chapter five

Synaptic transmission

Study objectives

- Describe the mechanism by which chemical synapses function
- Describe the effects of neurotransmitter binding to its receptors on the postsynaptic neuron
- Compare and contrast excitatory synapses and inhibitory synapses
- Distinguish between an excitatory postsynaptic potential (EPSP) and an inhibitory postsynaptic potential (IPSP)
- Describe how neurotransmitters are removed from the synaptic cleft
- Explain how temporal summation and spatial summation take place
- Distinguish between convergence and divergence
- Understand how pH and hypoxia affect synaptic transmission
- Describe the potential mechanisms by which drugs, toxins, and diseases affect synaptic transmission
- Explain why synaptic transmission is unidirectional
- Distinguish between an agonist and an antagonist

5.1 Introduction

The function of a neuron is to communicate or relay information to another cell by way of an electrical impulse. A *synapse* is the site where the impulse is transmitted from one cell to the next. A neuron may terminate on a muscle cell, a glandular cell, or another neuron. The discussion in this chapter will focus on neuron-to-neuron transmission. At these types of synapses, the *presynaptic neuron* transmits the impulse *toward* the synapse, and the *postsynaptic neuron* transmits the impulse *away* from the synapse. Specifically, the axon terminal of the presynaptic neuron comes into contact with the cell body or the dendrites of the postsynaptic neuron. Most neurons, particularly in the central nervous system (CNS), receive thousands of inputs. As will become evident, the transmission of the impulse at the synapse is *unidirectional*, and the presynaptic neuron influences the activity of the postsynaptic neuron only.

5.2 Chemical synapses

Most of the synapses in the nervous system are *chemical synapses* where the presynaptic neuron and the postsynaptic neuron are not in direct contact but instead are separated by a narrow (20 to 50 nanometers) space called the *synaptic cleft*. This space prevents the direct spread of the electrical impulse

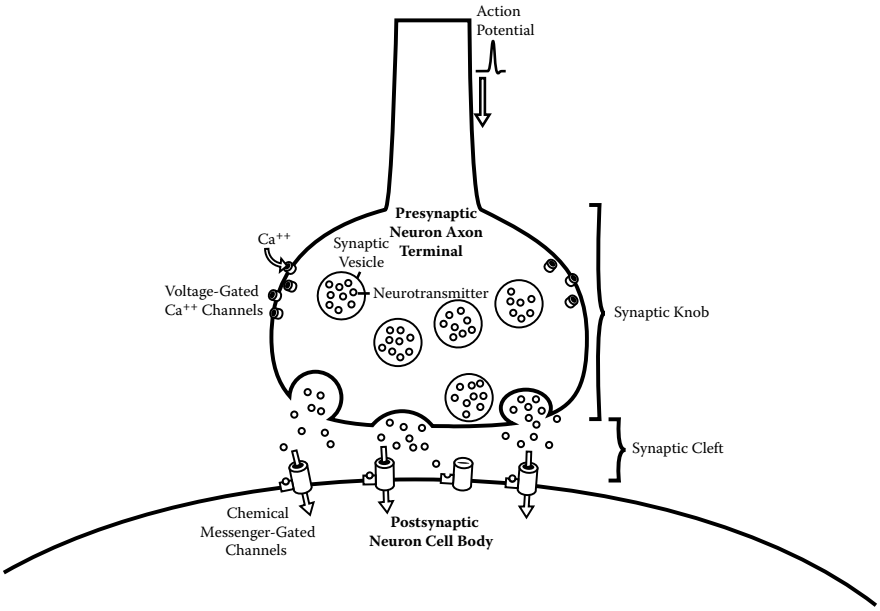


Figure 5.1 Mechanism of action at a chemical synapse. The arrival of an action potential at the axon terminal causes voltage-gated Ca^{++} channels to open. The resulting increase in concentration of Ca^{++} ions in the intracellular fluid facilitates the exocytosis of the neurotransmitter into the synaptic cleft. Binding of the neurotransmitter to its specific receptor on the postsynaptic neuron alters the permeability of the membrane to one or more ions, thus causing a change in the membrane potential and the generation of a graded potential in this neuron.

from one cell to the next. Instead, a chemical referred to as a *neurotransmitter* is released from the presynaptic neuron. The neurotransmitter diffuses across the synaptic cleft, binds to its specific receptor, and alters the electrical activity of the postsynaptic neuron.

The mechanism of action of a chemical synapse is shown in Figure 5.1. The axon terminal broadens to form a swelling referred to as the *synaptic knob*. Within the synaptic knob are many mitochondria that supply the energy for synaptic function. There are also many *synaptic vesicles* that store the preformed neurotransmitter. Also found in the membrane of the synaptic knob are *voltage-gated Ca^{++} channels*. When the electrical impulse, or action potential, has been transmitted along the length of the axon and reaches the axon terminal, the accompanying change in voltage, or depolarization, causes the voltage-gated Ca^{++} channels to open. As calcium is in a greater concentration in the extracellular fluid compared to the intracellular fluid, Ca^{++} ions enter the cell down their concentration gradient. The Ca^{++} ions then induce the release of the neurotransmitter from the synaptic vesicles into the synaptic cleft by causing the vesicles to fuse with the presynaptic membrane and, thereby, facilitating the process of exocytosis. The region of the synaptic

knob where there is an abundance of voltage-gated Ca^{++} channels and where the vesicles tend to fuse with the membrane is referred to as the *active zone.* The neurotransmitter molecules diffuse across the cleft and bind to specific receptors on the membrane of the postsynaptic neuron. This binding of the neurotransmitter may alter the permeability of the postsynaptic neuron to one or more ions or it may elicit biochemical changes within the cell by way of G proteins. The activation of G proteins also opens or closes ion channels. As always, a change in ion permeability results in a change in the membrane potential of the cell. This change in membrane potential at the synapse, which represents the information as received by the postsynaptic neuron, is in the form of a *graded potential* only. At any given synapse, the change in membrane potential is not great enough to reach threshold and generate an action potential. Instead, many graded potentials generated at one or more synapses are conducted over the cell membrane toward the axon hillock. If the depolarization caused by multiple graded potentials added together is sufficient for the axon hillock to reach threshold, then an action potential is generated here.

There are two important characteristics of synaptic function. First, the transmission of the electrical impulse is *unidirectional*, from the presynaptic neuron to the postsynaptic neuron. This is because only the presynaptic neuron releases neurotransmitter and only the postsynaptic neuron possesses the receptor that will elicit the expected response. Second, the strength of the response is variable and depends on the amount of neurotransmitter released into the synaptic cleft. As mentioned previously, neurotransmitter is preformed and stored in vesicles in the synaptic knob. Each vesicle contains a fixed amount of neurotransmitter, which is referred to as a *quantum*. Therefore, the amount of neurotransmitter released depends on the number of vesicles that fuse with the membrane in response to the stimulating action potentials. As more neurotransmitter is released to bind with its receptors, the greater is the response in the postsynaptic neuron.

There are two types of synapses: *excitatory synapses* and *inhibitory synapses*. At an *excitatory synapse*, binding of the neurotransmitter to its receptor causes an increase in the permeability of the membrane to Na^+ ions and K^+ ions through *chemical messenger-gated channels* that are closely associated with the receptor. As a result, Na^+ ions enter the cell down both their concentration and electrical gradients, and K^+ ions leave the cell down their concentration gradient only. Because there are two forces causing the inward diffusion of sodium and only one force causing the outward diffusion of potassium, the influx of Na^+ ions is significantly greater than the efflux of K^+ ions. This greater movement of positive (+) charges into the cell results in a small depolarization of the neuron and is referred to as an *excitatory postsynaptic potential* (EPSP). An EPSP is a graded potential only. A single action potential occurring at a single excitatory synapse opens too few Na^+ channels to depolarize the membrane all of the way to threshold; however, it does bring the membrane potential closer toward it. This increases the likelihood that

PHARMACY APPLICATION: ALZHEIMER'S DISEASE

Alzheimer's disease (AD) is a common, chronic, progressive neurologic disorder. Early-onset AD occurs in patients younger than 65 years of age. However, late-onset AD is more prevalent, accounting for 70% of the cases. This disorder affects approximately 4 million Americans, yet the cause remains unknown. It is characterized by the loss of neurons in the cerebral cortex as well as the presence of neurofibrillary tangles and plaques. Neurochemically, AD is associated with a decrease in the level of choline acetyltransferase in the brain. This enzyme is needed for the synthesis of the neurotransmitter, acetylcholine, which is associated with memory. In patients with AD, the deficiency in choline acetyltransferase (and, therefore, acetylcholine) is directly correlated to the severity of the dementia.

Currently there is no cure for AD. However, medications are available to slow the progression of the disease. These drugs have been designed to enhance the level of available acetylcholine in the brain. Long-lasting acetylcholinesterase inhibitors are drugs that inhibit the enzyme that degrades acetylcholine. As a result, more acetylcholine is available in the synapse to carry out its effects. These medications include Cognex® (tacrine), Aricept® (donepezil), Excelon® (rivastigmine), and Razadyne® (galantamine).

subsequent stimuli will continue the depolarization to threshold and that an action potential will be generated by the postsynaptic neuron.

At an *inhibitory synapse*, binding of the neurotransmitter to its receptor causes an increase in the permeability of the membrane to either K^+ ions or, more commonly, Cl^- ions through chemical messenger-gated channels (sodium channels are not affected). As a result, K^+ ions may leave the cell down their concentration gradient carrying positive (+) charges outward, or Cl^- ions may enter the cell down their concentration gradient carrying negative (–) charges inward. In either case, the neuron becomes more negative inside relative to the outside, and the membrane is now hyperpolarized. This small hyperpolarization is referred to as an *inhibitory postsynaptic potential* (IPSP). The movement of the membrane potential farther away from threshold decreases the likelihood that an action potential will be generated by the postsynaptic neuron.

Almost invariably, a neuron is genetically programmed to synthesize and release only a single type of neurotransmitter. Therefore, a given synapse is either always excitatory or always inhibitory. Once a neurotransmitter has bound to its receptor on the postsynaptic neuron and has caused its effect, it is important to inactivate it or remove it from the synapse in order to

prevent it from continuing its activity indefinitely. Several mechanisms to carry this out have been identified:

- Passive diffusion of the neurotransmitter away from the synaptic cleft.
- Destruction of the neurotransmitter by enzymes located in the synaptic cleft or in the plasma membranes of the presynaptic or postsynaptic neurons.
- Active reuptake of the neurotransmitter into the synaptic knob of the presynaptic neuron for reuse or enzymatic destruction.

5.3 Summation

As previously mentioned, a single action potential at a single synapse results in a graded potential only — either an EPSP or an IPSP. Therefore, the generation of an action potential in the postsynaptic neuron requires the addition or *summation* of a sufficient number of excitatory inputs to depolarize this neuron to threshold. There are two types of summation that may occur: *temporal summation* or *spatial summation*.

Temporal summation occurs when multiple EPSPs (or IPSPs) produced by a *single* presynaptic neuron in close sequence exert their effect on the membrane potential of the postsynaptic neuron. For example, an action potential in the presynaptic neuron produces an EPSP and partial depolarization of the postsynaptic neuron (see Figure 5.2). While the postsynaptic neuron is still depolarized, a second action potential in the presynaptic neuron produces another EPSP in the postsynaptic neuron that adds to the first EPSP and further depolarizes this neuron. As more and more EPSPs add together, the membrane depolarizes closer and closer toward threshold until an action potential is generated. Although temporal summation is illustrated in Figure 5.2 with the summation of relatively few EPSPs, in actuality, the addition of up to 50 EPSPs may be necessary to reach threshold. Because a presynaptic neuron may generate up to 500 action potentials per second, temporal summation occurs quite readily. The strength of the signal to the postsynaptic neuron is, therefore, influenced by the *frequency of nerve impulses* generated by the presynaptic neuron.

Spatial summation occurs when multiple EPSPs (or IPSPs), produced by *many* presynaptic neurons, exert their effects on the membrane potential of the postsynaptic neuron simultaneously. For example, Figure 5.3 depicts a single postsynaptic neuron that is innervated by three presynaptic neurons. Inputs from presynaptic neurons A and B are excitatory, and the input from presynaptic neuron C is inhibitory. Once again, single action potentials in either neuron A or B produce individual EPSPs that are insufficient to depolarize the postsynaptic neuron to threshold. However, if EPSPs from neurons A and B are produced at the same time, the depolarizations add together and the membrane potential of the postsynaptic neuron reaches threshold, resulting in the generation of an action potential. Inputs from neurons A

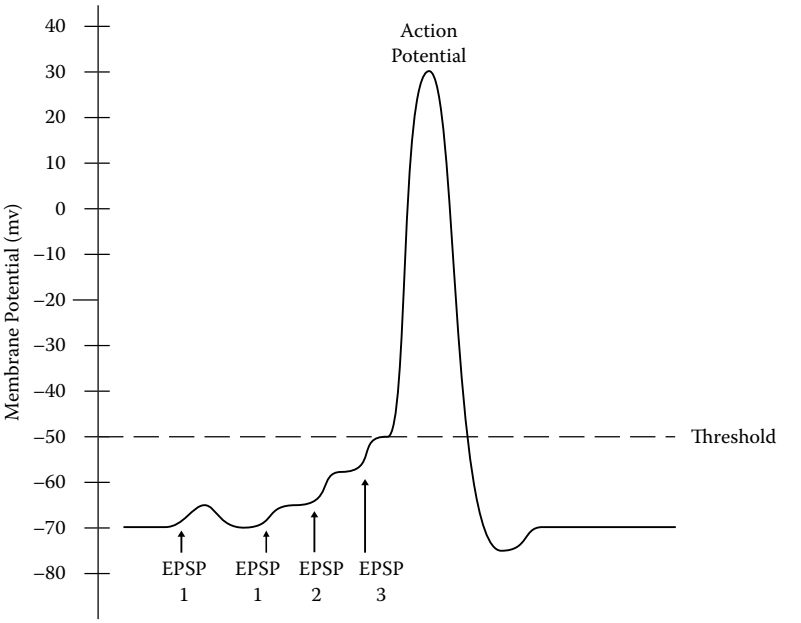


Figure 5.2 Temporal summation. Multiple EPSPs produced by a single presynaptic neuron in close sequence may add together to depolarize the postsynaptic neuron to threshold and generate an action potential.

(excitatory) and C (inhibitory) occurring simultaneously may, in effect, cancel each other out, resulting in little or no change in the membrane potential of the postsynaptic neuron. As with temporal summation, this example has been simplified to clearly illustrate the concept. In actuality, a large number of excitatory inputs from different presynaptic neurons are necessary to depolarize the postsynaptic neuron to threshold. Because a typical neuronal cell body receives thousands of presynaptic inputs, spatial summation also occurs quite readily. The number of presynaptic neurons that are active simultaneously, therefore, influences the strength of the signal to the postsynaptic neuron. Under normal physiological conditions, temporal summation and spatial summation may occur concurrently.

5.4 Interconnections between neurons

The interconnections or communication among neurons in humans is very extensive. Imagine the complexity of the electrical activity that may occur among 100 billion neurons in the human brain, where each of these neurons provides input to and receives input from hundreds of other neurons. It is the diversity of these interconnections that accounts for the uniqueness of many abstract neurological phenomena in individuals such as intellect,

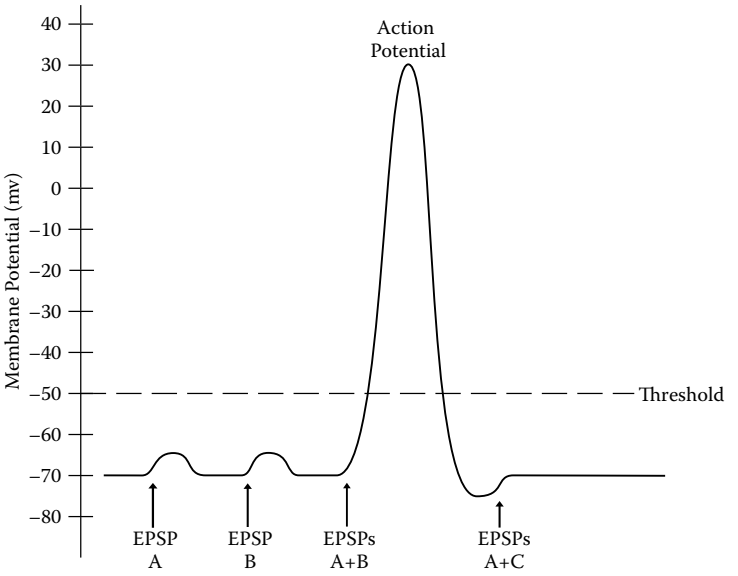


Figure 5.3 Spatial summation. Multiple EPSPs or IPSPs produced by many presynaptic neurons simultaneously may add together to alter the membrane potential of the postsynaptic neuron. Sufficient excitatory input (A and B) will depolarize the membrane to threshold and generate an action potential. The simultaneous arrival of excitatory and inhibitory inputs (A and C) may cause them to cancel each other out so that the membrane potential does not change.

personality, and memory. There are two types of interconnections: *convergence* and *divergence*.

Convergence occurs when the axon terminals of many presynaptic neurons all synapse with a single postsynaptic neuron. As discussed previously, spatial summation of nerve impulses relies on the presence of convergence. *Divergence* occurs when the axon of a single presynaptic neuron branches and synapses with multiple postsynaptic neurons. In this way, activity in a single nerve fiber can affect several regions of the nervous system, each with a different function, at the same time.

5.5 *Factors affecting synaptic transmission*

There are several factors that influence the synaptic transmission of electrical impulses:

- pH of the interstitial fluid
- Hypoxia
- Drugs, toxins, and diseases

Neurons are very sensitive to changes in the *pH of the interstitial fluid* surrounding them. Usually, the pH of arterial blood is 7.4. Under conditions of *alkalosis* where the pH increases, the excitability of neurons also increases, rendering them more likely to generate action potentials. This inappropriate stimulation of the nervous system may lead to seizures, particularly in epileptics who are predisposed to having them. Under conditions of *acidosis* where the pH decreases, the excitability of neurons is depressed, rendering them less likely to generate action potentials. This lack of stimulation of the nervous system may lead to a comatose state. Severe diabetic acidosis or acidosis associated with end-stage renal failure will often lead to coma.

Neuronal function depends on a constant supply of oxygen. *Hypoxia*, a decrease in oxygen availability, depresses neuronal activity. Interruption of blood flow to the brain for only a few seconds leads to unconsciousness. A prolonged lack of blood flow, which is characteristic of stroke, leads to permanent brain damage in the affected area.

Many *drugs, toxins,* and *diseases* also exert their clinical effects by altering some phase of synaptic activity. In fact, drugs are capable of altering literally every level of neuronal and synaptic function. There are four mechanisms by which these effects may occur:

1. Altered release of a neurotransmitter.
2. Altered interaction of a neurotransmitter with its receptor.
3. Altered removal of a neurotransmitter from the synaptic cleft.
4. Replacement of a deficient neurotransmitter.

5.5.1 Altered release

Tetanus is an infectious disease caused by the bacterium *Clostridium tetani*. This bacterium produces a neurotoxin active on inhibitory synapses in the spinal cord. Motor neurons, the neurons that supply skeletal muscle and cause contraction, have cell bodies that lie in the spinal cord. Under normal circumstances, these motor neurons receive both excitatory and inhibitory inputs from various sources. It is the balance of these inputs that results in the appropriate degree of muscle tone or muscle contraction. Tetanus toxin prevents the release of gamma amino butyric acid (GABA), an important neurotransmitter active at the inhibitory synapses. Elimination of the inhibitory inputs results in unchecked or unmodulated excitatory input to the motor neurons. The resulting uncontrolled muscle spasms initially occur in the muscles of the jaw, giving rise to the expression *lockjaw*. The muscle spasms eventually affect the respiratory muscles, which prevents inspiration and leads to death due to asphyxiation.

5.5.2 Altered interaction of a neurotransmitter with its receptor

There are several ways in which the interaction of a neurotransmitter with its receptor may be altered pharmacologically. One such mechanism involves the administration of *antagonists,* drugs that bind to a given receptor and prevent the action of the neurotransmitter but, by classical definition, initiate no other effect. An interesting clinical example of this form of therapy involves schizophrenia, a severe mental disorder characterized by delusions, hallucinations, social withdrawal, and disorganized speech and behavior. Although the precise cause of schizophrenia is unknown, its pathophysiology appears to involve neuronal pathways that release excessive amounts of the neurotransmitter dopamine. Antipsychotic drugs, such as Thorazine® (chlorpromazine) and Haldol® (haloperidol), minimize the symptoms of schizophrenia by blocking dopamine receptors, thereby preventing the excess dopamine from exerting its effects.

An *agonist* is a drug that binds to a given receptor and stimulates it. In other words, agonists mimic the effect of endogenous neurotransmitters. Albuterol, the active ingredient in medications such as Ventolin®, is a β_2-adrenergic receptor agonist. Stimulation of these receptors in the lungs causes the airways to dilate. Therefore, albuterol is effective in reversing the bronchospasm and dyspnea (difficulty in breathing) associated with asthma.

Another mechanism by which neurotransmitter/receptor interaction may be altered involves the administration of drugs that *facilitate* the binding of the endogenously produced neurotransmitter to its receptor. Once again, the neurotransmitter used as an example is GABA, the most prevalent inhibitory neurotransmitter in the nervous system. Not only does it contribute to the regulation of skeletal muscle tone by inhibiting the activity of motor neurons, it is also involved in the regulation of mood and emotions by acting as a CNS depressant. The benzodiazepines, which are antianxiety drugs and include Valium® (diazepam) and Ativan® (lorazepam), act by binding to a specific site on the GABA receptor. This binding causes a conformational change in the receptor protein that enhances the binding of GABA. As more GABA binds to the receptors, its effectiveness in the CNS is increased, and anxiety is decreased.

5.5.3 Altered removal of a neurotransmitter from the synaptic cleft

The third mechanism by which drugs may alter synaptic activity involves changes in neurotransmitter reuptake or degradation. Very well-known examples of drugs in this category are Prozac® (fluoxetine) and the more recently prescribed Cymbalta® (duloxetine) and Wellbutrin® (bupropion), which are used to treat depression. Although the complete etiology is unknown, it is widely accepted that depression involves a deficiency of monoamine neurotransmitters (e.g., norepinephrine and serotonin) in the

CNS. Prozac, a selective serotonin reuptake inhibitor, prevents the removal of serotonin from the synaptic cleft. Cymbalta prevents the reuptake of both serotonin and norepinephrine, and Wellbutrin prevents the reuptake of dopamine and norepinephrine. As a result, the concentration and, therefore, the activity of these neurotransmitters in the brain are enhanced, and the symptoms of depression are relieved.

5.5.4 Replacement of a deficient neurotransmitter

Finally, synaptic activity may be altered by the replacement of a deficient neurotransmitter. This form of drug therapy is effective in the treatment of Parkinson's disease. The pathophysiology of Parkinson's involves the progressive destruction of dopaminergic (dopamine-releasing) neurons resulting in a deficiency of dopamine in certain areas in the brain. In addition to neuronal pathways involved in the regulation of mood and emotion, dopamine is released by neurons that inhibit skeletal muscle contraction. Because motor neurons normally receive both excitatory and inhibitory inputs, the inhibition provided by the dopaminergic pathways results in smooth, precise muscle contractions. In the patient with Parkinson's disease, this loss of inhibition leads to increased muscle tone, or muscle rigidity, and resting tremors. These symptoms are alleviated by administering levodopa (L-dopa), a precursor for dopamine. L-dopa is taken up by the axon terminals of dopaminergic neurons and used to form dopamine. Interestingly, in some patients, a side effect of this dopamine replacement therapy is the development of symptoms characteristic of schizophrenia. Recall that this mental disorder is caused by overactive dopaminergic neurons. On the other hand, drugs used to treat schizophrenia, dopamine receptor antagonists, may elicit the symptoms of Parkinson's disease.

Medical terminology

Acidosis (ăs″ĭ-dō′sĭs): An increase in the concentration of hydrogen ions in the arterial blood such that the pH is less than 7.4.

Agonist (ăg′ŏn-ĭst): A chemical substance or drug that binds to a receptor and elicits the same effects as an endogenously produced substance or neurotransmitter.

Alkalosis (ălk″ă-lō′sĭs): A decrease in the concentration of hydrogen ions in the arterial blood such that the pH is greater than 7.4.

Antagonist (ăn-tăg′ŏn-ĭst): A chemical substance or drug that binds to a receptor and blocks or prevents the effects of an endogenously produced substance or neurotransmitter.

Convergence (cŏn-vĕr′gĕns): Condition where more than one presynaptic neuron synapses with and influences one postsynaptic neuron.

Divergence (dī-vĕr′gĕns): Condition where one presynaptic neuron branches, synapsing with and influencing more than one postsynaptic neuron.

Facilitation (fă-sĭ″ĭ-tā′shŭn): Condition where a postsynaptic neuron has received a subthreshold stimulus and is, therefore, partially depolarized and more likely to generate an action potential upon subsequent stimulation.

Hypoxia (hī-pŏks′ē-ă): An abnormal decrease in oxygen in the tissues.

Quantum (kwŏn′tŭm): A fixed or definite amount.

Synapse (sĭn′aps): Point of contact between two neurons, typically including the axon terminal of one neuron and the cell body or dendrites of the second.

Synaptic knob (sĭn-ăp′tĭc nŏb): Region of the presynaptic axon terminal that comes into close apposition with the postsynaptic neuron.

Unidirectional (ū″nē-dī-rĕk′shŭn-ŏl): In one direction only.

Vesicle (vĕs′ĭ-kl): Membranous sac that stores preformed neurotransmitter in the synaptic knob.

Bibliography

Baldessarini, R., Therapy of Depression and Anxiety Disorders, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 17.

Baldessarini, R., and Tarazi, F. I., Pharmacotherapy of Psychosis and Mania, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 18.

Bear, M. F., Connors, B. W., and Paradiso, M. A., Chemical Control of the Brain and Behavior, in *Neuroscience, Exploring the Brain,* 3rd ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2007, chap. 15.

Bloom, F., Neurotransmission and the Central Nervous System, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 12.

Boss, B. J., Concepts of Neurological Dysfunction, in *Pathophysiology, The Biologic Basis for Disease in Adults & Children,* Mosby, St. Louis, MO, 2002, chap. 15.

Buxton, I. L. O., Pharmacokinetics and Pharmacodynamics: The Dynamics of Drug Absorption, Distribution, Action and Elimination, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 1.

Finley, P. R., Selective Serotonin Re-uptake Inhibitors: Pharmacologic Profiles and Potential Therapeutic Distinctions. *Annals of Pharmacotherapy*, 28(12): 1359–1369, 1994.

Garoutte, B., *Neuromuscular Physiology*, Mill Valley Medical, Millbrae, CA, 1996.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hanson, M., *Pathophysiology, Foundations of Disease and Clinical Intervention*, W.B. Saunders, Philadelphia, PA, 1998.

Kane, J. M., Schizophrenia. *New England Journal of Medicine*, 334(1): 34–41, 1996.

Lombard, J. H., and Rusch, N. J., Cells, Nerves and Muscles, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 1.

Pasch, S. K., Disorders of Thought, Mood, and Memory, in *Pathophysiology, Concept of Altered Health States*, Porth, C.M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 53.

Rockhold, R. W., The Chemical Basis for Neuronal Communication, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 4.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Standaert, D. G., and Young, A. B., Treatment of Central Nervous System Degenerative Disorders, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 20.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E., Raff, H., and Strang, K., *Vander's Human Physiology, The Mechanisms of Body Function,* 10th ed., McGraw-Hill, New York, 2006.

chapter six

The nervous system

Study objectives

- Describe the organization of the nervous system including the central nervous system (CNS) and the peripheral nervous system (PNS)
- Distinguish between the three types of neurons: afferent neurons, efferent neurons, and interneurons
- Describe the functions of the neuroglial cells
- List the three major levels of CNS function and describe their activities
- Distinguish between the three types of tracts in the CNS: projection tracts, association tracts, and commissural tracts
- Describe the activity of each of the functional areas of the cerebral cortex
- Explain how language is processed in the cerebral cortex
- Describe the lateral specialization that occurs in the cerebral cortex
- Describe the function of the basal ganglia, the thalamus, the hypothalamus, and the brainstem
- Describe the function of the limbic system
- Distinguish between the three regions of the cerebellum and their functions
- Compare and contrast the exchange of materials between the blood and peripheral tissues with that of the blood and the brain
- Explain the functions of the blood–brain barrier
- Explain the functions of cerebrospinal fluid

6.1 Introduction

The nervous system is one of the two regulatory systems in the human body that influences the activity of all of the other organ systems. It consists of literally billions of neurons that are interconnected in a highly organized manner to form circuits. It is the number of neurons and the manner in which they are interconnected in a given circuit that distinguishes one region of the brain from another, and the brain of one individual from that of another. In addition, *plasticity*, the ability to alter circuit connections and function in response to sensory input and experiences, adds further complexity and distinctiveness to our neurological responses and behavior.

The nervous system is divided into two anatomically distinct regions: the *central nervous system* (CNS) and the *peripheral nervous system* (PNS). The *CNS*

consists of the brain and the spinal cord. The *PNS* consists of the 12 pairs of cranial nerves that arise from the brainstem and the 31 pairs of spinal nerves that arise from the spinal cord. These peripheral nerves carry information between the CNS and the tissues of the body. The PNS consists of two divisions: the *afferent division* and the *efferent division*.

The *afferent division* carries sensory information toward the CNS, and the *efferent division* carries motor information away from the CNS toward the effector tissues (muscles and glands). The efferent division is further divided into two components: the somatic nervous system that consists of the motor neurons that innervate skeletal muscle, and the autonomic nervous system that innervates cardiac muscle, smooth muscle, and glands.

6.2 Classes of neurons

There are three functional classes of neurons in the human nervous system:

1. Afferent neurons
2. Efferent neurons
3. Interneurons

Afferent neurons lie predominantly in the PNS (see Figure 6.1). Each has a sensory receptor that is activated by a particular type of stimulus, a cell body located adjacent to the spinal cord, and an axon. The *peripheral axon* extends from the receptor to the cell body, and the *central axon* continues from the cell body into the spinal cord. Afferent neurons transmit information from receptors in the periphery of the body to the CNS.

Efferent neurons also lie predominantly in the PNS. In this case, the cell bodies are found in the CNS in either the spinal cord or the brainstem, and the axons extend out into the periphery of the body where they innervate the effector tissues. By way of convergence, the centrally located cell bodies may receive inputs from several different regions of the brain and the spinal cord that will influence their activity. Efferent neurons transmit information from the CNS to muscles and glands.

The third class of neurons includes the *interneurons*, which lie entirely within the CNS. As the human brain and spinal cord contain well over 100 billion neurons, interneurons account for approximately 99% of all the neurons in the body taken together. Interneurons lie between afferent and efferent neurons and are responsible for integrating sensory input and coordinating a motor response. In the simplest condition, interneurons process responses at the level of the spinal cord in the form of *reflexes* that are automatic, stereotyped responses to given stimuli. For example, stimulation of pain receptors generates action potentials in their associated afferent neurons. These impulses are transmitted to the spinal cord where the afferent neurons stimulate interneurons. These interneurons then stimulate

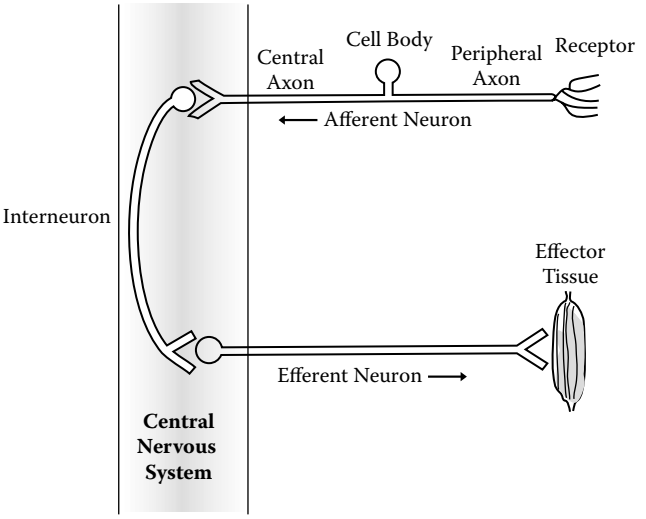


Figure 6.1 Types of neurons. Afferent neurons, which transmit impulses toward the CNS and efferent neurons, which transmit impulses away from the CNS, lie predominantly in the PNS. Interneurons, which process sensory input and coordinate motor responses, lie entirely within the CNS.

efferent neurons that cause skeletal muscle contraction in the affected area to remove the body part from the painful stimulus. This *withdrawal reflex* involves comparatively few interneurons and does not require any input from higher nervous centers in the brain. On the other hand, a response to some other stimulus may involve more sophisticated neurological phenomena, such as memory, motivation, judgment, and intellect. This type of response is not automatic, is clearly far more complex, and may require the activity of millions of interneurons in many regions of the brain prior to the stimulation of motor neurons to carry out the desired effect.

6.3 *Major levels of central nervous system (CNS) function*

There are three major levels of CNS function:

1. Spinal cord
2. Brainstem
3. Cerebrum and the cerebral cortex

The *spinal cord* is the most anatomically inferior portion of the CNS, and its functions are at the lowest level of sophistication (see Table 6.1). As mentioned previously, the spinal cord receives sensory input from the periphery

Table 6.1 Major Levels of CNS Function

Spinal Cord	Processes reflexes, transmits nerve impulses to and from brain
Brainstem	Receives sensory input and initiates motor output, controls life-sustaining processes (e.g., respiration, circulation, digestion)
Cerebrum and Cerebral Cortex	Processes, integrates, and analyzes information; involved with the highest levels of cognition, voluntary initiation of movement, sensory perception, and language

of the body and contains the cell bodies of motor neurons responsible for both voluntary and involuntary movements. Once again, the involuntary and neurologically simple reflexes are processed entirely at the level of the spinal cord. Voluntary, deliberate movements are initiated and controlled by thought processes in the cerebrum. The second important function of the spinal cord is to *transmit nerve impulses* to and from the brain. *Ascending pathways* carry sensory input to higher levels of the CNS, and *descending pathways* carry impulses from the brain to motor neurons in the spinal cord.

The *brainstem*, which consists of the medulla, pons, and midbrain, is, in evolutionary terms, the oldest and smallest region of the brain. Continuous with the spinal cord, the brainstem receives sensory input and initiates motor output by way of cranial nerves III through XII, which are functionally analogous to the 31 pairs of spinal nerves. Whereas the spinal cord processes sensory and motor activities in the trunk of the body and the limbs, the brainstem processes these activities primarily in the head, neck, and face. The brainstem also controls many basic life-sustaining processes including respiration, circulation, and digestion. Even with the loss of higher cognitive function, this lower level of the brain can sustain these bodily functions essential for survival.

The *cerebrum* and the *cerebral cortex*, which account for 80% of the total brain weight in humans, constitute the highest functional level of the CNS. As the species becomes more cognitively sophisticated, the cerebral cortex becomes larger and more highly folded. These convolutions or folds serve to increase the surface area of the cerebral cortex, thus allowing for a greater number of neurons. For example, the total surface area of the cerebral cortex for the rat, cat, and human are 6 cm^2, 83 cm^2, and 2500 cm^2, respectively. It is not unexpected that the cerebrum is the most highly developed in the human. Responsible for the highest levels of processing, integration, and analysis of information, the cerebral cortex plays an important role in the most elaborate neurological functions, including intellect, thought, personality, voluntary initiation of movement, final sensory perception, and language.

Table 6.2 Adult Brain Structures

Forebrain • Cerebrum	→ Cerebral Cortex
• Basal Ganglia • Thalamus • Hypothalamus	→ Subcortical Structures (embedded within cerebrum)
Midbrain Hindbrain • Pons • Medulla	→ Brainstem
• Cerebellum	

6.4 The brain

The brain is the integrative portion of the nervous system that serves to receive, process, and store sensory information and then plan and orchestrate the appropriate motor response. It is divided into several anatomically and functionally distinct regions (see Table 6.2). The forebrain consists of the cerebrum, the basal ganglia, the thalamus, and the hypothalamus. The midbrain along with the pons and the medulla of the hindbrain compose the functional region referred to as the brainstem. The cerebellum is also considered a component of the hindbrain but is functionally distinct from the brainstem. A summary of the functions of the major components of the brain is found in Table 6.3.

6.4.1 Cerebrum

The *cerebrum* is composed of two hemispheres, the left and the right, that are anatomically connected to ensure communication between them. Two types of tissue compose each hemisphere (see Figure 6.2): *gray matter* and *white matter*.

The *gray matter*, which contains the cell bodies of neurons, is on the outer surface of the cerebrum and forms the *cerebral cortex*. The *white matter*, which is composed of the myelinated axons of neurons, is found underlying the cortex in the core of the cerebrum. These axons are bundled together according to function and organized into units referred to as *tracts*. There are three types of tracts in the cerebrum:

1. Projection tracts
2. Association tracts
3. Commissural tracts

Table 6.3 Summary of the Functions of the Major Components of the Brain

Forebrain

Cerebral Cortex

- Sensory perception
- Voluntary movement of skeletal muscle
- Language
- Intellect, personality, judgment, behavior, memory

Basal Ganglia

- Inhibition of skeletal muscle tone
- Coordination of slow, sustained movements

Thalamus

- Relay station for sensory input
- Role in awareness

Hypothalamus

- Regulation of the autonomic nervous system
- Regulation of endocrine system function
- Thermoregulation
- Regulation of plasma volume, plasma osmolarity, thirst, hunger, urine output
- Influence on behavior and emotions

Brainstem

Medulla

- Cardiovascular, respiratory, digestive control centers

Pons

- Relay station for cerebrum and cerebellum
- Assists medulla in control of breathing

Midbrain

- Contributes to control of eye movements
- Relay station for auditory and visual reflexes

Reticular Formation

- Cortical awareness, ability to direct attention, sleep
- Coordination of orofacial motor activities
- Coordination of eating and breathing
- Response to pain

Cerebellum

- Maintenance of balance
- Control of eye movements

Table 6.3 (continued) Summary of the Functions of the Major Components of the Brain

Cerebellum (continued)
• Regulation of skeletal muscle tone • Coordination of skilled voluntary movements • Contributes to planning, programming, and initiation of voluntary skeletal muscle activity • Procedural memory and motor learning

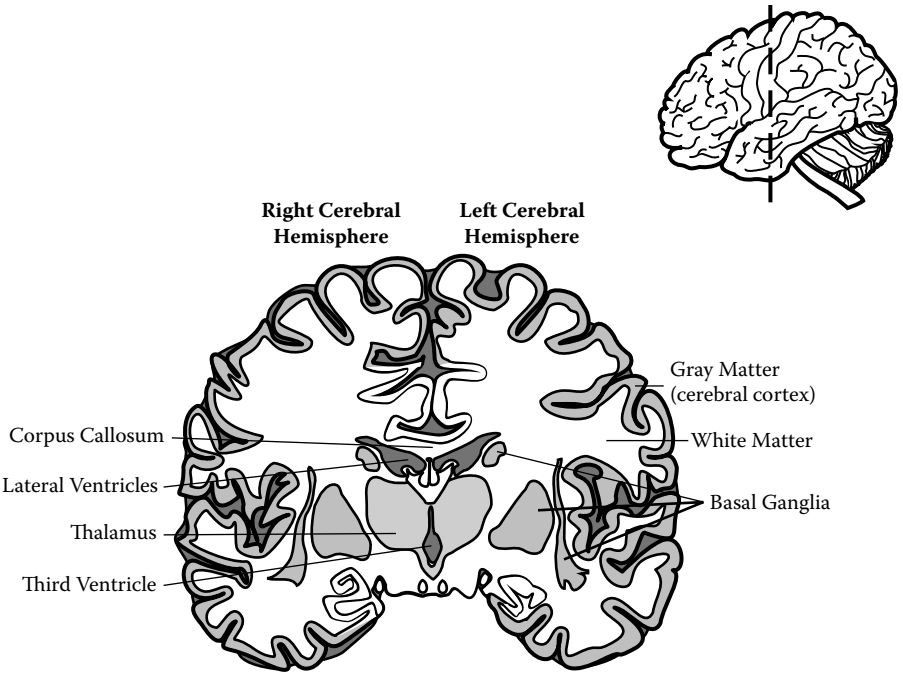


Figure 6.2 Frontal section of the brain. The cerebrum is composed of two types of tissue: the internal white matter and the external gray matter which forms the cerebral cortex. Embedded within the cerebral hemispheres are other masses of gray matter, the basal ganglia, and the thalamus. The ventricles are filled with cerebrospinal fluid (CSF).

Projection tracts may be *descending* and carry motor nerve impulses from the cerebral cortex to lower regions of the brain or spinal cord or they may be *ascending* and carry sensory impulses from lower regions of the brain or spinal cord to the cortex. *Association tracts* transmit nerve impulses from one functional region of the cerebral cortex to another within the same hemisphere. *Commissural tracts* transmit impulses from one hemisphere to the other. The primary example of this type of tract is the corpus callosum, the thick band

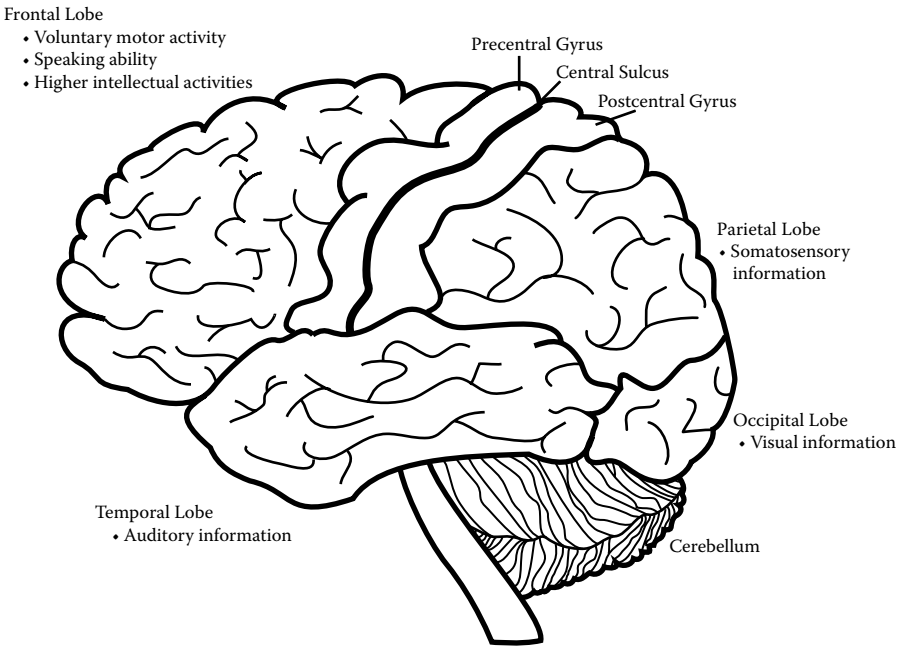


Figure 6.3 Lateral view of the four lobes of the cerebral cortex.

of tissue connecting the left and the right hemispheres consisting of more than 100 million neurons. The communication provided by each of these types of tracts facilitates the integration, processing, and storage of information among the various regions of the brain.

The cerebral cortex is not a smooth surface but, instead, is highly folded and has a furrowed appearance (see Figure 6.3). A convolution formed by these folds is referred to as a *gyrus* (plural: *gyri).* Each gyrus is separated from another by a *sulcus* (plural: *sulci),* which is a shallow groove, or a *fissure*, which is a deeper groove. The functional importance of gyri, sulci, and fissures is that they significantly increase the surface area of the cerebral cortex, providing space for a greater number of neurons.

Both hemispheres of the cerebrum consist of four lobes:

1. Frontal lobes
2. Parietal lobes
3. Occipital lobes
4. Temporal lobes

Named for the bones of the cranium under which they lie, the lobes are conspicuously defined by prominent sulci of the cortex, which have a relatively constant position in human brains. Each lobe is specialized for different activities (see Figure 6.3 and Table 6.4). The *frontal lobes*, which are

Table 6.4 Functions of the Lobes of the Cerebral Cortex

Frontal Lobe
• Voluntary movement of skeletal muscle
• Speaking ability
• Higher intellectual abilities, personality, judgment, behavior
Parietal Lobe
• Processing and integration of somatosensory information
• Understanding spoken and written language
• Formulating coherent patterns of speech (or writing) to express thoughts and emotions
Occipital Lobe
• Processing and integration of visual information
Temporal Lobe
• Processing and integration of auditory information

located in the anterior portions of the hemispheres, are responsible for voluntary motor activity, speaking ability, and higher intellectual activities. The *parietal lobes*, which are posterior to the frontal lobes, process and integrate sensory information. The *occipital lobes* in the posterior-most aspects of the cerebrum process visual information, and the *temporal lobes*, located laterally, process auditory information.

6.4.2 Functional regions of the cerebral cortex

The cerebral cortex is organized into several functionally discrete areas (see Figure 6.4). However, it is important to remember that no single area functions in isolation. The activity in each area depends on neurons in other areas for incoming and outgoing messages.

The *somatosensory cortex* is located in the postcentral gyrus which is the anterior-most region of the parietal lobes (see Figure 6.3). This region contains the terminations of ascending pathways that transmit nerve impulses concerning temperature, touch, pressure, pain, and proprioception. The latter is the awareness of posture, movement, changes in equilibrium, and the position of one's body parts, particularly in reference to surrounding objects. As such, the somatosensory cortex is the site for initial cerebral processing of these types of inputs.

Each section of this region of cortex receives sensory input from a specific area of the body in a highly organized and sequential manner. It is organized in a "foot-to-tongue" pattern along the medial-to-lateral axis (top to bottom of the gyrus). Interestingly, the size of the region of the cortex devoted to different areas of the body is quite disproportionate. For example, the trunk of the body and the legs are not densely innervated with sensory neurons.

Cortical Area	Function
Sensory Input	Relayed from afferent neuronal receptors.
↓	
Primary Sensory Areas	Initial cortical processing of sensory input.
↓	
Unimodal Association Areas	Further processing of information from a single sensory modality.
↓	
Multimodal Sensory Association Areas	Highest level of processing, integration, and interpretation of diverse sensory input for planning purposeful action.
↓	
Multimodal Motor Association Areas	Neuronal programming of movements according to cortical and subcortical input.
↓	
Primary Motor Cortex	Transmission of impulses to somatic efferent motor neurons in spinal cord to initiate voluntary contraction of skeletal muscle.

Figure 6.4 Potential route of transmission of electrical impulses through association pathways of the cerebral cortex.

As a result, the axonal terminations of the pathways originating in these body parts are limited in number and take up only a small portion of the somatosensory cortex. Conversely, the face, the tongue, and the hands are very densely innervated with sensory neurons. Therefore, the terminations of pathways originating in these body parts are numerous and are represented in a much larger portion of the somatosensory cortex. The proportion of cortex devoted to a given body part is determined by the degree of sensory perception associated with that body part as well as the importance of the sensory input from that part of the body. The somatosensory cortex not only localizes the source of sensory input, but it also perceives the intensity of the stimulus.

These ascending sensory pathways cross from one side of the CNS to the other so that sensory input from the left side of the body is transmitted to the somatosensory cortex of the right cerebral hemisphere and vice versa. Therefore, damage to this region of cortex in a given hemisphere results in sensory deficits such as numbness and tingling in the opposite side of the body.

In addition to the somatosensory cortex, there are special senses areas in the cerebral cortex involved with the primary or initial processing of a specific type of stimulus. The *primary visual cortex* (sight) is located in the occipital lobes, the *primary auditory cortex* (hearing) and the *primary olfactory cortex* (smell) are located in the temporal lobes, and the *primary gustatory cortex* (taste) is located at the base of the somatosensory cortex in the parietal lobes.

Each of these primary areas is surrounded by a "higher-order" sensory area or a *unimodal association area* that further integrates information from

a single sensory modality and provides more complex aspects of the input. For example, the primary visual cortex is the first site of processing of visual information. Association tracts originating in this area then project to the surrounding unimodal association area for higher-level processing of this visual input. In fact, the visual unimodal association cortex occupies the remaining portion of the occipital lobes.

The *posterior parietal cortex* is located posterior to the somatosensory cortex and serves as its unimodal association area. In addition to the further processing of somatosensory input, information from the somatosensory cortex is integrated with visual inputs in this region, as association tracts from both of these functional areas terminate here. This activity is important for planning complex movements and for hand (proprioception)–eye (visual) coordination.

The unimodal association areas, in turn, project to *multimodal sensory association* areas that integrate information about more than one sensory modality. Activity in the multimodal sensory association areas helps to create a complete understanding of one's surroundings. In these areas, the highest level of cognitive brain function takes place. Not only do these areas process, integrate, and interpret several different forms of sensory information, they then link this data to the planning of movement and goal-directed action.

The *prefrontal multimodal association area* is located in the anterior-most region of the frontal lobe. It is involved with some of the most distinctly human intellectual traits, including memory and planning of motor activity, long-term planning and judgement, foresight, a sense of purpose, a sense of responsibility, a sense of social propriety, personality traits, and behavior. Consistent with this notion, lesions to this association area result in profound cognitive deficits, impaired motor activity, and changes in personality and social behavior. These patients do not respond to environmental stimuli in a way similar to healthy individuals. They have a lack of consistency of purpose, a lack of foresight, and a lack of ambition. In other words, they tend to achieve less in life; an outcome that suggests their ability to plan and organize everyday activities is impaired. Interestingly, however, their general intelligence, perception, and long-term memory are rather intact.

The *posterior multimodal association area* is located at the junction of the parietal, temporal, and occipital lobes. It pools and integrates somatic, auditory, and visual stimuli for complex perceptual processing. As such, this area is involved primarily with visuospatial localization, language, and attention. Lesions here interfere with awareness of one's body position and of the space in which it moves as well as the ability to integrate and make sense of the elements of a visual scene. In other words, these patients have normal visual acuity but cannot focus on an object of interest.

The *limbic multimodal association area* is partially located in each of the temporal, parietal, and frontal lobes. It is concerned with emotional expression and memory storage. Although these functions appear to be unrelated, it is important to remember that the emotional impact of an event is a major

determinant of whether the event is remembered. Damage here may cause profound changes in emotional expression with little change in perception or intelligence. Tumors in this region have been associated with emotional disturbances including fear, irritability, and depression. Once again, it is important to remember that while each of these multimodal association areas has their own characteristic functions, they are highly interconnected and work together toward an end result.

The multimodal sensory association areas then project to the *multimodal motor association areas* located in the frontal lobes, including the *premotor cortex* and the *supplementary motor cortex*. Neurons here are active during preparation for a movement. These regions receive input from the basal ganglia, the cerebellum, the somatosensory cortex, and the posterior multimodal association cortex (all of which provide information about the ongoing movement) as well as the prefrontal multimodal association area. As such, these areas are important in the programming of complex sequences of movements and in orienting the body and limbs toward a specific target. For example, the motor planning for opening a door includes turning the body toward the door and extending the arm and hand toward the doorknob. Lesions in these multimodal motor association areas interfere with the coordination and the performance of complex integrated movements.

Following the development of the motor program, neurons originating in the multimodal motor association areas transmit impulses by way of association tracts to the neurons of the primary motor cortex. The *primary motor cortex* is located in the precentral gyrus which is the posterior-most region of the frontal lobe adjacent to the multimodal motor association areas (see Figure 6.3). This area initiates voluntary contractions of specific skeletal muscles. Neurons whose cell bodies reside here transmit impulses by way of descending projection tracts to the spinal cord, where they synapse with the alpha motor neurons (which innervate skeletal muscles).

As with the somatosensory cortex, neurons here are highly organized with each section of the cortex innervating specific body parts in a sequential manner. Once again, the map is laid out in a "foot-to-tongue" pattern. Also like the somatosensory cortex, the size of the region of the primary motor cortex devoted to different parts of the body is quite disproportionate. Large portions of the primary motor cortex innervate the muscles of the hands, which perform complex movements, as well as the muscles responsible for speech and eating. On the other hand, little cortex is devoted to motor pathways terminating in the trunk of the body or the lower extremities, which are not capable of complex movements. Therefore, the distortions in cortical representation parallel the importance of a particular part of the body in terms of complexity of motor skills. Electrical activity in a specific region of the primary motor cortex will result in the contraction of its associated muscle, and recent studies have shown that many muscles, particularly those distal muscles of the upper extremities, are regulated from more than

one cortical location. Therefore, in addition to the area of the cortex devoted to the arm, other arm-specific neurons are located diffusely in other areas of the motor cortex.

A third similarity between the primary motor cortex and the somatosensory cortex is that the projection tracts cross from one side of the CNS to the other. Therefore, activity of motor neurons in the left cerebral hemisphere causes muscle contraction on the right side of the body and vice versa. As the commands for muscle contraction originate in the primary motor cortex, lesions in this region of cortex in a given hemisphere will result in paralysis in the opposite side of the body.

The exchange of information among individuals is largely limited to species with advanced nervous systems and is found predominantly in birds and mammals. In humans, communication takes place primarily through *language* or the use of spoken or written words to convey a message. The processing of language requires a large network of interacting brain areas, both cortical and subcortical. However, the two predominant cortical areas are Wernicke's area and Broca's area. In approximately 96% of people, these cortical areas for language skills are found only in the left hemisphere. Even languages such as American Sign Language, which rely on visuomotor abilities instead of auditory speech abilities, depend primarily on the left hemisphere.

Sensory input to the language areas comes from either the auditory cortex (hearing) or the visual cortex (reading). This input goes first to *Wernicke's area*, which is located in the left cerebral cortex near the junction between the parietal, temporal, and occipital lobes. This area is involved with language comprehension and is important for understanding both spoken and written messages. It is also responsible for formulating coherent patterns of speech. In other words, this area enables an individual to attach meaning to words and to choose the appropriate words to convey their thoughts. Impulses are then transmitted to *Broca's area* which is located in the left frontal lobe in close association with the motor areas of the cortex that control the muscles necessary for articulation. Broca's area is, therefore, responsible for the mechanical aspects of speaking.

A patient with a lesion in Wernicke's area is unable to understand any spoken or visual information. Furthermore, the patient's own speech, while fluent, is unintelligible because of frequent errors in the choice of words. This condition is known as *receptive aphasia*. On the other hand, a patient with a lesion in Broca's area is able to understand spoken and written language but is unable to express his or her response in a normal manner. Speech in this patient is nonfluent and requires great effort, as he or she cannot establish the proper motor command to articulate the desired words. This condition is known as *expressive aphasia*.

6.4.3 Basal ganglia

The *basal ganglia* consist of several nuclei or masses of gray matter embedded within the white matter of each cerebral hemisphere (see Figure 6.2). As with the cerebral cortex, this gray matter consists of functional aggregations of neuronal cell bodies. An important function of the basal ganglia involves their contribution to the control of voluntary movement. The axons of neurons originating in the primary motor cortex travel through descending projection tracts to the spinal cord where they stimulate motor neurons to cause skeletal muscle contraction. At the same time, by way of divergence, these neurons transmit impulses to the basal ganglia. These impulses form the primary source of input to these structures. In turn, the basal ganglia send impulses to the brainstem, which also transmits impulses to motor neurons in the spinal cord as well as the thalamus, which then transmits impulses back to the motor areas of the cerebral cortex.

The activity of the basal ganglia tends to be inhibitory. The thalamus positively reinforces motor activity in the cerebral cortex. Impulses from the basal ganglia modulate this effect. Through their inputs to the brainstem and, ultimately the motor neurons in the spinal cord, the basal ganglia inhibit muscle tone (recall that the degree of skeletal muscle contraction and tone is determined by the summation of excitatory and inhibitory inputs to the motor neurons). They also contribute to the coordination of slow, sustained contractions, especially those related to posture and body support. Motor disturbances associated with the basal ganglia include tremor and other involuntary movements; changes in posture and muscle tone, such as muscle rigidity; and slowness of movement without paralysis. Thus, disorders of the basal ganglia may result in either diminished movement (hypokinesias, such as Parkinson's disease) or excessive movement (hyperkinesias, such as Huntington's disease).

6.4.4 Thalamus

The *thalamus* is located between the cerebrum and the brainstem. Lying along the midline of the brain, it consists of two oval-shaped masses of gray matter, one in each cerebral hemisphere (see Figure 6.2). The thalamus is often described as a *relay station,* as ascending tracts transmitting upward from the spinal cord, as well as sensory tracts from the eyes and the ears, extending ultimately to the cerebral cortex pass through it. In fact, all sensory fiber tracts, except olfactory tracts, that transmit impulses to the cerebral cortex first synapse with neurons in the thalamus. Therefore, the thalamus may be considered the functional "gateway" to the cerebral cortex.

The thalamus acts as a *filter* for information to the cortex by either preventing or enhancing the passage of specific information depending upon its significance to the individual. In fact, more than 99% of all sensory information transmitted toward the brain is discarded, as it is considered irrelevant

and unimportant. This selection activity is accomplished largely at the level of the thalamus. As such, it plays an important role in general arousal and focused attention. For example, an individual may easily sleep through the noise of city traffic as this sensory input has been previously determined to be unimportant. However, this same individual may also be immediately aroused by the ever so quiet arrival of a teenager home late in the evening. This sensory is quite important to a parent and passes readily through the filter of the thalamus.

As mentioned previously in the discussion of the basal ganglia, the thalamus also plays a role in the regulation of skeletal muscle contraction by positively reinforcing voluntary motor activity initiated by the cerebral cortex.

6.4.5 Hypothalamus

As its name suggests, the *hypothalamus* lies beneath the thalamus and above the pituitary gland. Although it is quite small, accounting for only about 4 grams of the total 1400 grams of the adult human brain (<1% of the total brain mass), it plays a vital role in the maintenance of homeostasis in the body. It is composed of numerous cell groups and fiber pathways, each with a specific function.

The hypothalamus plays a particularly important role in regulating the autonomic nervous system, which innervates cardiac muscle, smooth muscle, and glands. Many of these effects involve ascending or descending pathways of the cerebral cortex passing through the hypothalamus. (Regulation of autonomic nervous system activity is discussed in detail in Chapter 9.) Endocrine activity is also regulated by the hypothalamus by way of its control over pituitary gland secretion. (Hypothalamic regulation of endocrine secretion is discussed in detail in Chapter 10.)

Recent studies have demonstrated that the hypothalamus serves to integrate autonomic nervous system responses and endocrine function with behavior, especially behavior associated with basic homeostatic requirements. The hypothalamus provides this integrative function by regulating the following:

- Blood pressure and electrolyte composition by regulating mechanisms involved with urine output, thirst, salt appetite, maintenance of plasma osmolarity, and vascular smooth muscle tone.
- Body temperature by regulating metabolic thermogenesis (e.g., shivering), cutaneous vasoconstriction or vasodilation, and behaviors that cause an individual to seek a warmer or cooler environment.
- Energy metabolism by regulating food intake, digestion, and metabolic rate.
- Reproduction by way of hormonal control of sexual activity, pregnancy, and lactation.

- Responses to stress by altering blood flow to skeletal muscles and other tissues as well as enhancing the secretion of hormones from the adrenal cortex (glucocorticoids) whose metabolic activities enable the body to physically cope with stress.

The hypothalamus regulates these physiological parameters by a three-step process involving negative feedback mechanisms (Chapter 1). First, the hypothalamus has access to and monitors sensory information from the entire body. Second, it compares this information to the various biological set points that have been established for optimal cellular function. Finally, if a deviation from set point for a given parameter is detected, the hypothalamus elicits a variety of autonomic, endocrine, and behavioral responses to return the parameter to its set point and reestablish homeostasis. For example, blood glucose levels are monitored by the hypothalamus. When blood glucose is low (<50 mg glucose/100 ml blood), it mediates the sensation of hunger to drive the individual to ingest food.

6.4.6 Brainstem

The functional region known as the *brainstem* consists of the midbrain, and the pons and medulla of the hindbrain. It is continuous with the spinal cord and serves as an important connection between the brain and the spinal cord as all sensory and motor pathways pass through it. The brainstem consists of numerous neuronal clusters or *centers*, each of which controls vital, life-supporting processes.

The *medulla,* which is immediately superior to and continuous with the spinal cord, contains control centers for subconscious, involuntary functions, such as cardiovascular activity, respiration, swallowing, and vomiting. The primary function of the *pons* is to serve as a relay for the transfer of information between the cerebrum and the cerebellum. Along with the medulla, it also contributes to the control of breathing. Continuing rostrally from the medulla and pons, the *midbrain* controls eye movement and relays signals for auditory and visual reflexes. It also provides linkages between components of the motor system, including the cerebellum, the basal ganglia, and the cerebrum.

In addition, the brainstem contains a diffuse network of neurons known as the *reticular formation*. This network is best known for its role in cortical alertness, ability to direct attention, and sleep. It is also involved with the coordination of orofacial motor activities; in particular, those involved with eating and the generation of emotional facial expressions. Other functions include the coordination of eating and breathing, blood pressure regulation, and the response to pain. In addition to forming a diffuse network with far-reaching functions in the brainstem, some reticular formation neurons may cluster together, forming nuclei or control centers. These centers contribute to the regulation of the gastrointestinal system (including swallowing

and vomiting), the respiratory system (including the initiation and modulation of respiratory rhythm, coughing, hiccupping, and sneezing), and the cardiovascular system.

6.4.7 Cerebellum

The *cerebellum* (Latin, "little brain") is part of the hindbrain and is attached to the dorsal surface of the upper region of the brainstem. Although it constitutes only 10% of the total volume of the brain, it contains more than half of all its neurons. Its surface consists of a thin cortex of gray matter with extensive folding, a core of white matter, and three pairs of nuclei embedded within it. The cerebellum is immature at birth but continues to develop throughout childhood and adolescence.

The specialized function of the cerebellum is to coordinate movement by evaluating differences between intended movement and actual movement. It carries out this activity while a movement is in progress as well as during repetitions of the same movement. Three important aspects of the cerebellum's organization enable it to carry out this function. First, it receives extensive sensory input from somatic receptors in the periphery of the body (proprioceptors) and from receptors in the inner ear providing information regarding equilibrium and balance. Second, output from the cerebellum is transmitted to the premotor and motor systems of the cerebral cortex and the brainstem, systems that control spinal interneurons and motor neurons. Finally, circuits within the cerebellum exhibit significant plasticity, which is necessary for motor adaptation and learning. Examples of motor learning include riding a bicycle, playing a musical instrument, and throwing a football.

The cerebellum consists of three functionally distinct parts:

1. Vestibulocerebellum
2. Spinocerebellum
3. Cerebrocerebellum

The *vestibulocerebellum* receives sensory input regarding motion of the head and its position relative to gravity as well as visual input. Outputs control axial muscles (primarily head and neck) and limb extensors assuring balance while standing still and during movement. Outputs also control eye movements and coordinate movement of the head and eyes. Lesions here affect an individual's balance. The ability to use the incoming sensory information to control eye movements when the head is rotating and movements of the limbs and body during standing and walking is also impaired.

The *spinocerebellum* influences muscle tone and coordinates skilled voluntary movements. It receives sensory input from interneurons in the spinal cord transmitting somatic information, in particular, from muscle and joint proprioceptors providing data regarding body movements and positions that are actually taking place. It also receives input from the cortical motor

areas providing information regarding the intended or desired movement. The spinocerebellum then compares these inputs. If the actual status of a body part differs from the intended status, the spinocerebellum transmits impulses back to the motor areas of the brain to make the appropriate adjustments in the activation of the associated skeletal muscles.

The *cerebrocerebellum* is involved with the planning, programming, and initiation of voluntary activity. It also participates in procedural memories or motor learning. This region of the cerebellum receives input from and provides output to the cortical motor areas directly. Lesions of the cerebrocerebellum cause delays in initiating movements and irregularities in the timing of multistep movements.

Disorders of the human cerebellum result in three types of abnormalities. The first is *hypotonia* or reduced muscle tone. Another includes abnormalities in the execution of voluntary movements or *ataxia* (defective muscular coordination). The third type of muscular malfunction is *intention tremors*. These tremors differ from the resting tremors of Parkinson's disease in that they occur *during* a movement and are most pronounced at the end of the movement when the patient attempts to terminate it.

PHARMACY APPLICATION: CENTRALLY ACTING DRUGS

Combinations of centrally acting drugs are frequently used to achieve a desired therapeutic effect, particularly when the agents used have different mechanisms of action. For example, a patient with Parkinson's disease may be treated with one drug that blocks the effects of the neurotransmitter, acetylcholine, and a second drug that enhances the activity of another neurotransmitter, dopamine. However, potentially detrimental effects may occur when the agents used have additive effects. The effect of a CNS stimulant or depressant is additive with the effects of all other categories of stimulant and depressant drugs. For example, the combination of benzodiazepines (diazepam, Valium) or barbiturates (pentobarbital, Nembutal®) with ethanol is not only additive, it may be fatal. Each of these drugs, especially the barbiturates, has a depressant effect on the respiratory center in the brainstem, such that high doses may cause the cessation of breathing. The effect of a CNS drug is also additive with the physiological state of the patient. For example, anesthetics and antianxiety drugs are less effective in a hyperexcitable patient compared to a healthy patient.

6.5 Glial cells

Glial cells outnumber neurons in the nervous system and account for 90% of the cells within the CNS. However, because they do not branch as much as neurons, these cells occupy only 50% of the volume of the brain. Glial

cells do not transmit electrical impulses. Instead, these cells serve as the connective tissue of the CNS providing structural and metabolic support for the neurons. As such, glial cells maintain an appropriate microenvironment essential for optimal neuronal function.

There are several types of glial cells in the CNS:

- Astrocytes
- Oligodendrocytes
- Microglia
- Ependymal cells

The most abundant type of glial cell is the *astrocyte*. Found throughout the CNS, astrocytes are highly branched cells with processes that extend to the surfaces of neuronal dendrites and cell bodies. Expansions of these processes, referred to as *end-feet,* also form a complete lining around the outer surface of the brain and spinal cord, where these structures come into contact with the pia mater, the innermost meningeal membrane that encloses the CNS. In addition, end-feet are found surrounding the small blood vessels of the CNS. Astrocytes have many important functions within the brain:

- Provide a pathway for neuronal migration during fetal development.
- Enhance synapse formation and synaptic transmission.
- Provide structural support by holding neurons together in their proper framework.
- Induce formation of the blood–brain barrier in CNS capillaries.
- Secrete growth factors necessary for neuronal function.
- Form an astrocytic scar in response to brain injury.
- Maintain electrolyte concentration (in particular, K^+ ion) and the pH of the interstitial fluid in the CNS.
- Participate in neurotransmitter metabolism and removal.

Oligodendrocytes are functionally analogous to Schwann cells in the peripheral nervous system. These glial cells are responsible for the formation of the myelin sheath that surrounds the neuronal axons in the CNS. However, in contrast to the Schwann cell, oligodendrocytes send out several processes that enable them to provide myelin to many axons at the same time.

Microglia are the immune effector cells of the CNS and are the predominant glial cells involved with CNS inflammation. Functionally analogous to macrophages in the periphery of the body, these glial cells serve as phagocytic scavengers in the brain. Like the macrophage, microglia are derived from monocytes. These cells then migrate into the CNS during embryonic development.

In a healthy individual, microglia account for only 1% of the cells in the CNS. In this case, they are considered to be quiet or resting and remain stationary until activated by infection or injury. When activated, the microglia

become highly mobile and migrate toward the site of damage. At that time, they proliferate and phagocytose cellular debris.

Ependymal cells line the internal cavities of the CNS including the ventricles of the brain and the central canal of the spinal cord. As such, these cells form a selectively permeable epithelial layer that separates the fluid compartments of the CNS (CSF versus interstitial fluid). In the ventricles, the ependymal cells contribute to the formation of CSF. Furthermore, these cells contain cilia whose beating contributes to the flow of CSF throughout the ventricles. Ependymal cells also serve as neural stem cells with the potential of forming other glial cells.

6.6 Blood–brain barrier

The movement of substances between the blood and the extracellular fluid surrounding the cells in most tissues of the body occurs very readily. This exchange takes place at the level of the capillaries, the smallest blood vessels in the cardiovascular system whose walls are formed by a single layer of endothelial cells. Lipid-soluble substances are able to move across this layer of endothelial cells at any point because they can move directly *through* the plasma membrane by passing between the phospholipid molecules of the bilayer. The movement of water-soluble substances is limited to the multiple pores found *between* the cells; however, it also takes place rapidly and efficiently.

This nonselective exchange of materials, which includes all substances except plasma proteins, does not occur in all vascular beds, however. Many substances found in the blood are potentially harmful to the CNS. Therefore, the brain and the spinal cord are protected from these substances by the *blood–brain barrier*. In the capillaries of the brain and spinal cord, there are no pores between the endothelial cells. Instead, there are *tight junctions* that fuse the cells together. As mentioned previously, astrocytes also play a role in the formation of the blood–brain barrier. As a result, exchange between the blood and the extracellular fluid of the brain is altered. Lipid-soluble substances, such as oxygen, carbon dioxide, steroid hormones, most anesthetics, and alcohol, continue to move directly through the plasma membrane and, therefore, remain very permeable. Because the blood–brain barrier anatomically prevents the movement of materials between the cells, it is impermeable to water-soluble substances, such as glucose, amino acids, and ions. These substances are exchanged between the blood and the extracellular fluid of the brain by way of highly selective membrane-bound protein carriers.

There are several benefits to the presence of this barrier. It protects the neurons of the CNS from fluctuations in plasma components. For example, a change in the potassium ion concentration could alter neuronal function due to its effect on membrane potential. Second, the barrier minimizes the possibility that harmful blood-borne substances reach the CNS. Finally, it prevents

any blood-borne substances that could function as neurotransmitters from reaching the brain and causing inappropriate neuronal stimulation.

The blood–brain barrier exists in capillaries in all areas of the brain and spinal cord except the hypothalamus and some regions of the brainstem. The absence of the barrier in a given region coincides with the function of that area. For example, the hypothalamus contributes to homeostasis by monitoring the concentration of various blood-borne substances such as glucose and hormones. Glucose and amino acid-derived hormones are hydrophilic and would be unable to come into contact with hypothalamic neurons if the barrier were present. Another instance includes the vomit center of the medulla whose neurons detect the presence of potentially toxic substances in the blood. This center prevents the further absorption of these substances from the gastrointestinal tract by inducing vomiting. Once again, neurons in this region need to be exposed to any hydrophilic toxins in order to carry out this function.

PHARMACY APPLICATION:
ANTIHISTAMINES AND THE BLOOD–BRAIN BARRIER

Antihistamine drugs have long been used to treat symptoms of allergy such as sneezing, itching, watery discharge from the eyes and nose, and possibly wheezing. The older or first-generation histamine H_1 receptor antagonists such as Benadryl® and Tavist® effectively relieve these peripheral symptoms. However, these medications, of the drug class ethanolamines, are very lipophilic and readily cross the blood–brain barrier to interact with histamine H_1 receptors in the CNS as well. As a result, they also cause central effects such as diminished alertness, slowed reaction times, and sedation. The newer or nonsedating antihistamines such as Claritin® and Allegra® have been chemically designed to be less lipophilic and not cross the blood–brain barrier at therapeutic doses. Therefore, these medications, which are of the drug class piperidines, eliminate the peripheral symptoms of allergy without this depression of CNS activity.

6.7 Cerebrospinal fluid

Embedded within the brain are four *ventricles* or chambers that form a continuous fluid-filled system. In the roof of each of these ventricles is a network of capillaries referred to as the *choroid plexus*. It is from the choroid plexuses of the two lateral ventricles (one in each cerebral hemisphere) that *cerebrospinal fluid* is primarily derived. Due to the presence of the blood–brain barrier, the selective transport processes of the choroid plexus determine the composition of the CSF. Therefore, the composition of the CSF is markedly different from the composition of the plasma. However, the CSF is in equilibrium

with the interstitial fluid of the brain and contributes to the maintenance of a consistent chemical environment for the neurons, which serves to optimize their function.

The CSF flows through the ventricles, downward through the central canal of the spinal cord, and then upward back toward the brain through the subarachnoid space that completely surrounds the brain and spinal cord. As the CSF flows over the superior surface of the brain, it leaves the subarachnoid space and is absorbed into the venous system. Although CSF is actively secreted at a rate of 500 ml/day, the volume of this fluid in the system is approximately 140 ml. Therefore, the entire volume of CSF is turned over three to four times per day.

The one-way flow of the CSF and the constant turnover facilitate its important function of *removing potentially harmful brain metabolites*. The CSF also protects the brain from impact by serving as a *shock-absorbing system* that lies between the brain and its bony capsule. Finally, because the brain and the CSF have about the same specific gravity, the brain floats in this fluid. This *reduces the effective weight of the brain* from 1400 grams to less than 50 grams and prevents compression of neurons on the inferior surface of the brain.

Medical terminology

Afferent neuron (ăf′ĕr-ĕnt nĕr′ŏn): Nerve cell that transmits sensory information from the periphery of the body to the CNS.

Aphasia (ă-fā′zē-ă): Impairment in the ability to communicate through language.

Ataxia (ā-tăk′sē-ă): Defective skeletal muscle coordination.

Disproportionate (dĭs-prō-pŏr′shŭn-ăt): Unequal in size.

Efferent (ĕf′ĕr-ĕnt) neuron: Nerve cell that transmits motor information from the central nervous system to muscle cells or glands.

Fissure (fĭs′ĕr): Deep furrow in the brain.

Ganglion (găng′lē-ŏn): Mass of neuronal cell bodies in the peripheral nervous system.

Gray matter (grā măt′ĕr): Regions in cerebrum and cerebellum containing neuronal cell bodies and unmyelinated neurons.

Gyrus (gī′rŭs): Convolution of the cerebral cortex.

Hyperkinesia (hī″pĕr-kĭ-nē′zē-ă): Enhanced motor response or activity.

Hypokinesia (hī″pō- kĭ-nē′zē-ă): Reduced motor response or activity.

Hypotonia (hī″pō-tō′nē-ă): Reduced muscle tone.

Interneuron (ĭn′tĕr-nĕr″ŏn): Nerve cell located entirely within the central nervous system.

Multimodal (mŭlt″ē-mōd′ăl): Of more than one type.

Nucleus (nū′klē-ŭs): Mass of neuronal cell bodies in the central nervous system.

Plasticity (plăs-tĭs′ĭ-tē): Ability to be changed in form.

Proprioception (prō′prē-ō-cĕp″shŭn): Awareness of body position and movement.

Reflex (rē′flĕks): Predictable, stereotyped, involuntary response to a given stimulus.

Sulcus (sŭl′kŭs): Small groove or depression of the cerebral cortex.

Unimodal (ūn″ē-mōd′ăl): Of one type.

White matter (whīt măt′ĕr): Regions in cerebrum and cerebellum consisting of myelinated axons of neurons.

Bibliography

Amaral, D. G., The Anatomical Organization of the Nervous System, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 17.

Bear, M. F., Connors, B. W., and Paradiso, M. A., *Neuroscience, Exploring the Brain,* 3rd ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2007.

Bloom, F. E., Neurotransmission and the Central Nervous System, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 12.

Garoutte, B., *Neuromuscular Physiology*, 4th ed., Mill Valley Medical, Millbrae, CA, 1996.

Ghez, C., and Thach, W. T., The Cerebellum, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 42.

Laterra, L., and Goldstein, G. W., Ventricular Organization of Cerebral Spinal Fluid: Blood-Brain Barrier, Brain Edema, and Hydrocephalus, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, appendix B.

Lynch, J. C., The Cerebral Cortex, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 32.

Ma, T. P., The Basal Nuclei, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 26.

Mihailoff, G. A., and Haines, D. E., The Diencephalon, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 15.

Mihailoff, G. A., and Haines, D. E., Motor System II: Corticofugal Systems and the Control of Movement, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 25.

Naftel, J. P., Ard, M. D., Fratkin, J. D., and Hutchins, J. B., The Cell Biology of Neurons and Glia, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 2.

Nieuwenhuys, R., Ten Donkelaar, H. J., and Nicholson, C., *The Central Nervous System of Vertebrates*, vol. 3, Springer, Berlin, 1998.

Rechtschaffen, A., and Siegel, J., Sleep and Dreaming, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 47.

Saper, C. B., Brain Stem, Reflexive Behavior, and the Cranial Nerves, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 44.

Saper, C. B., Iversen, S., and Frackowiak, R., Integration of Sensory and Motor Function: The Association Areas of the Cerebral Cortex and the Cognitive Capabilities of the Brain, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 19.

Sherwood, L., *Human Physiology: From Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Sherwood, L., Klandorf, H., and Yancey, P. H., *Animal Physiology: From Genes to Organisms*, Brooks/Cole, Pacific Grove, CA, 2005.

Silverthorn, D. U., *Human Physiology: An Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Skidgel, R. A., and Erdos, E. G., Histamine, Bradykinin and Their Antagonists, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 24.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Warren, S., Capra, N. F., and Yezierski, R. P., The Somatosensory System I: Tactile Discrimination and Position Sense, in *Fundamental Neuroscience for Basic and Clinical Applications,* Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 17.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function*, McGraw-Hill, Boston, MA, 2006.

chapter seven

The spinal cord

Study objectives

- Define *cauda equina* and explain how it is formed
- Distinguish between a nerve and a tract
- Explain the function of the gray matter of the spinal cord
- Describe the location and function of each of the four types of neurons found in the gray matter of the spinal cord
- Explain the function of the white matter of the spinal cord
- Describe the composition of the ascending tracts including both the origin and termination of each of the neurons
- Distinguish between corticospinal tracts and multineuronal tracts
- Discuss the mechanisms by which spinal anesthesia and epidural anesthesia exert their effects
- Define the various categories of reflexes
- List the components of the reflex arc
- Explain the mechanism of the withdrawal reflex
- Explain the mechanism of the crossed-extensor reflex

7.1 Introduction

The lowest level of the central nervous system (CNS), both anatomically and functionally, is the *spinal cord*. Continuous with the brainstem, it exits the skull through the *foramen magnum*. The spinal cord then passes through the *vertebral canal* of the vertebral column to the level of the first or second lumbar vertebrae.

The spinal cord is divided into four anatomical regions: cervical, thoracic, lumbar, and sacral. These regions are named according to the vertebrae adjacent to them during embryonic development. Each region is subdivided into functional segments. A pair of spinal nerves extends from each segment (one nerve from the left side of the spinal cord and one nerve from the right side) and exits the CNS through the *intervertebral foramina*, or openings between adjacent vertebrae. There are a total of 31 pairs of spinal nerves:

- 8 Cervical
- 12 Thoracic
- 5 Lumbar
- 5 Sacral
- 1 Coccygeal

In terms of sensory input, each specific region of the body surface innervated by a given spinal nerve is referred to as a *dermatome.* These spinal nerves may also branch off and innervate a deeper internal organ. Dermatomes are highly organized and begin at the level of the head and proceed downward to the feet. Interestingly, dermatomes vary significantly in size and shape.

Spinal nerves arising from the *cervical level* of the cord are involved with sensory perception and motor function to dermatomes located at the back of the head, the neck, and the arms. Nerves arising from the *thoracic level* innervate dermatomes found on the anterior surface of the trunk (chest and abdomen) as well as the upper back. Spinal nerves from the *lumbar region* innervate dermatomes of the lower back and the anterior surface of the legs and feet. Finally, spinal nerves from the *sacral region* of the cord innervate dermatomes located at the external genitalia, the buttocks, the posterior surface of the legs, and the bottom of the feet.

Lesions of the spinal cord interrupt sensation and motor function. The affected regions of the body are those innervated by spinal nerves below the level of the lesion. For example, damage to the spinal cord at the functional level of T_{12}, or the twelfth thoracic spinal nerve, would result in paralysis and numbness of the lower back and legs. Interestingly, the phrenic nerves that innervate the diaphragm, the major muscle of inspiration, arise from spinal cord segments C_3 through C_5. Therefore, only lesions high in the cervical region will affect breathing.

The human spinal cord and the vertebral column initially grow at the same rate during embryonic development. In this way, the spinal segments and the vertebral bones for which they are named are aligned. Therefore, the spinal nerves emerge from the vertebral column through the intervertebral foramina at the same level as the spinal cord segment from which they arise. However, after the third month of gestation, each vertebral bone becomes larger compared to the associated spinal segment. Therefore, the vertebral column grows approximately 25 cm longer than the spinal cord. (This explains why the spinal cord extends only as far as the upper lumbar vertebrae.) As a result, the spinal cord segment from which each pair of spinal nerves arise is no longer aligned with its associated vertebral bone. Because the vertebral column is now longer than the spinal cord, the intervertebral foramina have shifted downward relative to their corresponding spinal cord segment. Therefore, the spinal nerve roots arising from each segment must extend downward through the vertebral canal to reach their points of exit. This is the case especially for the spinal nerves arising from the lumbar and sacral regions of the cord. As a result, only spinal nerve roots are found in the vertebral canal below the level of the first or second lumbar vertebrae. Because of its appearance, this bundle of nerve roots is collectively referred to as the *cauda equina*, or "horse's tail." A sample of cerebrospinal fluid may

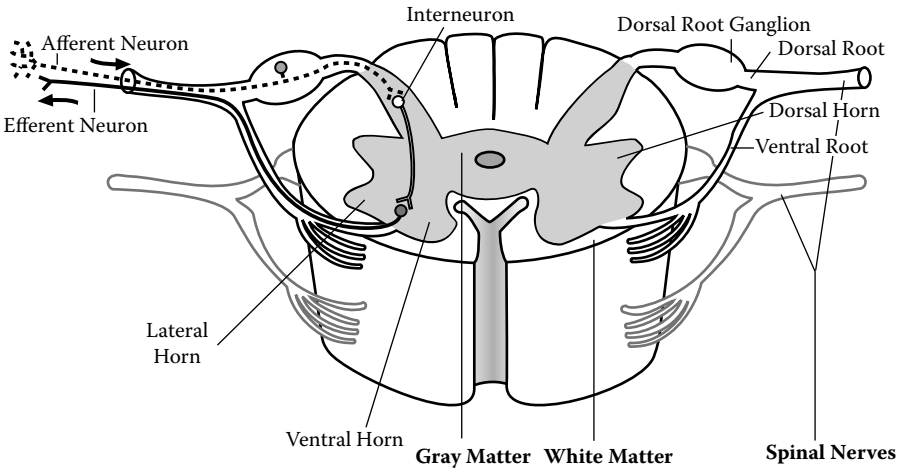


Figure 7.1 Cross-sectional view of the spinal cord. In contrast to the brain, the gray matter of the spinal cord is located internally, surrounded by the white matter. The gray matter consists of nerve cell bodies and unmyelinated interneuron fibers. This component of the spinal cord is divided into three regions: the dorsal horn, the lateral horn, and the ventral horn. The white matter consists of bundles of myelinated axons of neurons, or tracts. Each segment of the spinal cord gives rise to a pair of spinal nerves. These nerves contain both afferent and efferent neurons. Afferent neurons enter the spinal cord through the dorsal root, and efferent neurons exit the spinal cord through the ventral root.

be obtained from this region by way of a *lumbar puncture* or "spinal tap." A needle may be safely inserted into the vertebral canal without the possibility of penetrating the spinal cord. The spinal nerve roots are easily pushed aside by the needle, significantly reducing the possibility of puncturing one of these nerves.

The spinal nerves associate with the spinal cord by way of two branches, or roots: the *dorsal root* and the *ventral root*. The *dorsal root* contains afferent, or sensory, neurons. Impulses in these neurons travel from peripheral tissues toward the spinal cord. The dorsal root joins the spinal cord laterally, toward the posterior surface of the cord (Figure 7.1). The *ventral root* contains efferent, or motor, neurons. Impulses in these neurons travel away from the spinal cord toward the peripheral tissues. The ventral root exits the spinal cord laterally, toward the anterior surface of the cord.

At this point, it is important to note that a *nerve* is defined as a bundle of neuronal axons; some are afferent and some are efferent. A nerve does not consist of entire neurons, only their axons. Furthermore, nerves are found only in the peripheral nervous system (PNS). Bundles of neurons with similar functions located within the CNS are referred to as *tracts*. Therefore, technically speaking, there are no nerves within the brain or the spinal cord.

7.2 Functions of the spinal cord

The spinal cord is responsible for two vital CNS functions:

1. It conducts nerve impulses to and from the brain.
2. It processes sensory input from the skin, joints, and muscles of the trunk and limbs and initiates reflex responses to this input.

7.3 Composition of the spinal cord

The spinal cord consists of gray matter and white matter. The *gray matter* is composed of nerve cell bodies and unmyelinated interneuron fibers. The location of the gray matter in the spinal cord is opposite to that of the brain. In the brain, the gray matter of the cerebrum and the cerebellum is found externally forming a cortex, or covering, over the internally located white matter. In the spinal cord, the gray matter is found internally and is surrounded by the white matter.

The *white matter* is composed of the myelinated axons of neurons. These axons are grouped together according to function to form tracts. Neurons transmitting impulses toward the brain in the *ascending tracts* carry sensory information. Neurons transmitting impulses away from the brain in the *descending tracts* carry motor information.

7.3.1 Gray matter

A cross-sectional view of the spinal cord reveals that the gray matter has a butterfly or H-shape (see Figure 7.1). As such, on each side of the spinal cord, the gray matter is divided into three regions:

1. Dorsal horn (posterior, toward the back)
2. Ventral horn (anterior, toward the front)
3. Lateral horn (toward the side)

There are 1 billion neurons in the human spinal cord. Accordingly, each spinal segment contains millions of neurons within the gray matter. Functionally, there are four types of neurons (Table 7.1):

1. Second-order sensory neurons
2. Somatic motor neurons
3. Visceral motor neurons
4. Interneurons

The cell bodies of *second-order sensory neurons* are found in the dorsal horn. These neurons receive input from afferent neurons (*first-order sensory neurons*) entering the CNS from the periphery of the body through the dorsal

Table 7.1 Neurons Associated with the Spinal Cord

Neuron	Location of Cell Body	Function
First-order sensory	Dorsal root ganglion	Transmit sensory information toward spinal cord
Second-order sensory	Dorsal horn	Transmit sensory information to higher levels of the central nervous system
Somatic (α) motor	Ventral horn	Innervate skeletal muscle to cause contraction
Visceral motor	Lateral horn	Innervate cardiac muscle, smooth muscle, and glands to carry out autonomic nervous system functions
Interneurons	Throughout spinal cord gray matter	Interconnect sensory and motor neurons; integrative functions such as reflexes

root of the spinal nerve. The function of the second-order sensory neuron is to transmit nerve impulses to higher levels in the CNS. The axons of these neurons leave the gray matter and travel upward in the appropriate ascending tracts of the white matter.

The cell bodies of *somatic motor neurons* are found in the ventral horn. The axons of these neurons exit the CNS through the ventral root of the spinal nerve and innervate skeletal muscles. There are two types of motor neurons located in the ventral horn:

1. *Alpha motor neurons*: Innervate skeletal muscle fibers to cause contraction.
2. *Gamma motor neurons*: Innervate intrafusal fibers of the muscle spindle to cause contraction that allows the muscle spindle to remain sensitive to changes in muscle length.

The spatial organization of the cell bodies of the motor neurons follows a *proximal-distal rule*. Motor neurons that innervate the most proximal muscles (axial muscles of the neck and trunk) lie most medially in the gray matter. Motor neurons that innervate the most distal muscles (wrists, ankles, digits) lie most laterally in the gray matter.

The cell bodies of *visceral motor neurons* are found in the lateral horn. The axons of these neurons form the efferent nerve fibers of the autonomic nervous system (ANS). The ANS innervates cardiac muscle, smooth muscle, and glands (see Chapter 9). The axons of these neurons exit the spinal cord by way of the ventral root.

Interneurons are found in all areas of the spinal cord gray matter. These neurons are quite numerous, small, and highly excitable. Interneurons have many interconnections. They receive input from higher levels of the CNS as

well as from sensory neurons entering the CNS through the spinal nerves. Many interneurons in the spinal cord synapse with motor neurons in the ventral horn. These interconnections are responsible for the integrative functions of the spinal cord including reflexes.

Afferent neurons that transmit sensory information toward the spinal cord are referred to as *first-order sensory neurons*. The cell bodies of these neurons are found in the *dorsal root ganglia*. These ganglia form a swelling in each of the dorsal roots just outside of the spinal cord. The portion of the axon between the distal receptor and the cell body is referred to as the *peripheral axon,* and the portion of the axon between the cell body and the axon terminal within the CNS is referred to as the *central axon* (Figure 7.1).

Upon entering the spinal cord, the first-order sensory neurons may enter the gray matter. The neurons may then synapse with one or more of the following neurons:

- *Second-order sensory neuron*: Transmits impulses to higher levels of the CNS.
- *Alpha motor neuron*: Transmits impulses to skeletal muscles.
- *Interneurons*: Transmits impulses to motor neurons.

Synapses between the first-order sensory neurons and the alpha motor neurons, either directly or by way of interneurons, result in spinal cord reflexes. Reflexes are discussed in more detail in Section 7.4.

Alternatively, the first-order sensory neurons may initially enter the white matter of the spinal cord. In this case, the axons of these neurons may ascend the cord to the medulla or travel up or down the cord to a different spinal segment. Upon reaching its destination, the axon then enters the gray matter of the spinal cord and synapses with one or more of the neurons discussed above.

7.3.2 White matter

The white matter of the spinal cord consists of the myelinated axons of neurons. These axons may travel up the spinal cord to a higher spinal segment or to the brain. On the other hand, these axons may travel down the spinal cord to a lower spinal segment. The axons of neurons that carry similar types of impulses are bundled together to form tracts. *Ascending tracts* carry sensory information from the spinal cord toward the brain. *Descending tracts* carry motor impulses from the brain toward the interneurons of the spinal cord or the motor neurons in the lateral or ventral horns of the spinal cord gray matter. In general, these tracts are named according to their origin and termination. For example, the ventral spinocerebellar tract is an ascending tract carrying information regarding unconscious muscle sense (proprioception) from the spinal cord to the cerebellum. On the other

hand, the ventral corticospinal tract is a descending tract carrying information regarding voluntary muscle control from the cerebral cortex to the spinal cord.

7.3.2.1 Ascending tracts

These tracts contain three successive neurons:

1. First-order neurons
2. Second-order neurons
3. Third-order neurons

As discussed, the *first-order neuron* is the afferent neuron that transmits impulses from a peripheral receptor toward the CNS. Its cell body is located in the dorsal root ganglion. This neuron synapses with the *second-order neuron* whose cell body is located in the dorsal horn of the spinal cord or in the medulla of the brainstem. The second-order neuron then synapses with the *third-order neuron* whose cell body is located in the thalamus. Limited processing of sensory information takes place in the thalamus. Finally, the third-order neuron terminates in the somatosensory cortex where more complex cortical processing begins. Functions of the thalamus and the cerebral cortex are discussed in detail in Chapter 6.

All ascending tracts cross to the opposite side of the CNS. For example, sensory input entering the left side of the spinal cord ultimately terminates on the right side of the cerebral cortex. These tracts may cross at the level of entry into the spinal cord, a few segments above the level of entry, or within the medulla of the brainstem. The locations of specific ascending tracts are illustrated in Figure 7.2. A summary of the functions of each of these tracts is found in Table 7.2.

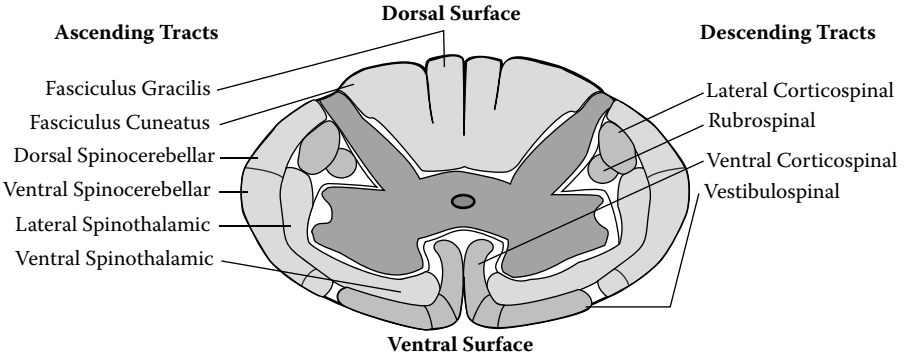


Figure 7.2 Ascending and descending tracts in the white matter of the spinal cord. Tracts are formed of bundles of neuronal axons that transmit similar types of information.

Table 7.2 Ascending and Descending Tracts in the White Matter of the Spinal Cord

Ascending Pathway	Function
Fasciculus Gracilis	Fine touch discrimination (ability to recognize the size, shape, and texture of objects and their movement across the skin), proprioception, vibration from legs and lower trunk; crossed
Fasciculus Cuneatus	Fine touch discrimination, proprioception, vibration from neck, arms, upper trunk; crossed
Dorsal Spinocerebellar	Proprioception (important for muscle tone and posture); uncrossed
Ventral Spinocerebellar	Proprioception; crossed
Lateral Spinothalamic	Pain, temperature; crossed
Ventral Spinothalamic	Light touch, pressure; crossed
Descending Pathway	**Function**
Lateral Corticospinal	Voluntary control of skeletal muscles; crossed
Rubrospinal	Originates in brainstem, subconscious control of skeletal muscle (muscle tone, posture); crossed
Ventral Corticospinal	Voluntary control of skeletal muscles; uncrossed
Vestibulospinal	Originates in brainstem, subconscious control of skeletal muscle (muscle tone, balance, equilibrium); uncrossed

PHARMACY APPLICATION: SPINAL ANESTHESIA

The injection of a local form of anesthetic into the cerebrospinal fluid surrounding the spinal cord causes spinal anesthesia. This injection is made below the level of the second lumbar vertebra in order to minimize direct nerve trauma. Spinal anesthesia is effective in the control of pain during lower body surgical procedures, such as knee surgery. Currently, the drugs most commonly used in the United States include lidocaine, bupivacaine, and tetracaine. The choice of anesthetic is determined by the duration of anesthesia required. Lidocaine is used for short procedures, bupivacaine is chosen for procedures of intermediate length, and tetracaine is used for procedures of long duration.

The mechanism of action of these anesthetics involves the blockade of sodium channels in the membrane of the second-order sensory neuron. The binding site for these anesthetics is on a subunit of the sodium channel located near the internal surface of the cell membrane. Therefore, the agent must enter the neuron in order to effectively block the sodium channel. Without the influx of sodium, neurons cannot depolarize and generate an action potential (see Chapter 4). Therefore, the second-order sensory neuron cannot be stimulated by impulses elicited by pain receptors associated with the first-order sensory neuron.

In other words, the pain signal is effectively interrupted at the level of the spinal cord and does not travel any higher in the CNS. In this way, the brain does not perceive pain.

Interestingly, the second-order sensory neurons are the neurons of the spinal cord gray matter which are most susceptible to the effects of spinal anesthesia. These neurons have a small diameter and are unmyelinated. The small diameter allows the drug to locate its binding site on the sodium channel more readily due to a higher concentration of the drug within the neuron. Furthermore, unmyelinated neurons have a greater number of sodium channels located over a larger surface area. Alpha motor neurons in the ventral horn are susceptible to these anesthetics only at high doses. This is because alpha motor neurons have a large diameter and are myelinated. The larger diameter results in a lower concentration of the drug within the neuron. Myelination limits the number and availability of sodium channels upon which the anesthetic can exert its effect.

7.3.2.2 Descending tracts

Voluntary movement of skeletal muscles is controlled by two types of descending tracts. Neurons in both of these tracts terminate on and influence the activity of the alpha motor neurons in the ventral horn. The two types of tracts include the *corticospinal* (*pyramidal*) *tracts* and the *multineuronal* (*extrapyramidal*) *tracts*.

The *corticospinal (pyramidal) tracts* originate in the cerebral cortex. Due to their triangular shape, the neurons of the primary motor cortex are referred to as *pyramidal cells*. Most of the axons of these neurons descend directly to the alpha motor neurons in the spinal cord. In other words, these are primarily monosynaptic pathways. This type of synaptic connection is particularly important for the movement of individual fingers. A primary function of these tracts is to regulate fine, discrete, voluntary movements of the hands and fingers. Recall the extensive area of the cerebral cortex devoted to the hands and fingers, which indicates the density of innervation of the muscles in these tissues (Chapter 6).

Another descending pathway that accompanies the corticospinal tracts is the *corticobulbar pathway.* This pathway begins in the cerebral cortex and travels to the brainstem (bulbar, pertaining to the brainstem). This pathway influences the neurons that innervate muscles of the eyes, face, tongue, and throat. As such, the corticobulbar pathway is the primary source of control for voluntary movement of the head and neck.

As with the ascending tracts, descending tracts cross from one side of the CNS to the other. Most of the tracts cross over in the medulla of the brainstem. Therefore, the right side of the brain influences the activity of the alpha motor neurons and, therefore, the skeletal muscles on the left side of the body.

The *multineuronal* (*extrapyramidal*) *tracts* originate in many regions of the brain including the motor regions of the cerebral cortex, the cerebellum, and the basal ganglia. Impulses from these various regions are transmitted to nuclei in the brainstem, in particular, the reticular formation and the vestibular nuclei. The axons of neurons in these nuclei descend to the alpha motor neurons in the spinal cord. Therefore, in contrast to the corticospinal tracts, these pathways are polysynaptic. The multineuronal tracts regulate overall body posture, balance, and walking. Specifically, these tracts control subconscious movements of large muscle groups in the trunk and in the limbs. Some of the pathways originating in the brainstem cross to the other side of the spinal cord to affect muscles on the opposite side of the body. However, most remain uncrossed (e.g., vestibulospinal tract) (Table 7.2).

These two types of descending motor tracts do not function in isolation. They are extensively interconnected and cooperate in the control of movement. For example, in order to grasp a doorknob to open a door, there is subconscious positioning of the body to face the door (multineuronal tracts) and extension of the arm toward the doorknob (corticospinal tracts).

The locations of specific descending tracts are illustrated in Figure 7.2. A summary of the functions of each of these tracts is found in Table 7.2.

PHARMACY APPLICATION: EPIDURAL ANESTHESIA

Epidural anesthesia is administered by injecting local anesthetic into the epidural space. Located outside of the spinal cord on its dorsal surface, the epidural space contains fat and is highly vascular. Therefore, this form of anesthesia can be administered safely at any level of the spinal cord. Furthermore, a catheter may be placed into the epidural space allowing for either continuous infusions or repeated bolus administrations of anesthetic.

The primary site of action of epidurally administered agents is on the spinal nerve roots. As with spinal anesthesia, the choice of drug to be used is determined primarily by the duration of anesthesia desired. However, when a catheter has been placed, short-acting drugs can be administered repeatedly. Bupivacaine is typically used when a long duration of surgical block is needed. Lidocaine is used most often for intermediate length procedures. Chloroprocaine is used when only a very short duration of anesthesia is required.

An important difference between epidural anesthesia and spinal anesthesia is that agents injected into the epidural space may readily enter the blood due to the presence of a rich venous plexus in this area. This is an important consideration when epidural anesthesia is used to control pain during labor and delivery. The agents used are able to cross the placenta, enter the fetal circulation, and exert a depressant effect on the neonate.

7.4 Spinal reflexes

Reflexes may be classified in several ways. They may be named according to the effector tissues that carry out the reflex response:

- *Skeletal muscle reflexes*: Control skeletal muscles.
- *Autonomic reflexes*: Control cardiac muscle, smooth muscle, and glands.

They may be named according to the region of the CNS that integrates incoming sensory information and elicits the reflex response:

- *Cranial reflexes*: Processed within the brain.
- *Spinal reflexes*: Processed at the level of the spinal cord.

Finally, reflexes may be either innate or learned:

- *Simple* or *basic reflexes*: Preprogrammed (built-in), unlearned responses (e.g., blinking, pulling a hand from a hot surface).
- *Acquired* or *conditioned reflexes*: Learned responses that require experience or training (e.g., driving a car, catching a ball).

This section examines the mechanism of simple spinal reflexes that control skeletal muscles.

A reflex occurs when a particular stimulus always elicits a particular response. This response is automatic and involuntary. In other words, it occurs without conscious effort. Therefore, reflexes are specific and predicable. Furthermore, they are often purposeful. For example, the withdrawal reflex causes a body part to be pulled away from a painful stimulus. In this way, tissue injury is avoided.

Spinal reflexes require no input from the brain as they are elicited entirely at the level of the spinal cord. However, while the reflex is underway, nervous impulses are also transmitted to the brain for further processing. In fact, input from the brain may modulate a reflex or alter the response to a stimulus through conscious effort.

A reflex response requires an intact neural pathway between the stimulated area and the responding muscle. This pathway is referred to as a *reflex arc* and includes the following components (see Figure 7.3):

- Sensory receptor
- Afferent or first-order sensory neuron
- Integrating center in the spinal cord (synapses)
- Efferent or motor neuron
- Effector tissue (skeletal muscle)

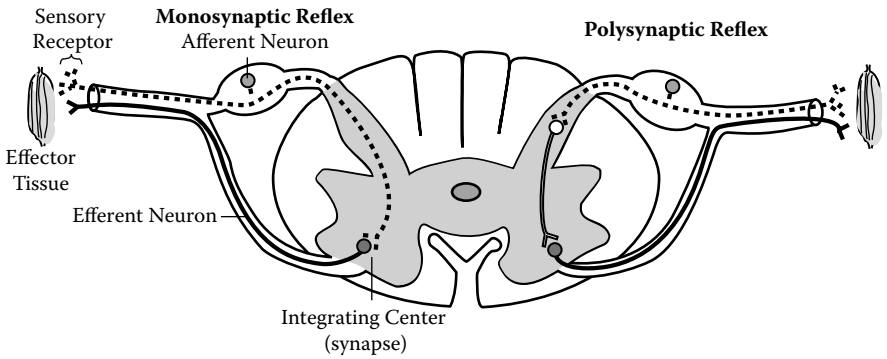


Figure 7.3 Components of a reflex arc. As illustrated by the components of the reflex arc, reflexes may be processed entirely at the level of the spinal cord with no need for input from the brain. A monosynaptic reflex has a single synapse between the afferent and the efferent neurons. A polysynaptic reflex has two or more synapses between the afferent and efferent neurons. In this case, interneurons lie between the sensory and motor neurons. The more interneurons involved, the more complex is the response.

A reflex is initiated by the stimulation of a *sensory receptor* located at the peripheral ending of an afferent or first-order sensory neuron. This *afferent neuron* transmits impulses to the spinal cord. Within the gray matter of the spinal cord, the afferent neuron synapses with other neurons. As such, the spinal cord serves as an *integrating center* for the sensory input. The afferent neuron synapses with the efferent or *motor neuron*. When the afferent neuron synapses directly with the motor neuron, it forms a *monosynaptic reflex*. An example of this type of reflex is the stretch reflex. When the afferent neuron synapses with an interneuron that then synapses with the motor neuron, it forms a *polysynaptic reflex*. An example of this type of reflex is the withdrawal reflex. Most reflexes are polysynaptic. The motor neuron then exits the spinal cord to innervate an *effector tissue* that carries out the reflex response.

7.4.1 Withdrawal reflex

The *withdrawal reflex* is elicited by a painful or tissue-damaging stimulus. The response is to quickly move the body part away from the source of the stimulus, usually by flexing a limb. Any of the major joints and, therefore, muscle groups, may be involved in a reflex depending upon the point of stimulation. For example, all of the joints of a limb are involved when a digit, such as a finger, is stimulated (e.g., finger, wrist, elbow, shoulder). Furthermore, the withdrawal reflex is a very powerful reflex and may override other nervous impulses, such as those regarding locomotion, or walking.

An example of the mechanism of the withdrawal reflex is illustrated in Figure 7.4. When a painful stimulus activates a sensory receptor on the right

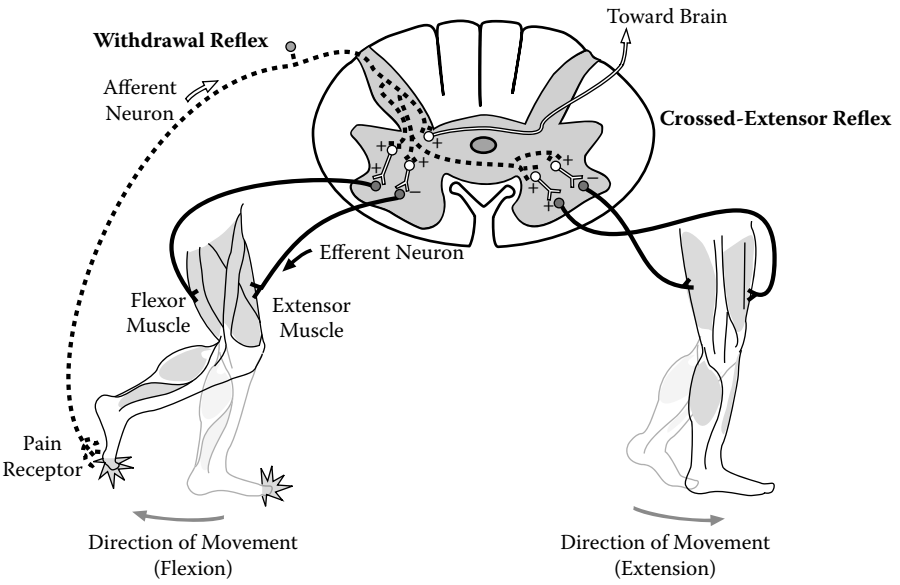


Figure 7.4 The withdrawal reflex coupled with the crossed-extensor reflex. A painful stimulus will elicit the withdrawal reflex. This reflex causes flexor muscles to contract and move the affected body part away from the stimulus. At the same time, the crossed-extensor reflex causes extensor muscles in the opposite limb to contract. The straightening of the opposite limb provides support for the body.

foot, action potentials are transmitted along the afferent neuron to the spinal cord. By way of divergence, this neuron synapses with several other neurons within the gray matter of the spinal cord:

- Excitatory interneuron
- Inhibitory interneuron
- Second-order sensory neuron

The *excitatory interneuron* then synapses with the alpha motor neuron that innervates the *flexor muscles* of the right leg. Consequently, stimulation of the excitatory interneuron leads to stimulation of the alpha motor neuron which then stimulates the flexor muscles to contract and pick up or withdraw the foot from the painful stimulus.

The *inhibitory interneuron* synapses with the alpha motor neuron that innervates the *extensor muscles* of the right leg. Therefore, stimulation of the inhibitory interneuron leads to the inhibition of the alpha motor neuron. As a result, the extensor muscles relax.

The flexor muscles and the extensor muscles are *antagonistic*. In other words, they cause opposite effects. Therefore, when one of these groups of muscles is activated, the other group must be inhibited. This is referred to as

reciprocal inhibition. In this way, activation of the withdrawal reflex leads to unimpeded flexion.

The *second-order sensory neuron* transmits impulses ultimately to the left side of the brain. This permits the awareness of pain, the identification of its source, and, if necessary, postural adjustment. As discussed, impulses in this pathway do not play a role in the reflex per se.

7.4.2 Crossed-extensor reflex

Where appropriate, the withdrawal reflex may be accompanied by the *crossed-extensor reflex*. In the example discussed, when the right leg is flexed or lifted, the left leg must be extended or straightened in order to support the body. In addition to stimulating interneurons on the right side of the spinal cord to influence skeletal muscle activity on the right side of the body, the afferent neuron may also stimulate interneurons on the left side of the spinal cord to influence skeletal muscle activity on the left side of the body. Once again, both excitatory and inhibitory interneurons are involved. However, in this case, these interneurons influence the activity of the opposite muscle groups. Stimulation of the excitatory interneuron on the left side of the spinal cord leads to the stimulation of the alpha motor neuron that innervates the extensor muscles. This causes the left leg to straighten. Stimulation of the inhibitory interneuron on the left side of the spinal cord leads to the inhibition of the alpha motor neuron that innervates the flexor muscles. This results in unimpeded extension of the left leg and the support of the body during the withdrawal of the right leg.

Medical terminology

Anesthesia (ăn″ĕs-thē′zē-ă): Localized or generalized loss of sensation following the administration of an anesthetic agent.

Antagonist (ăn-tăg′ō-nĭst): A muscle that counteracts the function of the prime mover or agonist muscle (e.g., the triceps brachii muscles are antagonistic to the action of the biceps brachii muscles).

Cauda equina (kaw′dă ē-kwīn′ă): Bundle of spinal nerve roots extending inferiorly through the vertebral canal below the level of the second lumbar vertebra.

Dermatome (dĕr′mă-tōm): A region of the body surface supplied by a single spinal nerve.

Epidural space (ĕp″ĭ-dūr′ăl spās): Area between the dura mater (outer meningeal covering of the spinal cord) and the vertebrae which is filled with fatty tissue and blood vessels.

Foramen magnum (for-ā′mĕn măg′nŭm): Large opening in the base of the occipital bone through which the spinal cord exits the skull.

Nerve (nĕrv): Bundle of neurons that transmits electrical impulses from the CNS (brainstem or spinal cord) to the organs and tissues of the body.

Reciprocal inhibition (rĭ-sĭp′rō-kăl ĭn″hĭ-bĭsh′ŭn): Inhibition of a muscle or muscle group that is antagonistic to the muscle or muscle group being stimulated.

Tract (trăkt): Bundle of neurons that transmits electrical impulses within the CNS (brain and spinal cord).

Unimpeded (ŭn-ĭm-pēd′ĕd): Unobstructed.

Bibliography

Amaral, D. G., The Anatomical Organization of the Nervous System, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 17.

Amaral, D. G., The Functional Organization of Perception and Movement, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 18.

Carroll, E. W., and Curtis, R. L., Organization and Control of Neural Function, in *Pathophysiology, Concepts of Altered Health States,* 7th ed., Porth, C. M., Ed., Lippincott-Raven, Philadelphia, PA, 2005, chap. 49.

Catterall, W. A., and Mackie, K., Local Anesthetics, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 14.

DeLong, M. R., The Basal Ganglia, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 43.

Ghez, C., and Krakauer, J., The Organization of Movement, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 33.

Krakauer, J., and Ghez, C., Voluntary Movement, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 38.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Shier, D., Butler, J., and Lewis, R., *Hole's Human Anatomy and Physiology,* 11th ed., McGraw-Hill, New York, 2007.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Turley, S., *Medical Language,* Pearson/Prentice Hall, Upper Saddle River, NJ, 2007.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter eight

Pain

Study objectives

- Describe the three types of nociceptors and the stimuli that activate them
- Distinguish between A-delta fibers and C fibers
- Compare and contrast fast pain and slow pain
- Distinguish between primary hyperalgesia and centrally mediated hyperalgesia
- Discuss the functions of the neurotransmitters, glutamate, and substance P
- Describe the pain pathway and the role that each stimulated region of the brain plays in the response to pain
- Explain how the endogenous analgesic system suppresses pain
- Describe the gate control theory and its contribution to the suppression of pain
- Distinguish between cutaneous pain and deep visceral pain
- Describe the mechanisms by which tissue ischemia and muscle spasm lead to pain
- Discuss the potential causes of visceral pain
- Describe the mechanism of referred pain
- Explain how phantom pain occurs
- List the properties of an ideal analgesic medication
- List and discuss the effects of nonnarcotic analgesic medications
- Discuss the effects of opioid analgesic medications
- Discuss the effects of adjuvant medications

8.1 Introduction

Sensations interpreted as *pain*, including burning, aching, stinging, and soreness, are the most distinctive forms of sensory input to the central nervous system. Pain serves an important protective function as it causes awareness of actual or potential tissue damage. Furthermore, it stimulates an individual to react to remove or withdraw from the source of the pain. Unlike other forms of sensory input, such as vision, hearing, and smell, pain has an urgent, primitive quality. This quality is responsible for the behavioral and emotional aspects of pain perception.

8.2 Nociceptors

Nociceptors (Latin *nocere,* "to hurt") are bare or free nerve endings. Therefore, they do not adapt, or stop responding, to sustained or repeated stimulation. This is beneficial, in that it keeps the individual aware of the damaging stimulus for as long as it persists. Nociceptors are widely distributed in the skin, dental pulp, bone, joints, muscle, blood vessels, meninges, and some internal organs. Stimulation of nociceptors involves the activation, or opening, of ion channels that alter membrane potential.

There are three major classes of nociceptors:

1. Thermal nociceptors
2. Mechanical nociceptors
3. Polymodal nociceptors

Thermal nociceptors are activated by extreme temperatures, especially heat. One group of these receptors is stimulated by noxious heat (>45°C). Heat-sensitive ion channels on these nociceptors are activated at this temperature causing depolarization and the generation of an action potential. A second group of thermal nociceptors is stimulated by noxious cold (<5°C). These are the temperatures at which the tissues begin to be damaged.

Mechanical nociceptors are activated by mechanical damage, such as cutting, pinching, or tissue distortion. They are also activated by intense pressure applied to the skin. The simple distortion of the nociceptor membrane activates mechanically gated ion channels, which results in depolarization and the generation of an action potential. Their firing rates increase with the destructiveness of the mechanical stimulus.

The majority of nociceptors are the *polymodal nociceptors*, which are activated by all types of damaging stimuli (thermal, mechanical, and chemical), including irritating exogenous substances that may penetrate the skin. Endogenous substances that may stimulate these receptors to elicit pain include potassium released from damaged cells, bradykinin, histamine, substance P, acids, and proteolytic enzymes (see Table 8.1). With the polymodal nociceptors, the chemicals may activate ion channels on the nociceptors to cause depolarization and generate an action potential. Stimulation of polymodal nociceptors elicits sensations of slow, burning pain.

There are two types of afferent neurons associated with nociceptors: *A-delta fibers* and *C fibers*. Thermal nociceptors and mechanical nociceptors are associated with *A-delta fibers*. These are small, myelinated fibers that transmit impulses at a rate of 5 to 30 meters per second. Polymodal nociceptors are associated with *C fibers*. These are small, unmyelinated fibers that transmit impulses at a rate generally less than 1 meter per second (range of 0.5 to 2.0 meters per second).

Table 8.1 Endogenous Chemicals Activating or Sensitizing Nociceptors

Chemical	Source	Enzyme Involved in Synthesis	Effect on First-Order Sensory Neuron	Pharmacological Intervention
Potassium	Damaged cells		Activation	
Serotonin	Platelets	Tryptophan hydroxylase	Activation	
Bradykinin	Plasma kininogen	Kallikrein	Activation	
Histamine	Mast cells		Activation	H1 receptor antagonists (e.g., diphenhydramine chloride, Benadryl)
Prostaglandins	Arachidonic acid/ damaged cells	Cyclooxygenase	Sensitization	Nonsteroidal anti-inflammatory drugs (e.g., aspirin, ibuprofen)
Leukotrienes	Arachidonic acid/ damaged cells	Lipoxygenase	Sensitization	
Substance P	First-order sensory neurons		Sensitization	Opioid receptor agonists (e.g., morphine)

8.3 Fast pain and slow pain

There are two types of pain: *fast pain* and *slow pain*. *Fast pain* may be described as sharp or prickling pain (see Table 8.2). This pain is perceived first (within 0.1 second) as it is carried by the more rapidly conducting A-delta fibers. Because fast pain is elicited by the stimulation of specific thermal or mechanical nociceptors, it is easily localized. This type of pain is not felt in most of the deeper tissues of the body. *Slow pain* may be described as dull, aching, or throbbing pain. This pain is perceived second (only after 1 second or more) as it is carried by C fibers. Slow pain persists longer and is typically more unpleasant. In fact, this pain tends to become greater over time. Slow pain is typically associated with tissue destruction. Noxious chemicals released from damaged cells or activated in the interstitial fluid can spread in the tissue causing a relatively diffuse stimulation of polymodal receptors. As a result, slow pain is poorly localized. It may occur in the skin as well as almost any deep tissue or organ.

Table 8.2 Characteristics of Fast and Slow Pain

Fast Pain	Slow Pain
Occurs first	Occurs second, persists longer
Sharp, prickling sensation	Dull, aching, throbbing sensation; more unpleasant
A-delta fibers	C fibers
Thermal or mechanical nociceptors	Polymodal nociceptors
Easily localized	Poorly localized

8.4 Hyperalgesia

An injured area is typically more sensitive to subsequent stimuli. As a result, painful stimuli, or even normally nonpainful stimuli, may cause an excessive pain response. There are two types of hyperalgesia: *primary hyperalgesia* and *centrally mediated hyperalgesia*.

An increase in the sensitivity of nociceptors is referred to as *primary hyperalgesia*. A classic example of hyperalgesia is a burn. Even light touch of a burned area may be painful.

The sensitization of nociceptors following tissue damage or inflammation results from a variety of chemicals released or activated in the injured area (see Table 8.1). These substances decrease the threshold for activation of the nociceptors. One such substance that seems to elicit more pain than the others is *bradykinin*. Activated by enzymes released from damaged cells, bradykinin causes pain by several mechanisms, including:

- Direct activation of A-delta and C fibers.
- Contribution, along with histamine, to the inflammatory response to tissue injury.
- Promotion of the synthesis and release of prostaglandins from nearby cells.

The *prostaglandins* sensitize all three types of pain receptors, which, in turn, will enhance the response to a noxious stimulus. In other words, it hurts more when prostaglandins are present. Aspirin and nonsteroidal anti-inflammatory drugs (NSAIDs) inhibit the synthesis of prostaglandins, which accounts, in part, for their analgesic effects.

Centrally mediated hyperalgesia involves the hyperexcitability of the second-order sensory neurons in the dorsal horn of the spinal cord. In the case of severe or persistent tissue injury, C fibers fire action potentials repetitively. As a result, the response of the second-order sensory neurons increases progressively. The mechanism of this enhanced response, also referred to as *wind-up,* depends on the release of the neurotransmitter *glutamate* from the C fibers. An excitatory neurotransmitter, glutamate stimulates the opening

of calcium channels gated by the *N*-methyl-D-aspartate (NMDA)-type glutamate receptor. Calcium influx ultimately leads to long-term biochemical changes and hyperexcitability of the second-order neuron.

8.5 Neurotransmitters of the nociceptive afferent fibers

There are two neurotransmitters released by the nociceptive afferent fibers in the dorsal horn of the spinal cord. These neurotransmitters, which stimulate the second-order sensory neurons, include *glutamate* and *substance P*.

The amino acid, *glutamate*, is the major neurotransmitter released by both A-delta fibers and C fibers. Glutamate binds to the AMPA-type glutamate receptor on the second-order sensory neuron to elicit action potentials and continue the transmission of the signal to higher levels of the central nervous system (CNS).

Substance P is released primarily from C fibers. Levels of this neurotransmitter increase significantly under conditions of persistent pain. Substance P also stimulates ascending pathways in the spinal cord. Furthermore, it appears to enhance and prolong the actions of glutamate.

8.6 Pain pathway

Stimulation of a nociceptor in the periphery of the body elicits action potentials in the *first-order neuron*. The signal is transmitted by this neuron to the *second-order neuron* in the dorsal horn of the spinal cord. From the spinal cord, the signal is transmitted to several regions of the brain. The more prominent ascending nociceptive pathway is the *spinothalamic tract*. Axons of the second-order sensory neurons immediately project to the contralateral (opposite) side of the spinal cord and ascend in the white matter, terminating in the *thalamus* (see Figure 8.1). The thalamus contributes to the basic sensation or awareness of pain only. The source of the painful stimulus cannot be determined by the thalamus.

The cell bodies of *third-order sensory* neurons are located in the thalamus. These neurons transmit the pain signal to the *somatosensory cortex*. The function of this region of the brain is to localize and perceive the intensity of the painful stimulus. Further transmission of the signal to the *association areas* of the cerebral cortex is important for the perception and the meaningfulness of the painful stimulus.

Other nerve signals are transmitted simultaneously from the spinal cord to the *reticular formation* of the brainstem by way of the *spinoreticular tract*. The reticular formation plays an important role in the response to pain. First, it facilitates avoidance reflexes at all levels of the spinal cord. Second, it is responsible for the significant arousal effects of pain. Signals from the reticular formation cause an increase in the electrical activity of the cerebral cortex associated with increased alertness. Furthermore, it sends nerve impulses to the *hypothalamus* to influence its functions associated with sudden alertness,

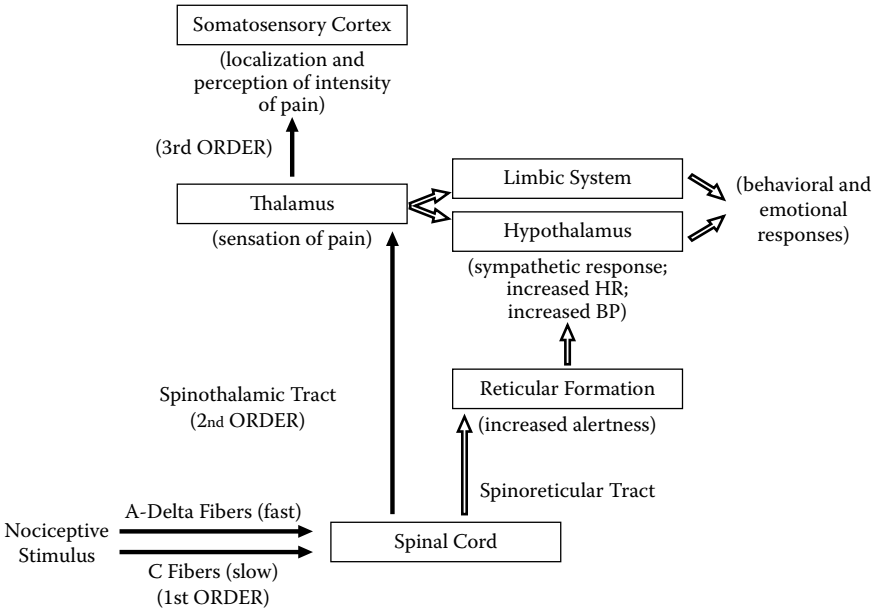


Figure 8.1 The pain pathway. The pain signal is transmitted to several regions of the brain including the thalamus, the reticular formation, the hypothalamus, the limbic system, and the somatosensory cortex. Each region carries out a specific aspect of the response to pain.

such as increased heart rate and blood pressure. These responses are mediated by the sympathetic nervous system.

Nerve signals from the thalamus and the reticular formation are transmitted to the *limbic system* as well as the hypothalamus. Together, these regions of the brain are responsible for the behavioral and emotional responses to pain. The limbic system, in particular, may be involved with the mood-altering and attention-narrowing effect of pain.

8.7 Endogenous analgesic system

The *endogenous analgesic system* is a built-in neuronal system that suppresses the transmission of nerve impulses in the pain pathway. It functions by way of the following neurotransmitters produced in the CNS:

- Endorphins
- Enkephalins
- Dynorphin

Endorphins are found primarily in the limbic system, the hypothalamus, and the brainstem. *Enkephalins* and *dynorphin* (in smaller quantities) are

found primarily in the *periaqueductal gray matter* (PAG) of the midbrain, the limbic system, and the hypothalamus. These endogenous substances mimic the effects of morphine and other opiate drugs (narcotics) at many points in the analgesic system, including in the dorsal horns of the spinal cord.

Opiate receptors are highly concentrated in the PAG area of the midbrain. Stimulation of this region produces long-lasting analgesia with no effect on the level of consciousness. For these reasons, the PAG area is often referred to as the *endogenous analgesia center*. The PAG area receives input from many regions of the CNS including the cerebral cortex, hypothalamus, reticular formation of the brainstem, and the spinal cord by way of the spinothalamic tracts. This region is also interconnected with the limbic system, which is responsible for the emotional response to pain.

There are three major components of the endogenous analgesic pathway:

1. Periaqueductal gray area
2. Nucleus raphe magnus
3. Pain inhibitory complex in the dorsal horns of the spinal cord

The endogenous analgesic pathway begins in the PAG area. Neurons of the PAG area descend to the *nucleus raphe magnus* (NRM) in the medulla (see Figure 8.2). Neurons of the NRM then descend to the dorsal horn of the spinal cord where they synapse with local spinal interneurons. The interneurons then synapse with the incoming pain fibers. Many of the neurons derived from the PAG area secrete enkephalin from their axon terminals in the NRM. The neurons derived from the NRM secrete serotonin from their axon terminals in the spinal cord. The serotonin stimulates the local cord interneurons to secrete enkephalin. The enkephalin then causes presynaptic inhibition of the incoming pain fibers. The binding of enkephalin to opiate receptors on these pain fibers blocks the calcium channels in the axon terminals. Because the influx of calcium is necessary for the exocytosis of the neurotransmitter, blocking these channels prevents the release of substance P. As a result, this system interrupts the pain signal at the level of the spinal cord.

The endogenous analgesic system is usually inactive. It remains unclear as to how this system becomes activated. Potential activating factors include exercise, stress, acupuncture, and hypnosis.

8.8 Gate control theory

The transmission of nerve impulses in the pain pathway may also be inhibited by the activity of *A-beta neurons* by way of a mechanism referred to as the *gate control theory*. Under normal conditions, activity in the second-order neurons of the pain pathway is blocked by tonically active *inhibitory interneurons* in the dorsal horn of the spinal cord (Figure 8.3a). When a painful stimulus

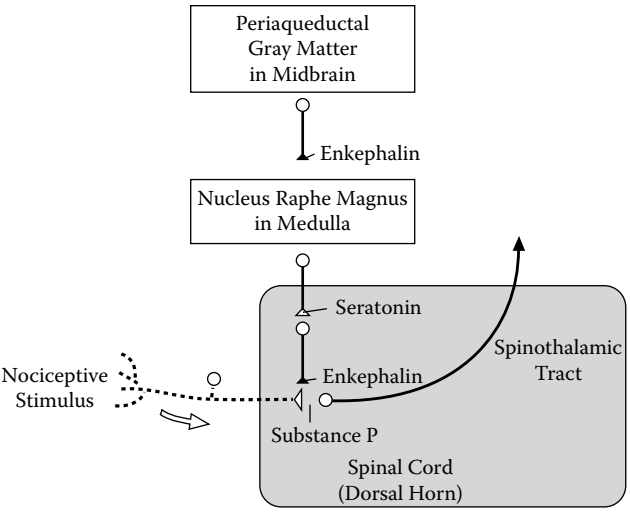


Figure 8.2 The endogenous analgesic system. The three major components of the endogenous analgesic system include the periaqueductal gray matter in the midbrain, the nucleus raphe magnus in the medulla, and the pain inhibitory complex in the dorsal horns of the spinal cord. This system causes presynaptic inhibition of the pain fibers entering the spinal cord. The binding of enkephalin to opiate receptors on the pain fibers prevents the release of the neurotransmitter, substance P. As a result, the pain signal is terminated in the spinal cord and does not ascend to higher centers in the CNS.

results in the transmission of action potentials to the spinal cord along the C fibers, the activity of the inhibitory interneuron ceases. By way of divergence, C fibers directly inhibit the inhibitory interneurons. Therefore, the C fibers not only stimulate the second-order neurons, they also prevent the inhibitory effect exerted upon those neurons, and a strong pain signal is transmitted to the brain (Figure 8.3b).

A-beta fibers are large-diameter fibers that transmit impulses at a rate of 30 to 70 m/sec. Like the C fibers, they are activated by mechanical stimuli, such as touch. A-beta fibers also synapse with the inhibitory interneurons in the dorsal horn of the spinal cord. However, in contrast to the C fibers that block activity in the inhibitory interneurons, A-beta fibers stimulate these interneurons (Figure 8.3c). As a result, the inhibitory effect that these neurons exert upon the second-order neurons in the pain pathway is renewed. Therefore, the inhibitory interneurons modulate, or reduce, the activation of the second-order neurons, and the pain signal transmitted to the brain becomes weaker.

The gate control theory explains the phenomenon that occurs by rubbing an injured body part. The rubbing activates the A-beta fibers and the perception of pain is reduced.

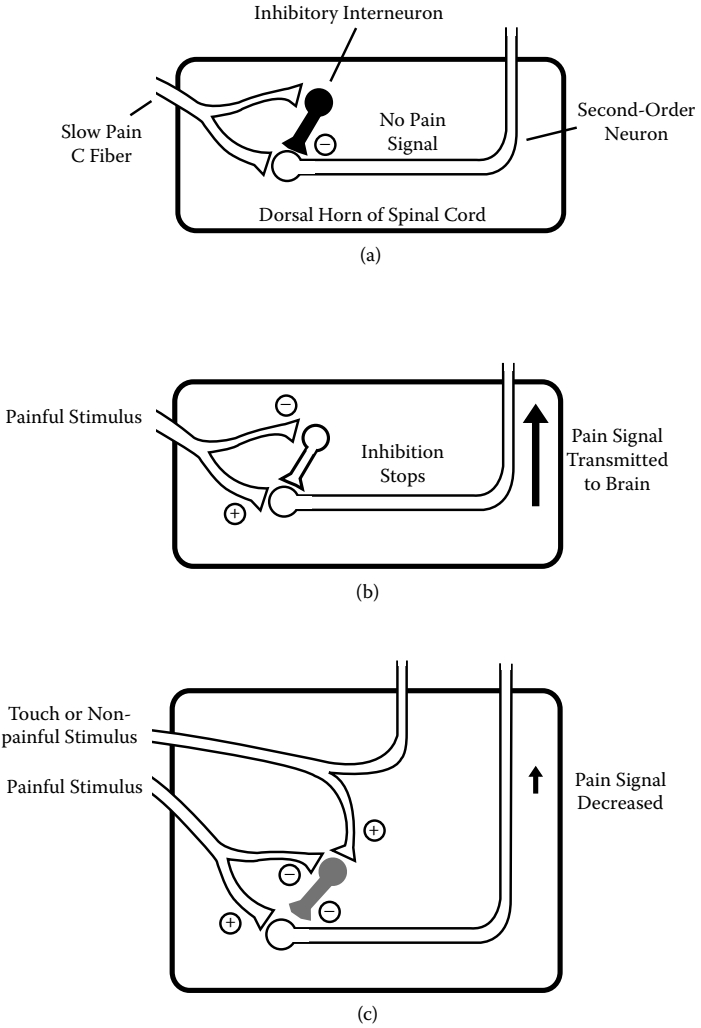


Figure 8.3 The gate control theory. The mechanism proposed by this theory involves the inhibition of the second-order neuron in the pain pathway. Cell bodies of these neurons are located in the dorsal horn of the spinal cord. (a) Under normal conditions, tonically active inhibitory interneurons in the dorsal horn of the spinal cord block the stimulation of the second-order neurons. This activity prevents the perception of pain. (b) When a strong pain stimulus reaches the spinal cord by way of the C fibers, the second-order neurons are stimulated. In addition, the C fibers block the activity of the inhibitory interneuron. As a result, the pain stimulus continues along the pain pathway to the brain. (c) The perception of pain may be reduced by simultaneous somatosensory input. The sensation of touch is transmitted by A-beta fibers. These fibers stimulate the inhibitory interneurons, and activity of the second-order neurons is reduced. The signals to the brain are decreased, and the perception of pain is decreased.

8.9 Cutaneous pain

Cutaneous pain is felt in superficial structures, such as the skin and subcutaneous tissues. A pinprick or a paper cut are examples of cutaneous pain. It is a sharp pain with a burning quality that may be easily localized. This pain may be abrupt or slow in onset.

8.10 Deep somatic pain

As its name implies, *deep somatic pain* is generated in deep body structures, such as the periosteum, muscles, tendons, joints, and blood vessels. This type of pain is more diffuse than cutaneous pain. It may be elicited by strong pressure, ischemia, and tissue damage.

8.10.1 Tissue ischemia

When blood flow to a tissue is decreased or interrupted, the tissue becomes painful within a few minutes. In fact, the greater the rate of metabolism in the tissue, the more rapid is the onset of pain. The causes of pain due to *tissue ischemia* include the following:

- Accumulation of *lactic acid* due to the anaerobic metabolism that occurs during ischemia.
- *Release and activation of noxious chemicals* in the area of tissue ischemia due to tissue damage (see Table 8.1).

The lactic acid and other noxious chemicals stimulate polymodal nociceptors.

8.10.2 Muscle spasm

The pain induced by *muscle spasm* results partially from the direct effect of tissue distortion on mechanical nociceptors. Muscle spasm also causes tissue ischemia. The increased muscle tension compresses blood vessels and decreases blood flow. Furthermore, the increased rate of metabolism associated with the spasm exacerbates the ischemia. As discussed above, ischemia leads to the stimulation of polymodal nociceptors.

8.11 Visceral pain

Visceral pain occurs in the organs and tissues of the thoracic and abdominal cavities. This type of pain has several interesting characteristics, including the following:

- Not all tissues evoke visceral pain (e.g., liver, lung parenchyma).
- It is diffuse and poorly localized.

- It often generates referred pain (discussed in Section 8.12).
- It is typically accompanied by autonomic nervous system responses (e.g., nausea, vomiting, sweating, pallor, increased blood pressure).

Visceral pain may be caused by several factors:

- Inflammation
- Chemical stimuli
- Spasm of a hollow organ
- Overdistension of a hollow organ

Inflammation of the appendix (appendicitis) and the gallbladder (cholecystitis) are common examples of visceral pain. Mechanical receptors are activated by the tissue distension associated with inflammation. In addition, inflammatory mediators, such as histamine and bradykinin, may activate polymodal nociceptors.

Chemical stimuli may include gastric acid (gastroesophageal reflux disease [GERD], gastric ulcer, duodenal ulcer) or those substances associated with tissue ischemia, tissue damage, and inflammation.

Spasm of the smooth muscle in the wall of a hollow organ causes pain due to the direct stimulation of mechanical nociceptors as well as ischemia-induced stimulation of polymodal nociceptors. This type of pain often occurs in the form of *cramps*. In other words, the pain increases to a high intensity and then subsides. This process occurs rhythmically, once every few minutes. Cramping pain frequently occurs in gastroenteritis, menstruation, and parturition (labor).

Overdistension of a hollow organ causes pain by the excessive stretch of the tissue and the stimulation of mechanical nociceptors. Overdistension may also cause collapse of the blood vessels resulting in the development of ischemic pain.

8.12 Referred pain

Referred pain is felt in a part of the body that is different from the actual tissue causing the pain. Typically, the pain is initiated in a visceral organ or tissue and referred to an area of the body surface. Classic examples of referred pain include *headache* and *angina*. Interestingly, the brain does not contain nociceptors. Therefore, the pain perceived as a headache originates in other tissues, such as the eyes, sinuses, muscles of the head and neck, and the meninges. Angina, or chest pain, is caused by coronary ischemia. It may be accompanied by pain referred to the neck, left shoulder, and left arm.

Referred pain most likely results from the convergence of visceral and somatic afferent fibers on the same second-order neurons in the dorsal horn of the spinal cord (see Figure 8.4). Therefore, the brain has no way of identifying the original source of the pain. Because superficial inputs normally

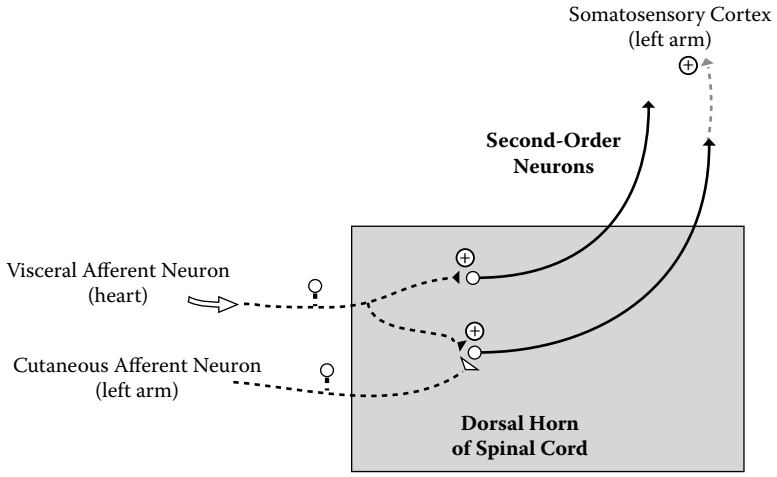


Figure 8.4 Referred pain. The mechanism of referred pain involves the convergence of visceral afferent neurons and cutaneous afferent neurons with the same second-order neurons in the dorsal horn of the spinal cord. In this example, the pain of angina that originates in the heart is referred to the left arm.

predominate over visceral inputs, higher centers may incorrectly attribute the pain to the skin instead of the deeper tissue.

8.13 Phantom pain

Phantom pain is pain that appears to arise from an amputated limb or body part. As many as 70% of amputees experience phantom pain. This pain may begin with sensations of tingling, heat, cold, or heaviness, followed by burning, cramping, or shooting pain. Phantom pain may disappear spontaneously or persist for many years.

The exact cause of phantom pain is not clearly understood. One proposed mechanism involves stimulation of the sensory pathway that had once originated in the amputated body part. An important point is that the sensory pathway originating in a given body part transmits impulses to the region of the somatosensory cortex devoted to that body part regardless of amputation. Stimulation at any point along this pathway results in the same sensation that would be produced by stimulation of the nociceptor in the body part. Following amputation of a body part, the ends of the first-order afferent nerves arising from that body part become trapped in the scar tissue of the stump. These afferent nerve endings exhibit increased sensitivity and are easily stimulated. Therefore, action potentials are generated at these nerve endings and are transmitted to the area of the somatosensory cortex devoted to the amputated body part. This results in the perception of pain arising from the amputated portion of the body.

A second theory of phantom pain suggests that the second-order neurons in the dorsal horn of the spinal cord become hyperactive. Spontaneous firing of these neurons causes the transmission of nerve impulses to the brain and the perception of pain.

8.14 Pharmacologic treatment of pain

An *analgesic drug* acts on the nervous system to suppress or eliminate pain without causing loss of consciousness. As such, an ideal analgesic would exhibit the following qualities:

- Potent
- Nonaddictive
- Minimal adverse effects
- Effective without altering the patient's state of awareness
- Does not cause tolerance
- Inexpensive

Pain medications may be divided into three categories:

1. Nonnarcotic analgesics
2. Opioid analgesics
3. Adjuvant analgesics

8.14.1 Nonnarcotic analgesics

The *nonnarcotic analgesics* include aspirin, NSAIDs, and acetaminophen. *Aspirin* acts peripherally to block the transmission of pain impulses. Furthermore, it reduces fever (antipyretic effects) and inflammation and inhibits the synthesis of the prostaglandins, which increase the sensitivity of the nociceptors.

The *NSAIDs* exert their analgesic effects primarily through the inhibition of cyclooxygenase, the rate-limiting enzyme for prostaglandin synthesis. Typical nonselective NSAIDs inhibit both cyclooxygenase 1 (COX-1; constitutive) and cyclooxygenase 2 (COX-2; induced in areas of inflammation). More recently, medications specific for COX-2, such as celecoxib (Celebrex®) have been developed. (Other medications, such as rofecoxib [Vioxx®] and valdecoxib [Bextra®] have been withdrawn from use due to adverse cardiovascular events reported in many patients.) The advantage of the COX-2 agents is that they reduce pain, fever, and inflammation without a high risk of unwanted gastric side effects that accompany COX-1 inhibition, particularly those leading to gastrointestinal irritation and gastric ulcers. (The protective effects of the COX-1-produced prostaglandins in the stomach are discussed in Chapter 20.)

Acetaminophen, another alternative to aspirin, is an effective analgesic and fever-reducing agent. However, at usually administered doses, this medication has no effect on inflammation.

8.14.2 Opioid analgesics

Medications with morphine-like actions are referred to as *opioid* or *narcotic agents*. Opioid drugs exert their effects through three major categories of opioid receptors: mu (μ), kappa (κ), and delta (Δ). Analgesia appears to involve μ receptors (largely at supraspinal sites) and κ receptors (principally within the spinal cord). *Morphine* produces analgesia through interaction with μ receptors. In fact, most clinically used opioids are relatively selective for μ receptors. Morphine can stimulate μ_2 receptors spinally or μ_1 receptors supraspinally. When given systemically, it acts predominantly through supraspinal μ_1 receptors. Other effects of μ receptor activation include respiratory depression, reduced gastrointestinal motility (leading to constipation), and feelings of well-being or euphoria.

Morphine may be administered orally, intravenously, or epidurally. An advantage of epidural administration is that it provides effective analgesia while minimizing the central depressant effects associated with systemic administration. The mechanism of action with the epidural route of administration involves the opiate receptors on the cell bodies of the first-order sensory neurons in the dorsal root ganglia as well as their axon terminals in the dorsal horn. Stimulation of these receptors inhibits the release of substance P and interrupts the transmission of the pain signal to the second-order sensory neuron.

Although opioid analgesics are very effective in relieving pain, they are also highly addictive. Furthermore, when needed for long-term use, there may be the development of tolerance. When this occurs, larger and larger doses of the narcotic agent are needed to elicit the same degree of pain relief.

8.14.3 Adjuvant analgesics

Adjuvant analgesics include medications such as *antidepressants* and *antiseizure medications*. The effectiveness of these agents may be due to the existence of nonendorphin synapses in the endogenous analgesic pathway. For example, the neurotransmitter serotonin has been shown to play a role in producing analgesia. Tricyclic antidepressant medications, such as imipramine, that block the removal of serotonin from the synapse, suppress pain in some individuals. Certain antiseizure medications, such as carbamazepine and phenytoin, have specific analgesic effects that are effective under certain conditions. For example, these medications, which suppress spontaneous neuronal firing, are particularly effective in the management of pain following nerve injury. Other agents, such as the *corticosteroids*, reduce pain by decreasing inflammation and the nociceptive stimuli responsible for the pain.

Medical terminology

Analgesia (ăn-ăl-jē′zē-ă): No perception of pain.
Analgesic (ăn-ăl-jē′sĭk): Medication that relieves pain.
Hyperalgesia (hī″pĕr-ăl-jē′zē-ă): Increased sensitivity to painful stimuli.
Ischemia (ĭs-kē′mē-ă): Temporary blood flow deficiency to an organ or tissue.
Narcotic (năr-kŏt′ĭk): Any drug derived from opium or opium-like compounds; analgesic drug that depresses the central nervous system.
Nociceptor (nō″sē-sĕp′tor): A free nerve ending responsive to painful stimuli.
Opioid (ō′pē-oyd): A synthetic narcotic not derived from opium which binds to and stimulates an opiate receptor (e.g., morphine).
Spasm (spăzm): An involuntary muscle contraction.

Bibliography

Basbaum, A. I., and Jessell, T. M., The Perception of Pain, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 24.

Bear, M. F., Connors, B. W., and Paradiso, M. A., *Neuroscience, Exploring the Brain,* 3rd ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2007.

Burke, A., Smyth, E. M., and Fitzgerald, G. A., Analgesic-Antipyretic and Antiinflammatory Agents; Pharmacotherapy of Gout, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 26.

Devine, E., Somatosensory Function, Pain, and Headache, in *Pathophysiology, Concepts of Altered Health States,* 7th ed., Porth, C. M., Ed., Lippincott-Raven, Philadelphia, PA, 2005, chap. 50.

Gardner, E. P., Martin, J. H., and Jessell, T. M., The Bodily Senses, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 22.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Rhoades, R., and Pflanzer, R., *Human Physiology*, 4th ed., Brooks/Cole, Pacific Grove, CA, 2003.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

chapter nine

The autonomic nervous system

Study objectives

- Explain how various regions of the central nervous system regulate autonomic nervous system function
- Explain how autonomic reflexes contribute to homeostasis
- Describe how the neuroeffector junction in the autonomic nervous system differs from that of a neuron-to-neuron synapse
- Compare and contrast the anatomical features of the sympathetic and parasympathetic systems
- For each neurotransmitter in the autonomic nervous system, list the neurons that release them and the type and location of receptors that bind with them
- Describe the mechanisms by which neurotransmitters are removed
- Distinguish between cholinergic and adrenergic receptors
- Describe the overall and specific functions of the sympathetic system
- Describe the overall and specific functions of the parasympathetic system
- Explain how the effects of the catecholamines differ from those of direct sympathetic stimulation

9.1 Introduction

The *autonomic nervous system* (ANS) is also known as the visceral or involuntary nervous system. It functions without conscious, voluntary control. Because it innervates cardiac muscle, smooth muscle, and various endocrine and exocrine glands, this nervous system influences the activity of most of the organ systems in the body. Therefore, it is evident that the ANS makes a significant contribution to homeostasis. The regulation of blood pressure, gastrointestinal responses to food, contraction of the urinary bladder, focusing of the eyes, and thermoregulation are just a few of the many homeostatic functions regulated by the ANS. Several distinguishing features of the ANS and the somatic nervous system, which innervates skeletal muscle, are summarized in Table 9.1.

Table 9.1 Distinguishing Features of the Somatic and Autonomic Nervous Systems

Somatic Nervous System	Autonomic Nervous System
Conscious control, voluntary	Unconscious control, involuntary
Skeletal muscle	All innervated structures except skeletal muscle (e.g., cardiac muscle, smooth muscle, glands)
Movement, respiration, posture	Visceral functions (e.g., cardiac activity, blood flow, digestion, etc.)
No peripheral ganglia; synapses located entirely within cerebrospinal axis	Peripheral ganglia located outside of cerebrospinal axis
Alpha motor neuron	Preganglionic neuron and postganglionic neuron
Myelinated, large diameter (9–13 μm)	Nonmyelinated, small diameter (~1 μm)
Neurotransmitter: acetylcholine only	Neurotransmitters: acetylcholine and norepinephrine
Cell bodies in ventral horn of spinal cord	Cell bodies in brainstem, lateral horn of spinal cord
Axon divides; each axon terminal innervates a single muscle fiber directly	No discreet innervation of individual effector cells
Motor end-plate or neuromuscular junction = axon terminal in close apposition to a specialized surface of the muscle cell membrane	Axon terminal with multiple varicosities releases neurotransmitter over a wide surface area affecting many tissue cells
No gap junctions between effector cells; no spread of electrical activity from one muscle fiber to another	Gap junctions allow the spread of electrical activity throughout the tissue

9.2 Regulation of autonomic nervous system activity

The efferent nervous activity of the ANS is regulated by several regions in the central nervous system (CNS):

- Hypothalamus and brainstem
- Cerebral cortex and limbic system
- Spinal cord

The efferent nervous activity of the ANS is largely regulated by autonomic reflexes. In many of these reflexes, sensory information is transmitted to homeostatic control centers, in particular, those located in the hypothalamus and the brainstem. Much of the sensory input from the thoracic and abdominal viscera is transmitted to the brainstem by afferent fibers of cranial

nerve X, the vagus nerve. Other cranial nerves also contribute sensory input to the hypothalamus and the brainstem. This input is integrated and a response is carried out by the transmission of nerve signals that modify the activity of preganglionic autonomic neurons. Many important variables in the body are monitored and regulated in the hypothalamus and the brainstem including heart rate, blood pressure, gastrointestinal peristalsis and glandular secretion, body temperature, hunger, thirst, plasma volume, and plasma osmolarity.

An example of this type of autonomic reflex is the baroreceptor reflex. Baroreceptors, located in some of the major systemic arteries, are sensory receptors that monitor blood pressure. If blood pressure decreases, the number of sensory impulses transmitted from the baroreceptors to the vasomotor center in the brainstem also decreases. As a result of this change in baroreceptor stimulation and sensory input to the brainstem, ANS activity to the heart and blood vessels is adjusted to increase heart rate and vascular resistance so that blood pressure increases and returns to its normal value.

These neural control centers in the hypothalamus and the brainstem may also be influenced by higher brain areas. Specifically, the cerebral cortex and the limbic system influence ANS activities associated with emotional responses by way of hypothalamic-brainstem pathways. For example, blushing during an embarrassing moment, a response most likely originating in the frontal association cortex, involves vasodilation of blood vessels to the face. Other emotional responses influenced by these higher brain areas include fainting, breaking out in a cold sweat, and a racing heart rate.

Some autonomic reflexes may be processed at the level of the spinal cord. These include the micturition reflex (urination) and the defecation reflex. Although these reflexes are subject to influence from higher nervous centers, they may occur without input from the brain.

9.3 *Efferent pathways of the autonomic nervous system*

The efferent pathways of the ANS consist of two neurons that transmit impulses from the CNS to the effector tissue. The *preganglionic neuron* originates in the CNS with its cell body in the lateral horn of the gray matter of the spinal cord or in the brainstem. The axon of this neuron travels to an autonomic ganglion located outside of the CNS where it synapses with a *postganglionic neuron*. It is this neuron that innervates the effector tissue. (A *ganglion* is a cluster of nerve cell bodies in the peripheral nervous system. A *nucleus* is a cluster of nerve cell bodies in the central nervous system.)

Synapses between the autonomic postganglionic neuron and the effector tissue, the *neuroeffector junction*, differ greatly from the neuron-to-neuron synapses discussed previously in Chapter 5 (see Table 9.1). The postganglionic fibers in the ANS do not terminate in a single swelling like the synaptic knob, and they do not synapse directly with the cells of a tissue. Instead, where the axons of these fibers enter a given tissue, they contain multiple

swellings called *varicosities*. When the neuron is stimulated, these varicosities release neurotransmitters along a significant length of the axon and, therefore, over a large surface area of the effector tissue. The neurotransmitter diffuses through the interstitial fluid to where its receptors are located in the tissue. This diffuse release of the neurotransmitter affects many tissue cells simultaneously. Furthermore, cardiac muscle and most smooth muscle have *gap junctions* between the cells. These specialized intercellular communications allow for the spread of electrical activity from one cell to the next. As a result, the discharge of a single autonomic nerve fiber to an effector tissue may alter the activity of the entire tissue.

9.4 Divisions of the autonomic nervous system

The ANS is composed of two anatomically and functionally distinct divisions, the *sympathetic system* and the *parasympathetic system*. Two important features of these divisions include *tonic activity* and *dual innervation*.

Both systems are *tonically active*. In other words, they provide some degree of nervous input to a given tissue at all times. Therefore, the frequency of discharge of neurons in both systems can either increase or decrease. As a result, tissue activity may be either enhanced or inhibited. This characteristic of the ANS improves its ability to more precisely regulate a tissue's function. Without tonic activity, nervous input to a tissue could only increase.

Many tissues are *innervated by both systems*. Because the sympathetic system and the parasympathetic system typically have opposing effects on a given tissue, increasing the activity of one system while simultaneously decreasing the activity of the other results in very rapid and precise control of a tissue's function. For example, sympathetic nervous stimulation increases heart rate, and parasympathetic nervous stimulation decreases heart rate. Under conditions of exercise where it is beneficial to increase heart rate and pump more blood to the skeletal muscles, there is a simultaneous increase in sympathetic activity and decrease in parasympathetic activity to the heart. Both of these effects serve to increase heart rate. Therefore, the activity of a given tissue is the result of the balance of the input from these two opposing systems. Several distinguishing features of these two divisions of the ANS are summarized in Table 9.2.

Each system is dominant under certain conditions. The sympathetic system predominates during emergency *fight-or-flight* reactions and during exercise. The overall effect of the sympathetic system under these conditions is to prepare the body for strenuous physical activity. More specifically, sympathetic nervous activity will increase the flow of blood that is well-oxygenated and rich in nutrients to the tissues that need it, in particular, the working skeletal muscles. The parasympathetic system predominates during quiet, resting conditions. The overall effect of the parasympathetic system under these conditions is to conserve and store energy and to regulate basic body functions such as digestion and urination (*rest and digest*).

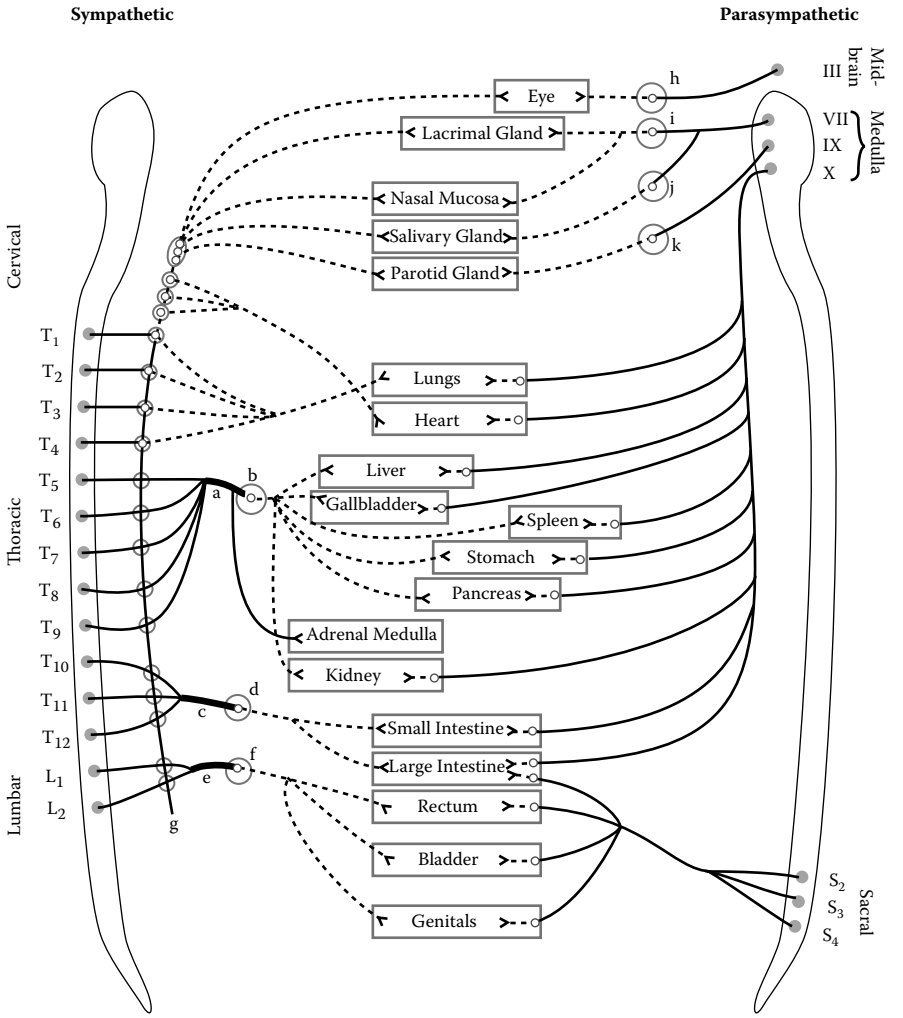
Table 9.2 Distinguishing Features of the Sympathetic and Parasympathetic Systems

Sympathetic System	Parasympathetic System
Originates in the thoracic and lumbar regions of the spinal cord (T_1–L_2)	Originates in the brainstem (cranial nerves III, VII, IX, and X) and the sacral region of the spinal cord (S_2–S_4)
Ganglia located in paravertebral sympathetic ganglion chain or in collateral ganglia	Terminal ganglia located near or embedded within target tissue
Short cholinergic preganglionic fibers Long adrenergic postganglionic fibers	Long cholinergic preganglionic fibers Short cholinergic postganglionic fibers
Ratio of preganglionic fibers to postganglionic fibers is 1:20	Ratio of preganglionic fibers to postganglionic fibers is 1:3
Divergence coordinates activity of neurons at multiple levels of the spinal cord	Limited divergence
Activity often involves mass discharge of entire system	Activity normally to discrete organs
Predominates during emergency "fight-or-flight" reactions and exercise	Predominates during quiet "rest and digest" conditions

9.5 Sympathetic division

The preganglionic neurons of the sympathetic system arise from the thoracic and lumbar regions of the spinal cord (segments T_1 through L_2) (see Figure 9.1). Most of these preganglionic axons are short and synapse with postganglionic neurons within ganglia found in the *sympathetic ganglion chains*. These ganglion chains, which run parallel immediately along either side of the spinal cord, each consist of 22 ganglia. The preganglionic neuron may exit the spinal cord and synapse with a postganglionic neuron in a ganglion at the same spinal cord level from which it arises. The preganglionic neuron may also travel more rostrally or caudally (upward or downward) in the ganglion chain to synapse with postganglionic neurons in ganglia at other levels. In fact, a single preganglionic neuron may synapse with several postganglionic neurons in many different ganglia. Overall, the ratio of preganglionic fibers to postganglionic fibers is about 1:20. The long postganglionic neurons originating in the ganglion chain then travel outward and terminate on the effector tissues. This divergence of the preganglionic neuron results in coordinated sympathetic stimulation to tissues throughout the body. The concurrent stimulation of many organs and tissues in the body is referred to as a *mass sympathetic discharge*.

Other preganglionic neurons exit the spinal cord and pass through the ganglion chain without synapsing with a postganglionic neuron. Instead, the axons of these neurons travel more peripherally and synapse with



a: Greater splanchnic nerve
b: Celiac ganglion
c: Lesser splanchnic nerve
d: Superior mesenteric ganglion
e: Lumbar splanchnic nerve
f: Inferior mesenteric ganglion
g: sympthetic ganglion chain
h: Ciliary ganglion
i: Pterygopalatine ganglion
j: Submandibular ganglion
k: Otic ganglion
Postganglionic nerve

Figure 9.1 The autonomic nervous system and its effector organs. The efferent pathways of this system consist of two neurons that transmit impulses from the central nervous system to the effector tissue, the preganglionic neuron (solid line), and the postganglionic neuron (dashed line). As illustrated, most tissues receive nervous input from both divisions of the autonomic nervous system — the sympathetic division and the parasympathetic division.

postganglionic neurons in one of the *sympathetic collateral ganglia* (see Figure 9.1). These ganglia are located about halfway between the CNS and the effector tissue.

Finally, the preganglionic neuron may travel to the *adrenal medulla* and synapse directly with this glandular tissue. The cells of the adrenal medulla have the same embryonic origin as neural tissue and, in fact, function as *modified postganglionic neurons*. Instead of the release of neurotransmitter directly at the synapse with an effector tissue, the secretory products of the adrenal medulla are picked up by the blood and travel throughout the body to all of the effector tissues of the sympathetic system.

An important feature of this system, which is quite distinct from the parasympathetic system, is that the postganglionic neurons of the sympathetic system travel within each of the 31 pairs of spinal nerves (see Chapter 7). Interestingly, 8% of the fibers that constitute a spinal nerve are sympathetic fibers. This allows for the distribution of sympathetic nerve fibers to the effectors of the skin, including blood vessels and sweat glands. In fact, most innervated blood vessels in the entire body, primarily arterioles and veins, receive only sympathetic nerve fibers. Therefore, vascular smooth muscle tone and sweating are regulated by the sympathetic system only. In addition, the sympathetic system innervates structures of the head (eye, salivary glands, mucus membranes of the nasal cavity), thoracic viscera (heart, lungs), and viscera of the abdominal and pelvic cavities (e.g., stomach, intestines, pancreas, spleen, adrenal medulla, kidneys, urinary bladder) (see Figure 9.1).

9.6 *Parasympathetic division*

The preganglionic neurons of the parasympathetic system arise from several nuclei of the brainstem and from the sacral region of the spinal cord (segments S_2 through S_4) (see Figure 9.1). The axons of the preganglionic neurons are quite long compared to those of the sympathetic system and synapse with postganglionic neurons within *terminal ganglia* that are close to or embedded within the effector tissues. The axons of the postganglionic neurons, which are very short, then provide input to the cells of that effector tissue.

The preganglionic neurons that arise from the brainstem exit the CNS through the cranial nerves. The oculomotor nerve (III) innervates the eyes; the facial nerve (VII) innervates the lacrimal glands, the salivary glands, and the mucus membranes of the nasal cavity; the glossopharyngeal nerve (IX) innervates the parotid (salivary) glands; and the vagus nerve (X) innervates the viscera of the thorax and the abdomen (e.g., heart, lungs, stomach, pancreas, small intestine, upper half of the large intestine, and liver). The physiological significance of cranial nerve X in terms of the influence of the parasympathetic system is clearly illustrated by its widespread distribution and the fact that 75% of all parasympathetic fibers are in the vagus nerve.

The preganglionic neurons that arise from the sacral region of the spinal cord exit the CNS and join together to form the pelvic nerves. These nerves innervate the viscera of the pelvic cavity (for example, lower half of the large intestine and organs of the renal and reproductive systems).

Because the terminal ganglia are located within the innervated tissue, there is typically little divergence in the parasympathetic system compared to the sympathetic system. In many organs, there is a 1:1 ratio of preganglionic fibers to postganglionic fibers. Therefore, the effects of the parasympathetic system tend to be more discrete and localized, with only specific tissues being stimulated at any given moment, compared to the sympathetic system where a more diffuse discharge is possible.

9.7 Neurotransmitters of the autonomic nervous system

The two most common neurotransmitters released by neurons of the ANS are *acetylcholine* (Ach), and *norepinephrine* (NE). Neurotransmitters are synthesized in the axon varicosities and stored in vesicles for subsequent release. Several distinguishing features of these neurotransmitters are summarized in Table 9.3. Nerve fibers that release acetylcholine are referred to as *cholinergic* fibers. These include all preganglionic fibers of the ANS, both sympathetic and parasympathetic systems; all postganglionic fibers of the parasympathetic system; and some sympathetic postganglionic fibers innervating sweat glands (see Figure 9.2). Nerve fibers that release norepinephrine are referred to as *adrenergic* fibers. All other sympathetic postganglionic fibers release norepinephrine.

As previously mentioned, the cells of the adrenal medulla are considered modified sympathetic postganglionic neurons. Instead of a neurotransmitter, these cells release *hormones* into the blood. Approximately 20% of the hormonal output of the adrenal medulla is norepinephrine. The remaining 80% is *epinephrine* (EPI). Unlike true postganglionic neurons in the sympathetic system, the adrenal medulla contains an enzyme that methylates norepinephrine to form epinephrine. The synthesis of epinephrine, also known as *adrenalin*, is enhanced under conditions of stress. These two hormones released by the adrenal medulla are collectively referred to as the *catecholamines*.

9.8 Termination of neurotransmitter activity

For any substance to serve effectively as a neurotransmitter, it must be rapidly removed or inactivated from the synapse or, in this case, the neuroeffector junction. This is necessary in order to allow new signals to get through and influence effector tissue function. Neurotransmitter activity may be terminated by three mechanisms:

Table 9.3 Distinguishing Features of the Neurotransmitters of the Autonomic Nervous System

Feature	Acetylcholine	Norepinephrine	Epinephrine[a]
Site of Release	All preganglionic neurons of the autonomic nervous system; all postganglionic neurons of the parasympathetic system; some sympathetic postganglionic neurons innervating sweat glands (alpha motor neurons innervating skeletal muscle)[b]	Most sympathetic postganglionic neurons; adrenal medulla (20% of secretion)	Adrenal medulla (80% of secretion)
Receptor	Nicotinic, muscarinic (cholinergic)	α_1, α_2, β_1 (adrenergic)	α_1, α_2, β_1, β_2 (adrenergic)
Termination of Activity	Enzymatic degradation by acetylcholinesterase	Reuptake into nerve terminals; diffusion out of synaptic cleft and uptake at extraneuronal sites; metabolic transformation by monoamine oxidase (within nerve terminal) or catechol-O-methyl-transferase within the liver	Metabolic transformation by catechol-O-methyltransferase within the liver

[a] Although epinephrine is not a direct neurotransmitter for the autonomic nervous system, its release from the adrenal medulla supplements the effects of a mass sympathetic discharge.
[b] Alpha motor neurons, a component of the somatic nervous system, also release acetylcholine as a neurotransmitter.

1. Enzymatic degradation
2. Reuptake into the neuron
3. Diffusion away from the synapse

The primary mechanism used by cholinergic synapses is *enzymatic degradation*. *Acetylcholinesterase* hydrolyzes acetylcholine to its component choline and acetate. It is one of the fastest-acting enzymes in the body, and

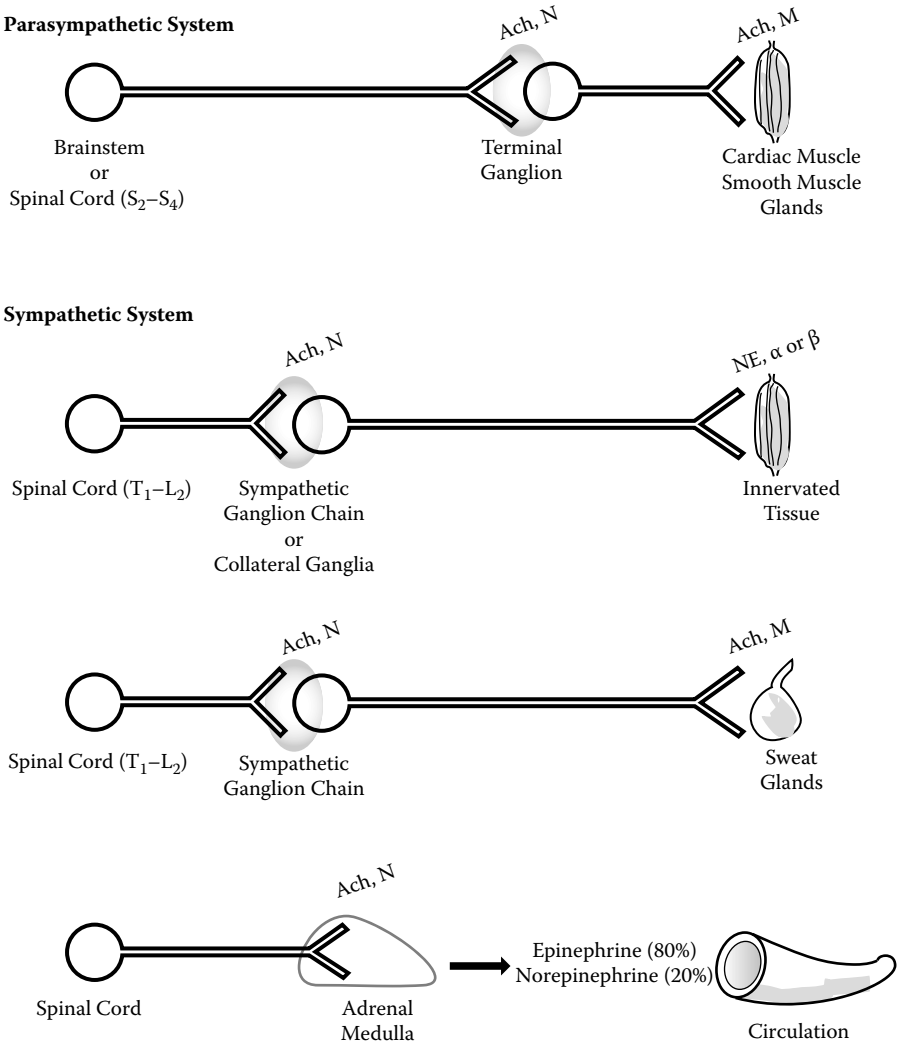


Figure 9.2 Autonomic nerve pathways. All preganglionic neurons release acetylcholine (Ach) that binds to nicotinic receptors (N) on the postganglionic neurons. All postganglionic neurons in the parasympathetic system and some sympathetic postganglionic neurons innervating sweat glands release Ach that binds to muscarinic (M) receptors on the cells of the effector tissues. All of the remaining postganglionic neurons of the sympathetic system release norepinephrine (NE) that binds to (α) or (β) receptors on the cells of the effector tissues. The cells of the adrenal medulla, which are modified postganglionic neurons in the sympathetic system, release epinephrine (EPI) and NE into the circulation.

acetylcholine removal occurs in less than 1 msec. The most important mechanism for the removal of norepinephrine from the neuroeffector junction is the *reuptake* of this neurotransmitter into the sympathetic postganglionic neuron that released it. Norepinephrine may then be metabolized intraneuronally by *monoamine oxidase* (MAO). The circulating catecholamines, epinephrine and norepinephrine, are inactivated by *catechol-O-methyltransferase* (COMT) in the liver. Of lesser importance in terms of the termination of neurotransmitter activity is the diffusion of the neurotransmitter away from the synapse. The neurotransmitter is then eliminated by extraneuronal sites.

9.9 Receptors for autonomic neurotransmitters

As discussed in the previous section, all of the effects of the ANS in tissues and organs throughout the body, including smooth muscle contraction or relaxation, alteration of myocardial activity, and increased or decreased glandular secretion, are carried out by only three substances: *acetylcholine*, *norepinephrine*, and *epinephrine*. Furthermore, each of these substances may stimulate activity in some tissues and inhibit activity in others. How can this wide variety of effects on many different tissues be carried out by so few neurotransmitters or hormones? The effect caused by any of these substances is determined by the receptor distribution in a particular tissue and the biochemical properties of the cells in that tissue, specifically, the second messenger and enzyme systems present within the cell.

The neurotransmitters of the ANS and the circulating catecholamines bind to specific receptors on the cell membranes of the effector tissue. All adrenergic receptors and muscarinic receptors are coupled to *G proteins* that are also embedded within the plasma membrane. Receptor stimulation causes activation of the G protein and the formation of an intracellular chemical, the *second messenger*. (The neurotransmitter molecule, which cannot enter the cell itself, is the *first messenger*.) The function of the intracellular second messenger molecules is to elicit tissue-specific biochemical events within the cell which alter the cell's activity. In this way, a given neurotransmitter may stimulate the same type of receptor on two different types of tissue and cause two different responses due to the presence of different biochemical pathways within each tissue. (This signal transduction/second messenger mechanism is discussed in more detail in Chapter 10.)

Acetylcholine binds to two types of cholinergic receptors: *nicotinic receptors* and *muscarinic receptors*.

Nicotinic receptors are found on the cell bodies of all postganglionic neurons in the ganglia of the ANS, both sympathetic and parasympathetic divisions. Acetylcholine released from the preganglionic neurons binds to these nicotinic receptors and causes a rapid increase in the cellular permeability to Na^+ ions and Ca^{++} ions. The resulting influx of these two cations causes depolarization and excitation of the postganglionic neurons of the ANS pathways.

Muscarinic receptors are found on the cell membranes of the effector tissues and are linked to G proteins and second messenger systems that carry out the intracellular effects. Acetylcholine released from all parasympathetic postganglionic neurons and some sympathetic postganglionic neurons traveling to sweat glands binds to these receptors. Muscarinic receptors may be either inhibitory or excitatory, depending on the tissue upon which they are found. For example, muscarinic receptor stimulation in the myocardium is inhibitory and decreases heart rate, but stimulation of muscarinic receptors in the lungs is excitatory, causing contraction of airway smooth muscle and bronchoconstriction.

There are two classes of *adrenergic receptors* for norepinephrine and epinephrine: *alpha* (α) and *beta* (β). Furthermore, there are at least two subtypes of receptors in each class: α_1, α_2, β_1, and β_2. All of these receptors are linked to G proteins and second messenger systems that carry out the intracellular effects.

Alpha receptors are the more abundant of the adrenergic receptors. Of the two subtypes, α_1 receptors are more widely distributed on the effector tissues. Alpha-1 receptor stimulation leads to an increase in intracellular calcium. As a result, these receptors tend to be excitatory. For example, stimulation of α_1 receptors causes contraction of vascular smooth muscle resulting in vasoconstriction as well as increased glandular secretion by way of exocytosis.

PHARMACY APPLICATION:
ALPHA-1 ADRENERGIC RECEPTOR ANTAGONISTS

Hypertension, or a chronic elevation in blood pressure, is a major risk factor for coronary artery disease, congestive heart failure, stroke, kidney failure, and retinopathy. An important cause of hypertension is excessive vascular smooth muscle tone or vasoconstriction. Prazosin, an α_1-adrenergic receptor antagonist, is very effective in the management of hypertension. Because α_1-receptor stimulation causes vasoconstriction, drugs that block these receptors result in vasodilation and a decrease in blood pressure (see Table 9.4).

Compared to α_1 receptors, α_2 receptors have only moderate distribution on the effector tissues. Alpha-2 receptor stimulation causes a decrease in cyclic adenosine monophosphate (cAMP) within the tissue's cells and, therefore, results in inhibitory effects such as smooth muscle relaxation and decreased glandular secretion. However, α_2 receptors have important presynaptic effects. Where α_1 receptors are found on the effector tissue cells at the neuroeffector junction, the α_2 receptors are found on the varicosities of the postganglionic neuron. Norepinephrine released from this neuron binds to not only the α_1 receptors on the effector tissue to cause some physiological effect, it also binds to the α_2 receptors on the postganglionic neuron. Alpha-2

Table 9.4 Agonists and Antagonists of Autonomic Receptors

Receptor	Neurotransmitter/ Catecholamine	Agonist	Antagonist
Nicotinic	Acetylcholine	Nicotine	Mecamylamine
Muscarinic	Acetylcholine	Bethanechol	Atropine, Scopolamine
Alpha one	NE, EPI	Phenylephrine	Prazosin
Alpha two	NE, EPI	Clonidine	Yohimbine
Beta 1 (selective)	NE, EPI	Dobutamine	Metoprolol
Beta 2 (selective)		Albuterol	
Beta (nonselective)	EPI	Isoproterenol, Epinephrine	Propranolol, Carvedilol

receptor stimulation results in *presynaptic inhibition* or in a decrease in the release of norepinephrine. (Although it is the *postganglionic* neuron, at the neuroeffector junction this neuron is considered *presynaptic*.) In this way, norepinephrine inhibits its own release from the sympathetic postganglionic neuron and controls its own activity. Both α_1 and α_2 receptors have equal affinity for norepinephrine released directly from sympathetic neurons as well as circulating epinephrine released from the adrenal medulla.

Stimulation of each type of β *receptor* leads to an increase in intracellular cAMP. Whether this results in an excitatory or an inhibitory response depends upon the specific cell type. As with α receptors, β receptors are also unevenly distributed, with β_2 receptors the more common subtype on the effector tissues. Beta-2 receptors tend to be inhibitory. For example, β_2 receptor stimulation causes relaxation of vascular smooth muscle and airway smooth muscle resulting in vasodilation and bronchodilation, respectively. Beta-2 receptors have a significantly greater affinity for epinephrine than for norepinephrine. Furthermore, terminations of sympathetic pathways are not found near these receptors. Therefore, β_2 receptors are stimulated only indirectly by circulating epinephrine instead of by direct sympathetic nervous activity.

Beta-1 receptors are the primary adrenergic receptors on the heart. (A small percentage of the adrenergic receptors on the myocardium are β_2.) Both subtypes of β receptors on the heart are excitatory, and stimulation leads to an increase in cardiac activity. Beta-1 receptors are also found on certain cells in the kidney. Epinephrine and norepinephrine have equal affinity for β_1 receptors.

Beta-3 (β_3) receptors are found primarily in adipose tissue. Stimulation of these receptors, which are innervated and have a stronger affinity for norepinephrine, causes lipolysis.

A summary of the autonomic receptors and their neurotransmitters, agonists, and antagonists is found in Table 9.4.

PHARMACY APPLICATION: SYMPATHOMIMETIC DRUGS

Sympathomimetic drugs are those that produce effects in a tissue resembling those caused from stimulation by the sympathetic nervous system. An important use for these drugs is in the treatment of bronchial asthma that is characterized by bronchospasm. As discussed, bronchodilation occurs following β_2-adrenergic receptor stimulation. Nonselective β receptor agonists, such as epinephrine and isoproterenol, are capable of causing bronchodilation. However, a potential problem with these drugs is that they stimulate *all* β receptors, including β_1 receptors on the heart. Therefore, an undesirable side effect of treatment with these nonselective agents is an increase in heart rate. Instead, β_2-selective drugs, such as albuterol, are chosen for this therapy. They are equally effective in causing bronchodilation without the adverse cardiac effects of the nonselective agents (see Table 9.4).

9.10 Functions of the autonomic nervous system

The two divisions of the ANS are dominant under different conditions. As stated previously, the sympathetic system is activated during emergency "fight-or-flight" reactions and during exercise. The parasympathetic system is predominant during quiet, resting conditions. As such, the physiological effects caused by each system are quite predictable. In other words, all of the changes in organ and tissue function induced by the sympathetic system work together to support strenuous physical activity, and the changes induced by the parasympathetic system are appropriate for when the body is resting. Several of the specific effects elicited by sympathetic and parasympathetic stimulation of various organs and tissues are summarized in Table 9.5.

The "fight-or-flight" reaction elicited by the sympathetic system is essentially a whole-body response. Changes in organ and tissue function throughout the body are coordinated so that there is an increase in the delivery of well-oxygenated, nutrient-rich blood to the working skeletal muscles. Both heart rate and myocardial contractility are increased so that the heart pumps more blood per minute. Sympathetic stimulation of vascular smooth muscle causes widespread vasoconstriction, particularly in the organs of the gastrointestinal system and in the kidneys. This vasoconstriction serves to "redirect" or redistribute the blood away from these metabolically inactive tissues and toward the contracting muscles. Bronchodilation in the lungs facilitates the movement of air in and out of the lungs so that the uptake of oxygen from the atmosphere and the elimination of carbon dioxide from the body are maximized. An enhanced rate of glycogenolysis (breakdown of glycogen into its component glucose molecules) and gluconeogenesis

Table 9.5 Effects of Autonomic Nerve Activity on Some Effector Tissues

Tissue	Sympathetic Receptor	Sympathetic Stimulation	Parasympathetic Stimulation
Eye			
Radial muscle of iris	α_1	Contraction (dilation of pupil; mydriasis)	—
Sphincter muscle of iris		—	Contraction (constriction of pupil; miosis)
Ciliary muscle	β_2	Relaxation for far vision	Contraction for near vision
Heart	β_1, β_2	↑ Heart rate ↑ Force of contraction ↑ Rate of conduction	↓ Heart rate ↓ Rate of conduction
Arterioles			
Skin	α_1	Strong constriction	—
Abdominal viscera	α_1	Strong constriction	—
Kidney	α_1	Strong constriction	—
Skeletal muscle	α_1, β_2	Weak constriction	—
Spleen	α_1	Contraction	—
Lungs			
Airways	β_2	Bronchodilation	Bronchoconstriction
Glands	α_1, β_2	↓ Secretion	↑ Secretion
Liver	α_1, β_2	Glycogenolysis Gluconeogenesis	↑ Secretion of bile
Adipose Tissue	β_3	Lipolysis	—
Sweat Glands	Muscarinic	Generalized sweating	—
	α_1	Localized sweating	—
Piloerector Muscles	α_1	Contraction (Erection of hair, goose bumps)	—
Adrenal Medullae	Nicotinic	↑ Secretion of epinephrine, norepinephrine	—
Salivary Glands	α_1, β_2	Small volume K^+ and water secretion	Large volume K^+ and water secretion; amylase secretion
Stomach			
Motility	α_1, β_2	Decreased	Increased

continued

Table 9.5 (continued) Effects of Autonomic Nerve Activity on Some Effector Tissues

Tissue	Sympathetic Receptor	Sympathetic Stimulation	Parasympathetic Stimulation
Sphincters	α_1	Contraction	Relaxation
Secretion			Stimulation
Intestine			
Motility	α_1, β_2	Decreased	Increased
Sphincters	α_1	Contraction	Relaxation
Secretion			Stimulation
Gallbladder	β_2	Relaxation	Contraction
Pancreas			
Exocrine	α	↓ Enzyme secretion	↑ Enzyme secretion
Endocrine (Islets β cells)	α	↓ Insulin secretion	↑ Insulin secretion
Urinary Bladder			
Detrusor muscle (bladder wall)	β_2	Relaxation	Contraction
Urethra sphincter		Contraction	Relaxation
Kidney	β_1	↑ Renin secretion	—

(formation of new glucose from noncarbohydrate sources) in the liver increases the concentration of glucose molecules in the blood. This is necessary for the brain, as glucose is the only nutrient molecule that it can utilize to form metabolic energy. An enhanced rate of lipolysis in adipose tissue increases the concentration of fatty acid molecules in the blood. Skeletal muscles then utilize these fatty acids to form metabolic energy for contraction. Generalized sweating elicited by the sympathetic system enables the individual to thermoregulate during these conditions of increased physical activity and heat production. Finally, the eye is adjusted such that the pupil dilates, letting more light in toward the retina (*mydriasis*), and the lens adapts for distance vision.

The parasympathetic system decreases heart rate, which helps to conserve energy under resting conditions. Salivary secretion is enhanced to facilitate the swallowing of food. Gastric motility and secretion are stimulated to begin the processing of ingested food. Intestinal motility and secretion are also stimulated to continue the processing and to facilitate the absorption of these nutrients. Exocrine and endocrine secretion from the pancreas is promoted. Enzymes released from the exocrine glands of the pancreas contribute to the chemical breakdown of the food in the intestine, and insulin released from the pancreatic islets promotes the storage of nutrient molecules within the

tissues once they are absorbed into the body. Another bodily maintenance type of function caused by the parasympathetic system is contraction of the urinary bladder which results in urination. Finally, the eye is adjusted such that the pupil contracts (*miosis*) and the lens adapts for near vision.

PHARMACY APPLICATION: CHOLINOMIMETIC DRUGS

Cholinomimetic drugs produce effects in a tissue resembling those caused from stimulation by the parasympathetic nervous system. These drugs have many important uses including the treatment of gastrointestinal and urinary tract disorders that involve depressed smooth muscle activity without obstruction. For example, postoperative ileus is characterized by a loss of tone or paralysis of the stomach or bowel following surgical manipulation. Urinary retention may also occur postoperatively, or it may be secondary to spinal cord injury or disease (neurogenic bladder). Typically, parasympathetic stimulation of the smooth muscle in each of these organ systems causes contraction to maintain gastrointestinal motility as well as urination. There are two different approaches in the pharmacotherapy of these disorders. One type of agent would be a muscarinic receptor agonist that would mimic the effect of the parasympathetic neurotransmitter, acetylcholine, and stimulate smooth muscle contraction. One of the most widely used agents in this category is bethanechol, which can be given subcutaneously (see Table 9.4). Another approach is to increase the concentration and, therefore, activity of endogenously produced acetylcholine in the neuroeffector junction. Administration of an acetylcholinesterase inhibitor prevents the degradation and removal of neuronally released acetylcholine. In this case, neostigmine is the most widely used agent. Neostigmine may be given either subcutaneously or orally.

PHARMACY APPLICATION: MUSCARINIC RECEPTOR ANTAGONISTS

Inspection of the retina during an ophthalmoscopic examination is greatly facilitated by mydriasis, or the dilation of the pupil. Parasympathetic stimulation of the circular muscle layer in the iris causes contraction and a decrease in the diameter of the pupil. Administration of a muscarinic receptor antagonist, such as atropine or scopolamine, prevents this smooth muscle contraction. As a result, sympathetic stimulation of the radial muscle layer is unopposed. This causes an increase in the diameter of the pupil. These agents are given in the form of eyedrops that act locally and limit the possibility of systemic side effects (see Table 9.4).

9.11 Adrenal medulla

A mass sympathetic discharge, which typically occurs during the "fight-or-flight" response and during exercise, involves the simultaneous stimulation of organs and tissues throughout the body. Included among these tissues are the adrenal medullae that release epinephrine and norepinephrine into the blood. In large part, the indirect effects of these catecholamines are similar to and, therefore, reinforce those of direct sympathetic stimulation. However, there are some important differences in the effects of the circulating catecholamines and those of norepinephrine released from sympathetic nerves, including:

- Duration of activity
- Breadth of activity
- Affinity for β_2 receptors

The *duration of activity* of the catecholamines is significantly longer than that of neuronally released norepinephrine. Therefore, the effects on the tissues are more prolonged. This difference has to do with the mechanism of inactivation of these substances. Norepinephrine is immediately removed from the neuroeffector junction by way of reuptake into the postganglionic neuron. This rapid removal limits the duration of the effect of this neurotransmitter. In contrast, there are no enzymes in the blood to degrade the catecholamines. Instead, the catecholamines are inactivated by COMT in the liver. As one might expect, the hepatic clearance of these hormones from the blood would require several passes through the circulation. Therefore, the catecholamines are available to cause their effects for a comparatively longer period of time (minutes as opposed to milliseconds).

Because they travel in the blood, organs and tissues throughout the body are exposed to the catecholamines. Therefore, they are capable of stimulating tissues that are not directly innervated by sympathetic nerve fibers, airway smooth muscle, hepatocytes, and adipose tissue, in particular. As a result, the catecholamines have a much wider *breadth of activity* compared to norepinephrine released from sympathetic nerves.

The third important feature that distinguishes the catecholamines from neuronally released norepinephrine involves epinephrine's *affinity for* β_2 *receptors*. Norepinephrine has a very limited affinity for these receptors. Therefore, circulating epinephrine causes effects that differ from those of direct sympathetic innervation, including:

- Greater stimulatory effect on the heart
- Relaxation of smooth muscle
  - Vascular
  - Bronchial

- Gastrointestinal
- Genitourinary

Epinephrine and norepinephrine have equal affinity for β_1 receptors, the predominant adrenergic receptor on the heart. However, the human heart also contains a small percentage of β_2 receptors that, like β_1 receptors, are excitatory. Therefore, epinephrine is capable of stimulating a greater number of receptors and of causing a *greater stimulatory effect on the myocardium*.

Beta-2 adrenergic receptors are also found on smooth muscle in several organ systems. These receptors tend to be inhibitory and cause *relaxation of the smooth muscle*. Vascular smooth muscle in skeletal muscle contains both α_1 and β_2 receptors. Norepinephrine, which stimulates only the excitatory α_1 receptors, causes strong vasoconstriction. However, epinephrine, which stimulates both types of receptors, causes only weak vasoconstriction. The vasodilation resulting from β_2 receptor stimulation opposes and, therefore, weakens the vasoconstriction resulting from α_1 receptor stimulation. Given that skeletal muscle may account for 40% of an adult's body weight, the potential difference in vasoconstriction, blood pressure, and the distribution of blood flow could be significant.

Another noteworthy example of the relaxation of smooth muscle by way of β_2 receptor stimulation involves the airways. Bronchodilation, or the opening of the airways, facilitates airflow in the lungs. Any direct sympathetic innervation to the lungs is irrelevant in this respect, as only circulating epinephrine is capable of stimulating these receptors on airway smooth muscle.

Medical terminology

adrenergic (ăd-rĕn-ĕr′jĭk): Refers to nerve fibers that release norepinephrine, receptors that bind with norepinephrine or epinephrine, or a substance that has catecholamine-like activity.

affinity (ă-fĭn′ĭ-tē): Attraction; ability to bind with a specific receptor.

cholinergic (kō″lĭn-ĕr′jĭk): Refers to nerve fibers that release acetylcholine, receptors that bind with acetylcholine, or a substance that has acetylcholine-like activity.

gap junction: Minute pore between muscle cells that allows for intercellular communication.

miosis (mī-ō′sĭs): Contraction of the pupil.

mydriasis (mĭd-rī′ă-sĭs): Dilation of the pupil.

tonic (tŏn′ĭk): Refers to a pattern of continuous, prolonged activity.

vagus (vā′gŭs): Cranial nerve X, the major parasympathetic nerve; it innervates structures in the thoracic and abdominal cavities.

varicosity (văr″ĭ-kŏs′ĭ-tē): Swelling at the terminal of a postganglionic axon from which a neurotransmitter is released.

Bibliography

Bear, M. F., Connors, B. W., and Paradiso, M. A., *Neuroscience, Exploring the Brain,* 3rd ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2007.

Brown, J. H., and Taylor, P., Muscarinic Receptor Agonists and Antagonists, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 7.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., Elsevier/Saunders, Philadelphia, PA, 2006.

Hoffman, B. B., Adrenoceptor-Activating and Other Sympathomimetic Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical Books/McGraw-Hill, New York, 2001, chap. 9.

Hoffman, B. B., Adrenoceptor Antagonist Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical Books/McGraw-Hill, New York, 2001, chap. 10.

Iversen, S., Iversen, L., and Saper, C., The Autonomic Nervous System and the Hypothalamus, in *Principles of Neuroscience,* 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 49.

Katzung, B. G., Introduction to Autonomic Pharmacology, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical Books/McGraw-Hill, New York, 2001, chap. 6.

Naftel, J. P., and Hardy, S. G. P., Visceral Motor Pathways, in *Fundamental Neuroscience for Basic and Clinical Applications,* 3rd ed., Haines, D. E., Ed., Churchill Livingstone/Elsevier, Philadelphia, PA, 2006, chap. 29.

Pappano, A. J., Cholinoceptor-Activating and Cholinesterase-Inhibiting Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical Books/McGraw-Hill, New York, 2001, chap. 7.

Pappano, A. J., Cholinoceptor-Blocking Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical Books/McGraw-Hill, New York, 2001, chap. 8.

Rhoades, R., and Pflanzer, R., *Human Physiology,* 4th ed., Thomson Learning, Pacific Grove, CA, 2003.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Taylor, P., Agents Acting at the Neuromuscular Junction and Autonomic Ganglia, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 9.

Westfall, T. C., and Westfall, D. P., Neurotransmission: The Autonomic and Somatic Motor Nervous Systems, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 6.

Westfall, T. C., and Westfall, D. P., Adrenergic Agonists and Antagonists, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 10.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter ten

The endocrine system

Study objectives

- Differentiate between the primary functions of the nervous system and the endocrine system
- Describe the biochemical and functional distinctions between steroid hormones, protein/peptide hormones, and amine hormones
- Explain the beneficial effects of the binding of hormones to plasma proteins
- Explain how hormones are eliminated
- Distinguish between a trophic and a nontrophic hormone
- Describe the three types of hormone interactions
- Explain the two primary mechanisms by which hormones carry out their effects
- Describe how the effects of hormones are amplified
- Describe how the pituitary gland is formed during embryonic development
- Describe the anatomical and functional relationships between the hypothalamus and the pituitary gland
- Explain how negative feedback mechanisms limit the release of hormones from the adenohypophysis
- List the functions and describe the mechanisms that regulate the release of hormones from the neurohypophysis
- List the functions and describe the mechanisms that regulate the release of hormones from the adenohypophysis
- Discuss the functions and the factors that regulate the release of the following hormones: thyroid hormones, calcitonin, parathyroid hormone, aldosterone, cortisol, adrenal androgens, insulin, and glucagon

10.1 Introduction

There are two major regulatory systems that contribute importantly to homeostasis: the *nervous system* and the *endocrine system*. In order to maintain relatively constant conditions in the internal environment of the body, each of these systems influences the activity of all of the other organ systems. The nervous system coordinates fast, precise responses, such as muscle contraction. As such, the electrical impulses generated by this system are very rapid and of short duration (msec). The endocrine system regulates

metabolic activity within the cells of the organs and tissues. In contrast to the nervous system, the endocrine system coordinates activities that require longer duration (hours, days) rather than speed. Examples of such activities include growth, the long-term regulation of plasma volume and blood pressure, and the coordination of menstrual cycles in females. The endocrine system carries out its effects through the production of *hormones*, chemical messengers that exert a regulatory effect on the cells of the body. Secreted from *endocrine glands*, which are ductless structures, hormones are released directly into the blood. They are then transported by the circulation to the tissues upon which they exert their effects. Because they travel in the blood, the serum concentrations of hormones are very low (10^{-11} to 10^{-9} M); therefore, these molecules must be very potent.

This chapter will consider the hormones secreted from several endocrine glands, including the pituitary, thyroid, parathyroid, and adrenal glands, as well as the pancreas. It is important to note, however, that hormones may be secreted from organs and tissues whose primary functions are other than endocrine in nature. For example, the hypothalamus synthesizes and releases many hormones that influence the secretion of hormones from the anterior pituitary gland. In addition, the hypothalamus synthesizes the hormones that are subsequently released from the posterior pituitary gland. Other tissues and their hormones include the following:

- *Heart*: Atrial natriuretic hormone (Chapter 21)
- *Kidneys*: Erythropoietin (Chapters 17 and 21)
- *Stomach*: Gastrin (Chapter 20)
- *Small intestine*: Secretin, cholecystokinin, gastric inhibitory peptide (Chapter 20)
- *Gonads* (*ovaries and testes*): Estrogen, progesterone, testosterone (Chapter 11)

Further discussion of these hormones may be found in subsequent chapters as indicated within the parentheses above.

Generally, a single hormone does not affect all of the body's cells. The tissues that respond to a hormone are referred to as the *target tissues*. The cells of these tissues possess specific receptors to which the hormone binds. This receptor binding then elicits a series of events that influences cellular activities.

10.2 Biochemical classification of hormones

Hormones are classified into three biochemical categories (see Table 10.1):

1. Steroids
2. Proteins/peptides
3. Amines

Table 10.1 Distinguishing Features of Steroid, Protein/Peptide, and Amine Hormones

Feature	Steroid Hormones	Protein/Peptide Hormones	Amine Hormones	
			Thyroid Hormones	Catecholamines
Synthesis	Synthesized on demand; derived from cholesterol	Synthesized in advance; derived from amino acids	Synthesized in advance; stored as part of thyroglobulin	Synthesized in advance; derived from tyrosine
Transport in Blood	Bound to carrier proteins	Soluble in plasma	Bound to carrier proteins	Soluble in plasma
Half-Life	Long	Short	Long	Short
Location of Receptor	Within cytoplasm	Membrane surface	Within cell nucleus	Membrane surface
Mechanism of Action	Gene activation	Second messenger systems	Gene activation	Second messenger systems
Target Tissue Response	Synthesis of new enzymes	Modification of existing enzymes	Synthesis of new enzymes	Modification of existing enzymes
Sites of Production	Adrenal cortex, testes, ovaries, placenta	Hypothalamus, pituitary gland, thyroid gland, parathyroid glands, pancreas	Thyroid gland	Adrenal medulla

Steroid hormones are produced by the adrenal cortex, testes, ovaries, and the placenta. Synthesized from cholesterol, these hormones are lipid soluble. Therefore, they cross cell membranes readily and bind to receptors found intracellularly. However, their lipid solubility renders them insoluble in blood. Therefore, these hormones are transported in the blood bound to proteins. This protein binding protects these hormones from rapid elimination and, therefore, steroid hormones tend to have relatively long half-lives.

Steroid hormones are not typically preformed and stored for future use within the endocrine gland. Because they are lipid soluble, they could readily diffuse out of the cells, and physiological regulation of their release would not be possible. Finally, steroid hormones are absorbed easily by the gastrointestinal tract and, therefore they may be administered orally.

The *protein/peptide hormones* are derived from amino acids. Most hormones are of this type. Peptide hormones contain fewer than 100 amino acids (e.g., antidiuretic hormone, oxytocin, glucagon). Protein hormones contain more than 100 amino acids (e.g., growth hormone). Glycoprotein hormones are peptide chains containing more than 100 amino acids as well as one or more carbohydrate groups (e.g., follicle-stimulating hormone, luteinizing hormone, thyroid-stimulating hormone).

These hormones are preformed and stored for future use in membrane-bound secretory granules. When needed, they are released by exocytosis. Protein/peptide hormones are water soluble, circulate in the blood predominantly in an unbound form, and, therefore, tend to have relatively short half-lives. Because these hormones are unable to cross the cell membranes of their target tissues, they bind to receptors on the membrane surface. Protein/peptide hormones cannot be administered orally as they are easily digested in the gastrointestinal tract. Instead, they are usually administered by injection (e.g., insulin). Small peptides are able to cross through mucus membranes and, therefore, may be given sublingually or intranasally. For example, Miacalcin®, the synthetic form of the hormone calcitonin, is prepared in the form of a nasal spray.

Amine hormones, which are all derived from the amino acid tyrosine, include the thyroid hormones and the catecholamines. The thyroid hormones tend to be biologically similar to the steroid hormones. They are mainly insoluble in the blood and are transported predominantly (>99%) bound to proteins. As such, these hormones have longer half-lives (triiodothyronine, T_3 = 24 hours; thyroxine, T_4 = 7 days). Furthermore, thyroid hormones cross cell membranes to bind with intracellular receptors, and they may be administered orally (e.g., Synthroid®). In contrast to the steroid hormones, however, the thyroid hormones have the unique property of being stored extracellularly in the thyroid gland as part of the thyroglobulin molecule.

The catecholamines are biologically similar to the protein/peptide hormones. These hormones are soluble in the blood and are transported in an unbound form. Therefore, the catecholamines have a relatively short half-life. Because these hormones do not cross cell membranes, they bind to receptors

on the membrane surface. Finally, the catecholamines are stored intracellularly in secretory granules for future use.

10.3 Transport of hormones

As discussed in the previous section, steroid and thyroid hormones are transported in the blood bound to plasma proteins. The serum concentrations of free hormone (H), plasma protein (P), and bound hormone (HP) are in equilibrium:

$$[H] \times [P] = [HP]$$

The total hormone concentration in the plasma is the sum of the free hormone and the bound hormone. When the concentration of the free form of a hormone decreases, then more of this hormone will be released from the binding proteins. Only the free hormone may diffuse across capillary walls and reach the target cells. Logically, the free hormone is the biologically active form. It binds to the target tissue to cause its actions, and it is involved with the negative feedback control of its secretion.

The binding of hormones to plasma proteins has several beneficial effects:

- Facilitation of transport
- Prolonged half-life
- Hormone reservoir

The steroid and thyroid hormones are minimally soluble in the blood. Binding to plasma proteins renders them water soluble and facilitates their transport. Second, protein binding prolongs the circulating half-life of these hormones. Because they are lipid soluble, they cross cell membranes easily. As the blood flows through the kidney, these hormones would enter cells or be filtered and lost to the urine if they were not held in the blood by the impermeable plasma proteins. Finally, the protein-bound form of the hormone serves as a "reservoir" of hormone that minimizes the changes in free hormone concentration when hormone secretion from its endocrine gland changes abruptly.

10.4 Hormone elimination

Once a hormone has carried out its effect, the plasma concentration of the hormone must return to its normal range. The concentration of a hormone in the plasma depends upon its *rate of secretion from the endocrine gland* and its *rate of removal from the blood*.

Removal of a hormone from the blood occurs by either *excretion* or *metabolic transformation*. These mechanisms are carried out primarily by the liver

and the kidneys. For example, hormones may be excreted by way of the bile (liver) or the urine (kidneys).

In some instances, hormones may be metabolized by the cells of their target tissue. Peptide hormones, in the form of hormone-receptor complexes, may be taken up by the target cell by way of endocytosis. The hormone is then degraded intracellularly, and the receptor is subsequently returned to the plasma membrane.

10.5 Functional classification of hormones

Hormones are classified into two functional categories:

1. Trophic hormones
2. Nontrophic hormones

A *trophic hormone* acts on another endocrine gland to stimulate the secretion of its hormone. For example, thyrotropin, or thyroid-stimulating hormone (TSH), stimulates the secretion of thyroid hormones from the thyroid gland. Adrenocorticotropin, or adrenocorticotropic hormone (ACTH), stimulates the adrenal cortex to secrete the hormone cortisol. Both of these trophic hormones, TSH and ACTH, are produced by the pituitary gland. In fact, many trophic hormones are secreted by the pituitary. It is for this reason that the pituitary gland is sometimes referred to as the "master gland," as its hormones regulate the activity of other endocrine glands.

A *nontrophic hormone* acts on nonendocrine target tissues. For example, parathormone (parathyroid hormone), released from the parathyroid glands, acts on bone tissue to stimulate the release of calcium into the blood. Aldosterone, released from the cortical region of the adrenal glands, acts on the kidney to stimulate the reabsorption of sodium into the blood.

10.6 Hormone interactions

Multiple hormones may affect a single target tissue simultaneously. Therefore, the response of the target tissue depends not only on the effects of each hormone individually, but also on the nature of the interaction of the hormones at the tissue. There are three types of hormone interactions:

1. Synergism
2. Permissiveness
3. Antagonism

Synergism occurs when two hormones interact at the target tissue such that the combination of their effects is more than additive. In other words, their combined effect is greater than the sum of their separate effects. For example, epinephrine, cortisol, and glucagon are three hormones that each

increase the level of blood glucose. The magnitude of their individual effects on glucose levels tends to be low to moderate. However, the activity of all three hormones simultaneously results in an increase in blood glucose that is several times greater than the sum of their individual effects.

In *permissiveness*, one hormone enhances the responsiveness of the target tissue to a second hormone. In other words, the first hormone increases the activity of the second hormone. For example, the normal maturation of the reproductive system requires not only reproductive hormones from the hypothalamus, the pituitary, and the gonads; it also requires the presence of thyroid hormone. Although thyroid hormone by itself has no effect on the reproductive system, if it is absent, then the development of this system is delayed. Therefore, thyroid hormone is considered to have a permissive effect on the reproductive hormones, facilitating their actions causing sexual maturation.

The mechanism of permissiveness involves the *up-regulation*, or an increase in the number, of hormone receptors. In the previous example, thyroid hormone up-regulates the receptors for sex hormones such as estrogen, progesterone, or testosterone.

Antagonism occurs when the actions of one hormone oppose the effects of another. For example, insulin decreases blood glucose and promotes the formation of fat. Glucagon, on the other hand, increases blood glucose and promotes the degradation of fat. Therefore, the effects of insulin and glucagon are antagonistic.

10.7 Mechanisms of hormone action

The binding of a hormone to its receptor initiates intracellular events that direct the hormone's action. Ultimately, all hormones produce their effects by altering intracellular protein activity. However, the mechanism by which this occurs depends on the location of the hormone receptor. Receptors are located on the cell surface, in the cytoplasm of the cell, or, in some cases, within the cell nucleus. As a result, there are two general mechanisms by which most hormones carry out their effects:

1. Signal transduction and second messenger systems
2. Gene activation

Protein/peptide hormones and the catecholamines are water-soluble substances and, accordingly, are unable to cross the plasma membrane to enter the cell. Therefore, these hormones must bind to their specific receptors on the cell surface. This receptor binding causes a response within the cell by way of *signal transduction* and the production of intracellular *second-messenger molecules*. The original, extracellular hormone is considered to be the first messenger because it carried the signal to the target tissue.

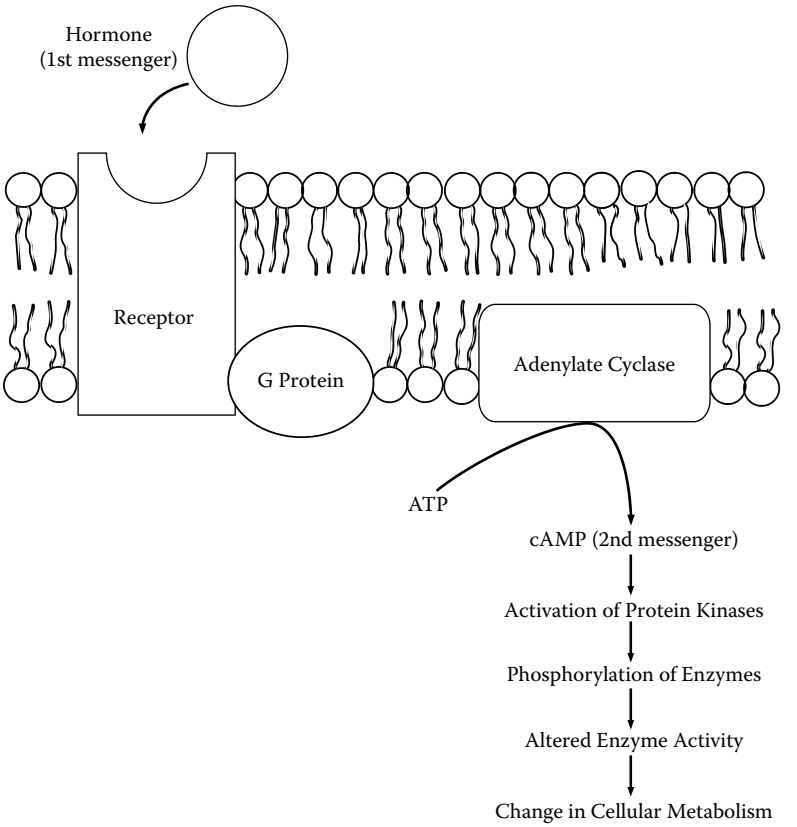


Figure 10.1 The cyclic adenosine monophosphate (cAMP) second-messenger system. The most common second-messenger system activated by the protein/peptide hormones and the catecholamines involves the formation of cAMP. This multistep process is initiated by the binding of the hormone (the first messenger) to its receptor on the cell surface. The subsequent increase in the formation of cAMP (the second messenger) leads to the alteration of enzyme activity within the cell. A change in the activity of these enzymes alters cellular metabolism.

The most common second messenger activated by the protein/peptide hormones and the catecholamines is *cyclic adenosine monophosphate* (cAMP). The pathway by which cAMP is formed and alters cellular function is illustrated in Figure 10.1. The process begins when the hormone binds to its receptor. These receptors are quite large and span the plasma membrane. On the cytoplasmic surface of the membrane, the receptor is associated with a *G protein* that serves as the transducer molecule. In other words, the G protein acts as an intermediary between the receptor and the second messengers that will alter cellular activity. These proteins are referred to as G proteins because they bind with guanosine nucleotides. In an unstimulated cell, the inactive G protein binds guanosine diphosphate (GDP). When the hormone

binds to its G-protein-associated receptor, the G protein releases GDP and binds with *guanosine triphosphate* (GTP) taken up from the cytoplasm. Upon binding with the GTP, the now activated G protein loses its affinity for the receptor and increases its affinity for the plasma-membrane-embedded enzyme, *adenylyl cyclase*. In turn, the adenylyl cyclase becomes activated and splits adenosine triphosphate (ATP) to form cAMP. The cAMP molecule serves as the second messenger that carries out the effects of the hormone inside the cell. The primary function of cAMP is to activate *protein kinase A*. This kinase then attaches phosphate groups to specific enzymatic proteins in the cytoplasm. The phosphorylation of these enzymes either enhances or inhibits their activity. This results in either the enhancement or the inhibition of specific cellular reactions and processes. Either way, cellular metabolism has been altered.

There are several noteworthy aspects of this mechanism of hormonal action:

- Onset of hormonal effects
- Multiple systems
- Cellular specificity
- Amplification of effect
- Prolonged action of hormones

First, the *onset of the response* of cells to the activation of second-messenger systems is comparatively rapid (within minutes). This mechanism involves changing the activity of *existing* enzymes rather than producing *new* enzymes, which is a more lengthy process. Second, there are *many signal transduction and second-messenger pathways*. For example, another signal transduction system involves the opening and closing of ion channels. Furthermore, some tissues use calcium as a second messenger and others use cyclic guanosine monophosphate (cGMP) or inositol triphosphate (IP_3). Third, the *cellular specificity* of a hormone's effect depends on the different kinds of enzyme activities that are ultimately modified in different target tissues. For example, antidiuretic hormone causes the reabsorption of water from the kidneys and the constriction of smooth muscle in blood vessels; two very different effects in two very different tissues caused by one hormone. Fourth, the effects elicited by second-messenger systems involve a multistep process. This is advantageous because at many of these steps a multiplying or cascading effect takes place which causes *amplification* of the initial signal. For example, one molecule of the hormone epinephrine binding to its receptor on a hepatocyte may result in the production of 10 million molecules of glucose. Finally, *hormonal action is prolonged*. Once an enzyme is activated, the effects are as long-lasting as the enzyme and no longer depend upon the presence of the initiating hormone.

PHARMACY APPLICATION: MULTIPLE TARGETS IN THE PHARMACOTHERAPY OF ASTHMA

Asthma is an inflammatory illness characterized by severe bronchoconstriction and wheezing. The pharmacotherapy for this disorder is directed toward a reduction in inflammation and bronchodilation. There are several different pharmacologic approaches and different drug classes that may be employed in treating the asthmatic patient. The two types of asthma medications discussed below are targeted for different points in the signal transduction and second-messenger pathway which contribute to the regulation of airway smooth muscle tone.

β_2 adrenergic receptor stimulation leads to increased levels of the second messenger, cAMP. An increase in cAMP causes bronchodilation. Albuterol (Proventil®, Ventolin®) is an example of a short-acting β_2 adrenergic receptor agonist. Long-acting medications include salmeterol xinafoate (Serevent®) and formoterol (Foradil®).

Theophylline, a methylxanthine, also increases the levels of the second messenger, cAMP. The mechanism by which it does so is uncertain. One proposed mechanism involves the inhibition of the enzyme phosphodiesterase. This enzyme breaks down cAMP. Therefore, inhibition of phosphodiesterase leads to an accumulation of cAMP and an increase in signal transduction through this pathway. Another proposed mechanism for the increase in cAMP in response to theophylline involves the antagonism of adenosine receptors. Adenosine receptor stimulation may lead to bronchoconstriction in asthmatics. Blockade of these receptors with theophylline would prevent this effect and lead to bronchodilation.

The importance of theophylline as a therapeutic agent has decreased over the last several years as other newly developed drugs, such as corticosteroids (fluticasone propionate, Flovent®) and leukotriene receptor antagonists (montelukast, Singulair®), have come into use. Research continues into the identification of other phosphodiesterase inhibitors that may be more potent bronchodilators with fewer side effects.

Steroid hormones and thyroid hormones carry out their effects by way of *gene activation*. In contrast to the protein/peptide hormones that alter *existing* enzyme activity, these hormones induce the synthesis of *new* enzymes, which then influence cellular metabolism.

Hormones in this category are lipophilic and easily enter the cells of the target tissue by diffusing through the plasma membrane. Typically, the receptors for steroid hormones are found in the cytoplasm. The hormone binds to its receptor forming a *hormone-receptor complex*. This complex then translocates to the cell nucleus. In some cases, unbound receptors are found in the nucleus, and hormone-receptor complexes may be formed there. For example, receptors for thyroid hormones are found within the nucleus.

In addition to binding with their specific hormones, these receptors are capable of binding to DNA at specific attachment sites referred to as the *hormone response elements* (HREs). Each of the steroid hormones binds with its receptor and attaches to a different HRE. A given HRE is located in a DNA span adjacent to the gene that will be transcribed. Binding of the hormone–receptor complex to the DNA activates specific genes within the target cell, resulting in the formation of *mRNA molecules*. The mRNA then diffuses into the cytoplasm and binds to a ribosome where protein synthesis takes place. These new proteins serve as enzymes that regulate cellular reactions and processes. Once again, intracellular metabolism has been altered.

As with signal transduction and second-messenger systems, the mechanism of gene activation allows for amplification of the hormone's effect. For example, a single hormone-activated gene induces the formation of many mRNA molecules, and each mRNA molecule may be used to synthesize many enzyme molecules. Furthermore, the effects of hormones using this mechanism are prolonged. As long as the newly synthesized enzyme is active, the effect of the initiating hormone persists.

10.8 The pituitary gland

The *pituitary gland*, or *hypophysis*, is located at the base of the brain just below the hypothalamus. It is composed of two functionally and anatomically distinct lobes (see Figure 10.2):

1. Neurohypophysis (posterior pituitary)
2. Adenohypophysis (anterior pituitary)

As its name implies, the *neurohypophysis* is derived embryonically from nervous tissue. It is essentially an outgrowth of the hypothalamus and is composed of bundles of axons, or neural tracts, of neurosecretory cells originating in two hypothalamic nuclei. These neurons are referred to as *neurosecretory cells* because they generate action potentials (*neuro-*) as well as synthesize hormones (*-secretory*). The cell bodies of the neurosecretory cells in the supraoptic nuclei produce primarily antidiuretic hormone (ADH), and the cell bodies of the paraventricular nuclei produce primarily oxytocin. These hormones are then transported down the axons to the neurohypophysis and stored in membrane-bound vesicles in the neuron terminals. Much like neurotransmitters, the hormones are released in response to the arrival of action potentials at the neuron terminal.

The *adenohypophysis* is derived embryonically from glandular tissue, specifically, *Rathke's pouch*. This tissue originates from the oropharynx, or the roof of the mouth. It then migrates toward the embryonic nervous tissue destined to form the neurohypophysis. When these two tissues come into contact, Rathke's pouch loses its connection with the roof of the mouth and the pituitary gland is formed.

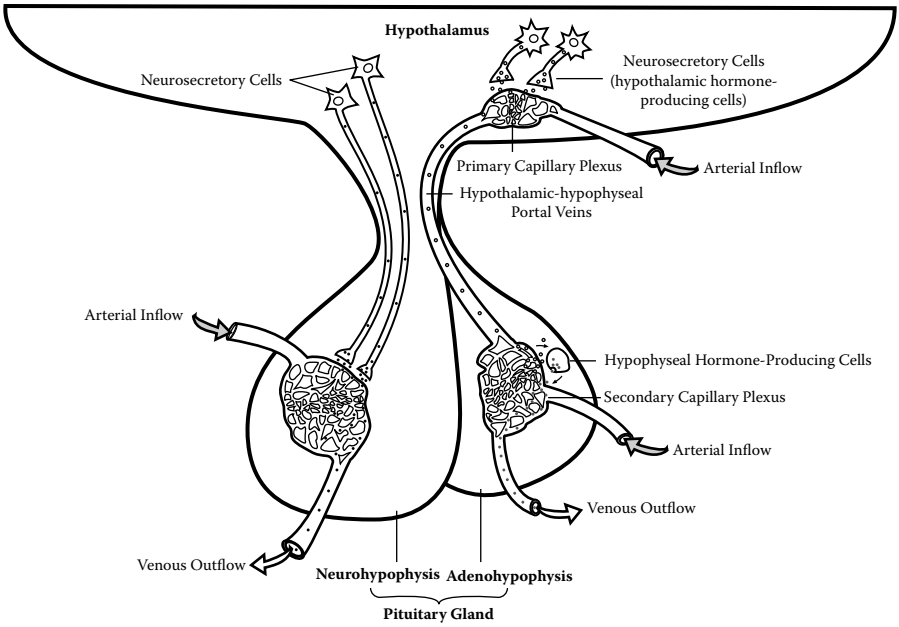


Figure 10.2 Anatomical and functional relationship between the hypothalamus and the pituitary gland. The neurohypophysis is derived from the hypothalamus. This anatomical connection allows the hypothalamus to directly influence the function of the neurohypophysis. Action potentials conducted by neurosecretory cells originating in the hypothalamus cause the release of hormones stored in the neurohypophysis. The adenohypophysis is derived from glandular tissue and, therefore, has no anatomical connection to the hypothalamus. The release of hormones from the adenohypophysis is regulated by hypothalamic hormones. These hormones are carried to the adenohypophysis through the hypothalamic-hypophyseal portal veins. Hypothalamic hormones enter the tissue of the adenohypophysis and influence the production of adenohypophyseal hormones. Hormones released from both regions of the pituitary gland, the neurohypophysis and the adenohypophysis, are removed from the pituitary gland by the venous outflow blood and transported to target tissues throughout the body.

Unlike the neurohypophysis that releases hormones originally synthesized in the hypothalamus, the adenohypophysis synthesizes its own hormones in specialized groups of cells. Similar to the neurohypophysis, however, the release of these hormones into the blood is regulated by the hypothalamus.

10.9 Relationship between the hypothalamus and the pituitary gland

The hypothalamus plays a very important role in the maintenance of homeostasis. It carries out this function in large part by regulating the activities of

the neurohypophysis and the adenohypophysis. For example, the hypothalamus processes signals from other regions of the nervous system including information regarding pain and emotional states, such as depression, anger, and excitement. In addition, because it is not protected by the blood–brain barrier, it monitors the composition of the blood and helps to regulate the concentration of nutrients, electrolytes, water, and hormones. In other words, it is an important processing center for information concerning the internal environment. This information is then used to control the release of hormones from the pituitary.

Due to their different embryonic origins, the neurohypophysis and the adenohypophysis are regulated by the hypothalamus using two very different mechanisms:

1. Nerve signals
2. Hormonal signals

As discussed previously, the neurohypophysis has a direct anatomical connection to the hypothalamus. Therefore, the hypothalamus regulates the release of hormones from the neurohypophysis by way of *nerve signals*. Action potentials generated by the neurosecretory cells originating in the hypothalamus are transmitted down the neuronal axons to the nerve terminals in the neurohypophysis and stimulate the release of the hormones into the blood. The tracts formed by these axons are referred to as the *hypothalamic-hypophyseal tracts* (see Figure 10.2). The action potentials are initiated by various forms of sensory input to the hypothalamus. Specific forms of sensory input that regulate the release of ADH and oxytocin are described in Section 10.11.

The adenohypophysis does not have a direct anatomical connection with the hypothalamus; therefore, regulation of hormone secretion by way of neuronal signals is not possible. Instead, these two structures are associated by a specialized circulatory system, and the secretion of hormones from the adenohypophysis is regulated by *hormonal signals* from the hypothalamus (see Figure 10.2). Systemic arterial blood is directed first to the hypothalamus. The exchange of materials between the blood and the interstitial fluid of the hypothalamus takes place at the *primary capillary plexus*. The blood then flows to the adenohypophysis through the *hypothalamic-hypophyseal portal veins*. Portal veins are blood vessels that connect two capillary beds. The second capillary bed in this system is the *secondary capillary plexus* located in the adenohypophysis.

Located in close proximity to the primary capillary plexus in the hypothalamus are specialized neurosecretory cells. In fact, the axons of these cells terminate on the capillaries. The neurosecretory cells synthesize two types of hormones: *releasing hormones* and *inhibiting hormones* (see Table 10.2). Each of these hormones helps to regulate the release of a particular hormone from the adenohypophysis. For example, thyrotropin-releasing hormone

Table 10.2 Summary of the Major Hormones

Location	Hormone	Target Tissues	Major Functions of Hormone
Hypothalamus	Releasing and inhibiting hormones (GnRH, TRH, CRH, PRF, PIH, GHRH, GHIH)	Adenohypophysis	Control of the release of hormones from the adenohypophysis
	Antidiuretic hormone (ADH)	Kidney	Promotes the reabsorption of water
		Arterioles	Vasoconstriction
	Oxytocin	Uterus	Contraction of smooth muscle
		Mammary glands	Ejection of milk
Adenohypophysis	Follicle-stimulating hormone (FSH)	Females: ovaries	Development of follicles; secretion of estrogen
		Males: testes	Spermatogenesis
	Luteinizing hormone (LH)	Females: ovaries	Rupture of follicle and ovulation; secretion of estrogen and progesterone from the corpus luteum
		Males: testes	Secretion of testosterone
	Thyroid-stimulating hormone (TSH)	Thyroid gland	Secretion of thyroid hormones (T_3, T_4)
	Adrenocorticotropic hormone (ACTH)	Adrenal cortex	Secretion of cortisol

	Prolactin (PRL)	Mammary glands	Breast development; lactation
	Growth hormone (GH)	Bone, visceral tissues	Growth of skeleton and visceral tissues; increase blood glucose and fatty acids; protein synthesis
Thyroid Gland	Triiodothyronine (T_3), Tetraiodothyronine (T_4)	Most tissues	Growth and maturation; normal neurological development and function; increase in metabolic rate
	Calcitonin	Bone	Decrease blood calcium
Parathyroid Glands	Parathyroid hormone (PTH)	Bone, kidneys, intestine	Increase blood calcium; decrease blood phosphate; activation of vitamin D_3
Adrenal Medulla	Epinephrine and norepinephrine	Adrenergic receptors throughout the body	"Fight-or-flight" response; reinforces the effects of the sympathetic nervous system
Adrenal Cortex	Mineralocorticoids (aldosterone)	Kidney	Reabsorption of sodium; excretion of potassium
	Glucocorticoids (cortisol)	Most tissues	Increase blood glucose and fatty acids; adaptation to stress
	Androgens	Various tissues	Secondary sex characteristics in females
Pancreas	Insulin	Most tissues	Cellular uptake, utilization and storage of glucose, fatty acids, and amino acids
	Glucagon	Most tissues	Increase blood glucose and fatty acids

produced by the neurosecretory cells of the hypothalamus stimulates the secretion of thyrotropin from the thyrotrope cells of the adenohypophysis. The hypothalamic releasing hormone is picked up by the primary capillary plexus, travels through the hypothalamic-hypophyseal portal veins to the anterior pituitary, leaves the blood by way of the secondary capillary plexus, and exerts its effect on the appropriate cells of the adenohypophysis. The hypophyseal hormone, in this case, thyrotropin, is then picked up by the secondary capillary plexus, removed from the pituitary by the venous blood, and delivered to its target tissue.

A noteworthy feature of this specialized circulation is that the regulatory hypothalamic hormones are delivered directly to the adenohypophysis by the portal system. Therefore, the concentration of these hormones remains very high because they are not diluted in the blood of the entire systemic circulation.

10.10 Negative feedback control of hormone release

In many cases, hormones released from the adenohypophysis are part of a three-hormone axis that includes:

1. The hypothalamic hormone
2. The adenohypophyseal hormone
3. The endocrine gland hormone

The hypothalamic hormone either stimulates or inhibits the secretion of the adenohypophyseal hormone. The trophic hormone from the adenohypophysis then stimulates the release of a hormone from another endocrine gland. This final endocrine gland hormone not only carries out its effects on its target tissues, it may also exert a *negative feedback* effect on the release of the hypothalamic and the adenohypophyseal hormones. In this way, this final hormone regulates its own release (see Figure 10.3). This process is referred to as *long-loop negative feedback*. The adenohypophyseal hormone may also exert a negative feedback effect on the hypothalamic hormone and limit its own release. This process is referred to as *short-loop negative feedback*.

10.11 Hormones of the neurohypophysis

10.11.1 Antidiuretic hormone (ADH)

ADH, also referred to as *vasopressin*, has two major effects, both of which are reflected by its names:

- Antidiuresis (decrease in urine formation by the kidney)
- Vasoconstriction of arterioles

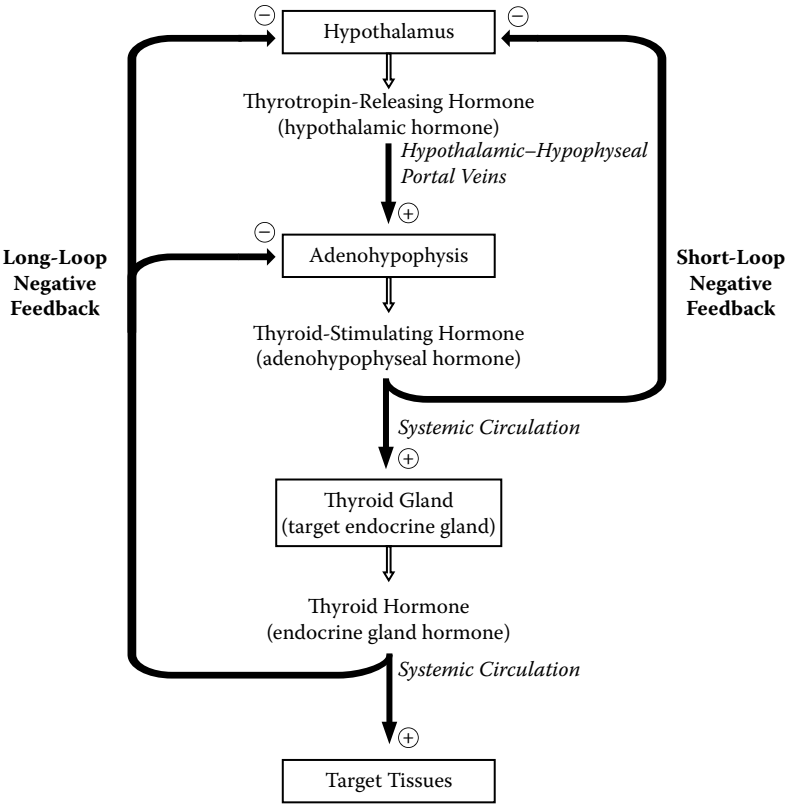


Figure 10.3 Negative-feedback regulation of hormone release. Hormones released from the adenohypophysis are often part of a three-hormone axis that includes the hypothalamic hormone, the adenohypophyseal hormone, and the target endocrine gland hormone. Long-loop negative feedback occurs when the final hormone in the axis inhibits the release of the hypothalamic or the adenohypophyseal hormones. Short-loop negative feedback occurs when the adenohypophyseal hormone inhibits the release of the hypothalamic hormone. Illustrated in this figure is the thyrotropin-releasing hormone/thyroid-stimulating hormone/thyroid hormone axis.

ADH promotes the reabsorption of water from the tubules of the kidney, or *antidiuresis*. Specifically, it acts on the collecting ducts and increases the number of water channels, or *aquaporins,* which increases the diffusion coefficient for water. This results in the conservation of water by the body and the production of a low volume of concentrated urine. This process may begin within 5 to 10 minutes of the presence of ADH. The reabsorbed water affects both plasma osmolarity and blood volume. The additional water dilutes the plasma and decreases plasma osmolarity. Furthermore, the added water expands blood volume. This effect of ADH on the kidney occurs at relatively low concentrations.

At higher concentrations, ADH causes *constriction of arterioles*. This vasoconstriction serves to increase blood pressure.

ADH secretion is regulated by several factors:

- Plasma osmolarity
- Blood volume
- Blood pressure
- Alcohol

The primary factor that influences ADH secretion is a change in *plasma osmolarity*. There are *osmoreceptors* in the hypothalamus which are located in close proximity to the ADH-producing neurosecretory cells. Stimulation of these osmoreceptors by an increase in plasma osmolarity results in stimulation of the neurosecretory cells, an increase in the frequency of action potentials in these cells, and the release of ADH from their axon terminals in the neurohypophysis. The water conserved due to the effect of ADH on the kidney helps to reduce plasma osmolarity or dilute the plasma back to normal.

The hypothalamic osmoreceptors have a threshold of 280 mOsM/L. Below this value, the osmoreceptors are not stimulated, and little or no ADH is secreted. Maximal ADH levels occur when plasma osmolarity is about 295 mOsM/L. Within this range, the regulatory system is very sensitive, with measurable increases in ADH secretion occurring in response to 1% changes in plasma osmolarity. The regulation of ADH secretion is an important mechanism by which a normal plasma osmolarity of 290 mOsM/L is maintained. (As discussed in Chapter 2, the maintenance of a normal plasma osmolarity is important in preventing the inappropriate movement of water into or out of the body's cells. This would lead to cellular swelling and possible lysis or cellular dehydration, respectively.)

Other factors regulating ADH secretion include blood volume and blood pressure. A decrease in *blood volume* of 10% to 25% causes an increase in ADH secretion that may be sufficient to cause vasoconstriction as well as antidiuresis. A decrease in *mean arterial blood pressure* of 5% or more also causes an increase in ADH secretion. The resulting water conservation and vasoconstriction help increase blood volume and blood pressure back to normal. Furthermore, the effect of blood pressure on ADH secretion may be correlated to the increase that occurs during sleep when blood pressure decreases. The result is the production of a low volume of highly concentrated urine that is less likely to elicit the micturition (urination) reflex and interrupt sleep.

In contrast to these physiological factors, *alcohol* inhibits the secretion of ADH, allowing for the loss of water from the kidneys. Therefore, the consumption of alcoholic beverages may actually lead to excessive water loss and dehydration instead of volume expansion.

PHARMACY APPLICATION: DIABETES INSIPIDUS

Diabetes insipidus is caused by a decrease in ADH secretion or a decrease in the kidney's response to ADH. This absolute or relative deficiency of ADH leads to excessive urination (polyuria) due to the patient's inability to reabsorb water from the kidneys. The excessive water loss leads to extreme thirst (polydipsia).

The preferred medication for treating diabetes insipidus is desmopressin acetate (synthetic 1-deamino-8-D-arginine vasopressin; DDAVP). Desmopressin may be prepared as a sterile solution for injection or as a nasal solution for intranasal administration. The duration of effect with continued intranasal dosing is from 5 to 24 hours. Desmopressin may also be administered orally. Because this drug is a peptide, it is rapidly inactivated by the enzyme trypsin. Therefore, oral doses of desmopressin are 10 to 20 times that of the intranasal dose. The typical nasal dosage is 10 to 40 mcg daily, and the oral dosage is 0.1 to 0.2 mg every 12 to 24 hours.

10.11.2 Oxytocin

The second hormone synthesized in the hypothalamus and secreted from the neurohypophysis is oxytocin. This hormone also exerts its major effects on two different target tissues. Oxytocin stimulates the *contraction of uterine smooth muscle* and the *contraction of myoepithelial cells* in the mammary glands.

Oxytocin stimulates *contraction of the smooth muscle in the wall of the uterus*. During labor, this facilitates the delivery of the fetus. During intercourse, this may facilitate the transport of the sperm through the female reproductive tract for fertilization of the ovum. Oxytocin also causes *contraction of the myoepithelial cells* surrounding the alveoli of the mammary glands. This results in "*milk-letdown*," or the expulsion of milk from deep within the gland into the larger ducts from which the milk can be obtained more readily by the nursing infant.

The secretion of oxytocin is regulated by *reflexes elicited by cervical stretch and by nursing*. Normally, as labor begins, the fetus is positioned head down. This orientation exerts pressure on the cervix and causes it to stretch. Sensory neurons in the cervix are thereby activated to transmit signals to the hypothalamus which will stimulate the release of oxytocin from the neurohypophysis. This hormone then enhances uterine contraction that causes further pressure and stretch of the cervix. Additional oxytocin is released, and so on, until there is adequate pressure buildup so that delivery takes place. In the lactating breast, nursing activates sensory neurons in the nipple to transmit signals to the hypothalamus to stimulate oxytocin release from

the neurohypophysis and, therefore, milk-letdown. Interestingly, this reflex may also be triggered through a conditioned response where the sight or sound of the hungry infant is sufficient to enhance oxytocin secretion. In contrast, the release of oxytocin from the neurohypophysis may be inhibited by pain, fear, or stress.

The function of oxytocin in males is not clearly understood.

PHARMACY APPLICATION: INDUCTION AND PREVENTION OF PARTURITION

Agents designed to stimulate uterine contractions are used to induce or augment parturition (labor). Circumstances that warrant the induction of labor include Rh problems (see Chapter 17), maternal diabetes, premature rupture of membranes, and preeclampsia. Synthetic oxytocin (Pitocin®, Syntocinon®) is the clinical drug of choice. This medication is administered intravenously.

Preterm labor may be treated with oxytocin receptor antagonists. The drug atosiban decreases the frequency of uterine contractions and reduces the likelihood of preterm delivery.

10.12 Hormones of the adenohypophysis

10.12.1 Gonadotropins

The gonadotropins, follicle-stimulating hormone and luteinizing hormone, exert their effects on the gonads (ovaries in the female and testes in the male). Taken together, the gonadotropins stimulate the gonads to *produce gametes* (*ova and sperm*) and *secrete sex hormones* (*estrogen, progesterone,* and *testosterone*).

Follicle-stimulating hormone (FSH), as its name indicates, stimulates the development of the ovarian follicles in females. It is within the follicles that the ova, or eggs, develop. This hormone also induces the secretion of estrogen from the follicle. In males, FSH acts on the Sertoli cells of the testes which are involved with the production of sperm, or spermatogenesis.

Luteinizing hormone (LH) is also named for its effects in the female which are to cause the rupture of the follicle to release the ovum and to cause the conversion of the ovarian follicle into a *corpus luteum* (Latin, *yellow body*). This hormone also induces the secretion of estrogen and progesterone from the corpus luteum. In males, LH acts on the Leydig cells of the testes to stimulate the secretion of testosterone.

Both FSH and LH are produced by the same cell type in the adenohypophysis, the gonadotrope. The release of FSH and LH is regulated by the hypothalamic releasing hormone, *gonadotropin-releasing hormone* (GnRH).

10.12.2 Thyroid-stimulating hormone (TSH)

Also known as thyrotropin, this hormone regulates the growth and metabolism of the thyroid gland. Furthermore, it stimulates the synthesis and release of the thyroid hormones, T_3 and T_4. The release of TSH from the thyrotrope cells of the adenohypophysis is induced by *thyrotropin-releasing hormone* (TRH).

10.12.3 Adrenocorticotropic hormone (ACTH)

Also known as adrenocorticotropin, this hormone stimulates growth and steroid production in the adrenal cortex. Specifically, it stimulates the secretion of cortisol and other glucocorticoids that are involved with carbohydrate metabolism. The release of ACTH from the adenohypophysis is influenced by more than one factor. *Corticotropin-releasing hormone* (CRH) from the hypothalamus stimulates the secretion of ACTH. In addition, ACTH secretion follows a *diurnal pattern*, with a peak in the early morning and a valley late in the evening.

10.12.4 Prolactin (PRL)

This hormone, which is produced by the lactotrope cells of the adenohypophysis, is involved with the initiation and maintenance of lactation in females. Its function in males is uncertain. Lactation involves three processes:

1. Mammogenesis
2. Lactogenesis
3. Galactopoeisis

Mammogenesis is the growth and development of the mammary glands that produce the milk. This process requires the actions of many hormones, including estrogens and progestins, in addition to PRL. *Lactogenesis* is the initiation of lactation. During pregnancy, lactation is inhibited by the high levels of estrogens and progestins. At delivery, the levels of these two hormones fall, allowing PRL to initiate lactation. *Galactopoeisis* is the maintenance of milk production. This process requires both PRL and oxytocin.

The release of prolactin from the adenohypophysis is normally inhibited by *prolactin-inhibiting hormone* (PIH) (*dopamine*) from the hypothalamus. Prolactin secretion is also controlled by *prolactin-releasing factor* (PRF). The release of PRF from the hypothalamus is mediated by reflexes elicited by suckling and breast stimulation.

10.12.5 Growth hormone (GH)

Also known as somatotropin, this is one of the few hormones that exerts its effects on organs and tissues throughout the body. Growth hormone is essential for the normal growth and development of the skeleton as well as visceral, or soft, tissues from birth until young adulthood. Growth of the skeleton involves both an increase in bone thickness and an increase in bone length. The mechanism of this growth involves the stimulation of osteoblast (bone-forming cell) activity and the proliferation of the epiphyseal cartilage in the ends of the long bones. The growth of visceral tissues occurs by *hyperplasia* (increasing the number of cells) and *hypertrophy* (increasing the size of cells). Growth hormone causes hyperplasia by stimulating cell division and by inhibiting apoptosis (programmed cell death). Growth hormone causes cellular hypertrophy by promoting protein synthesis and inhibiting protein degradation.

The growth-promoting effects of GH are carried out by *somatomedins*, which are peptides found in the blood. At least four somatomedins have been identified and described. The two most important somatomedins, which are structurally and functionally similar to insulin, are referred to as *insulin-like growth factors I and II* (IGF-I and IGF-II). Growth hormone stimulates the production of IGF-I in the liver, which is the predominant source of that found in the circulation. Local production of IGF-I also occurs in many target tissues. IGF-I is thought to mediate the growth-promoting effects of GH throughout life. Levels of both GH and IGF-I increase in parallel during puberty and other periods of growth in children. In contrast, IGF-II production does not depend on GH. Instead, IGF-II is thought to be important during fetal growth and development and is secreted in response to prolactin. The role of IGF-II in the adult is unknown.

Growth hormone also has many metabolic actions in the body:

- Protein metabolism
  - Increase tissue amino acid uptake
  - Increase DNA transcription
  - Increase RNA translation
  - Increase protein synthesis
  - Decrease catabolism of proteins and amino acids
- Lipid metabolism
  - Stimulation of lipolysis
  - Inhibition of lipogenesis
  - Increase in blood fatty acids
  - Increase in the utilization of fats over carbohydrates for energy production
- Carbohydrate metabolism
  - Decrease in glucose uptake and utilization by muscle
  - Increase in the hepatic output of glucose (glycogenolysis)
  - Increase in blood glucose

The net effects of these actions include enhanced growth due to protein synthesis, enhanced availability of fatty acids for use by skeletal muscle as an energy source, and glucose sparing for the brain which can use only this nutrient molecule as a source of energy.

The release of GH from the adenohypophysis is regulated by two hypothalamic hormones, *growth hormone-releasing hormone* (GHRH) and *growth hormone-inhibiting hormone* (GHIH) *(somatostatin*). Any factor or condition that enhances the secretion of GH could do so by stimulating GHRH release or by inhibiting GHIH release. The secretion of GH follows a diurnal rhythm with GH levels low and constant throughout the day and with a marked burst of GH secretion approximately 1 hour following the onset of sleep (deep or stage III and IV sleep). Other factors that stimulate GH secretion include exercise, stress, trauma, hypoglycemia, and starvation, especially with severe protein deficiency. Factors that inhibit GH secretion include hyperglycemia and aging. In most individuals, the production of GH decreases after 30 years of age and falls to about 25% of the adolescent level in very old age. This decrease in GH production is likely a critical factor in the loss of lean muscle mass at a rate of 5% per decade and the gain of body fat at the same rate after 40 years of age.

10.13 Thyroid gland

The thyroid gland is a butterfly-shaped structure lying over the anterior surface of the trachea just below the larynx. This gland produces two classes of hormones synthesized by two distinct cell types:

1. Thyroid hormones (T_3 and T_4), synthesized by follicular cells
2. Calcitonin, synthesized by parafollicular cells

10.13.1 Thyroid hormones

Internally, the thyroid consists of *follicles*, spherical structures whose walls are formed by a single layer of epithelial cells called *follicular cells*. The center of each follicle contains a homogenous gel referred to as *colloid*. Thyroid hormones are stored here as a component of the larger molecule, *thyroglobulin*. The amount of the thyroid hormones stored within the colloid is enough to supply the body for 2 to 3 months.

Thyroid hormones, derived from the amino acid tyrosine, are unique because they contain *iodine*. At this time, its incorporation into thyroid hormones is the only known use for iodine in the body. There are two thyroid hormones, named for the number of iodides added to the tyrosine residues of the thyroglobulin, *triiodothyronine* (T_3) and *tetraiodothyronine* (T_4) (*thyroxine*). Although significantly more T_4 is synthesized by the thyroid gland, T_3 is the active hormone. At the target tissue, T_4 is converted to the more potent T_3 by the removal of an iodine.

The thyroid hormones are lipophilic and are relatively insoluble in the plasma. Therefore, they are transported throughout the circulation bound to plasma proteins such as *thyroxine-binding globulin* (75%) and albumins (25%). Approximately 99.96% of circulating thyroxine is protein bound. Bound hormone is not available to cause any physiological effects; however, it is in equilibrium with the remaining 0.04% which is unbound. It is this free form of the hormone that is able to bind to receptors on the target tissues and cause its effects.

Thyroid hormone's primary action is to alter cellular metabolism. In fact, complete lack of thyroid hormone secretion causes the basal metabolic rate to fall 40% to 50% below normal. Significant excess of thyroid hormone secretion may increase the basal metabolic rate to 60% to 100% above normal. Some of the many metabolic effects of thyroid hormone in the body include the following:

- Growth and maturation
  - Perinatal lung maturation
  - Normal skeletal growth
- Neurological
  - Normal fetal and neonatal brain growth and development
  - Regulation of neuronal proliferation and differentiation, myelinogenesis, neuronal outgrowth, and synapse formation
  - Normal central nervous system (CNS) function in adults
- Sympathetic nervous system function
  - Increase in the number of β-adrenergic receptors
  - Increase in heart rate
  - Tremor
  - Sweating
- Cardiovascular system
  - Vasodilation
  - Increase in heart rate
  - Increase in myocardial contractility
  - Increase in cardiac output
- Metabolism
  - Increase in the number and activity of mitochondria
  - Increase in the rate of ATP formation
  - Increase in basal metabolic rate
  - Stimulation of all metabolic pathways, both anabolic and catabolic
  - Increase in carbohydrate uptake and utilization
  - Increase in glycolysis and gluconeogenesis
  - Increase in lipolysis and fatty acid mobilization
  - Increase in oxygen consumption
  - Increase in the rate and depth of breathing
  - Increase in heat production

As mentioned previously, thyroid hormones are secreted at a relatively steady rate. The secretion of hormones from the thyroid gland is regulated by negative feedback in the hypothalamic–pituitary–thyroid axis. The hypothalamus secretes thyrotropin-releasing hormone (TRH), which stimulates the release of TSH from the adenohypophysis of the pituitary. Thyroid-stimulating hormone then stimulates the release of T_3 and T_4 from the thyroid. In this hormone axis, negative-feedback inhibition is exerted primarily at the level of the pituitary. As the intracellular concentration of T_3 in the thyrotrope cells of the pituitary increases, then there is a decrease in the responsiveness of these TSH-producing cells to TRH. The mechanism of this decreased responsiveness involves the downregulation, or a decrease in the number, of TRH receptors. This results in a decrease in the secretion of TSH, and, consequently, a decrease in the secretion of T_3 and T_4. The excess of intracellular T_3 which elicits the negative feedback control of secretion comes from two sources: 80% from the removal of iodine from serum T_4 within the thyrotrope cells and 20% from serum T_3. Other factors that alter the release of TRH and, therefore, thyroid hormones, include exposure to cold (increased secretion) as well as anxiety (decreased secretion).

PHARMACY APPLICATION: HYPOTHYROIDISM

A deficiency in thyroid hormone production leads to hypothyroidism. In regions of the world where there is sufficient iodine in the diet, the most common cause of this disorder is Hashimoto's thyroiditis. This is an autoimmune disease characterized by the production of antibodies directed against the thyroid gland. Clinical manifestations of hypothyroidism are related to myxedema (fluid accumulation which is most obvious in the face) and the hypometabolic state. Specifically, symptoms include the gradual onset of weakness and fatigue, weight gain, and nervous system dysfunction (e.g., mental dullness, lethargy, impaired memory).

Hypothyroidism is treated pharmacologically with synthetic preparations of thyroid hormones. Levothyroxine sodium (L-T4, Synthroid, Levoxyl®) is typically administered in a tablet form. Because it is an amine hormone, absorption of this medication from the small intestine is variable and incomplete, such that 50% to 80% of the dosage is absorbed.

10.13.2 Calcitonin

This hormone, which is also secreted from the thyroid gland, is synthesized by the *parafollicular cells* (C cells) located between the follicles. The primary effect of *calcitonin* is to decrease the blood levels of calcium and phosphate. The mechanism of action involves the direct inhibition of osteoclast activity

which decreases bone resorption, or the breakdown of bone. This results in less demineralization of the bone and, therefore, a decrease in the release of calcium and phosphate from the bone into the blood. Calcitonin has no direct effect on bone formation by osteoblasts.

The release of calcitonin from the thyroid is regulated by plasma calcium levels through negative feedback. An increase in the level of calcium in the blood stimulates the secretion of calcitonin, and a decrease in the level of calcium in the blood inhibits secretion.

PHARMACY APPLICATION:
THERAPEUTIC EFFECTS OF CALCITONIN

The normal physiological effects of calcitonin are relatively weak. However, when used pharmacologically, the effects of this hormone are very important. Paget's disease is characterized by a significant increase in osteoclast activity and, therefore, a high rate of bone turnover and hypercalcemia. Because there is minimal species variation, human calcitonin or calcitonin from other species may be used to treat this disorder. Therefore, pharmacological intervention includes the administration of salmon calcitonin (Miacalcin®) that will depress the bone resorption and ease the symptoms of Paget's. Salmon calcitonin, which is 20 times more potent than human calcitonin, has also been approved for therapeutic use in patients with postmenopausal osteoporosis.

10.14 Parathyroid glands

There are four small parathyroid glands embedded on the posterior surface of the thyroid gland as it wraps around the trachea. *Parathyroid hormone* (PTH) (parathormone) is the principle regulator of calcium metabolism. The overall effects of PTH include an *increase in blood levels of calcium* and a *decrease in blood levels of phosphate*. PTH carries out these effects through multiple mechanisms of action:

- Decrease in calcium excretion in the urine
- Increase in phosphate excretion in the urine
- Increase in bone resorption
- Activation of vitamin D_3

Calcium is freely filtered along with the other components of the plasma through the nephrons of the kidney. Most of this calcium is reabsorbed back into the blood from the proximal tubule of the nephron. However, because the kidneys produce about 180 liters of filtrate per day, the amount of calcium that is filtered is substantial. Therefore, the physiological regulation of even a small percentage of calcium reabsorption may have a significant effect on

the amount of calcium in the blood. Parathyroid hormone acts primarily on the distal tubules and the collecting ducts to increase the reabsorption of calcium from these segments of the tubule and decrease the amount excreted in the urine. This activity conserves calcium and increases its concentration in the blood.

Phosphate, which is also freely filtered with the plasma through the nephrons of the kidney, is reabsorbed into the blood from the proximal tubule. Parathyroid hormone acts on this segment to decrease phosphate reabsorption and increase the amount excreted in the urine.

Parathyroid hormone stimulates bone resorption by increasing the number and the activity of osteoclasts. This demineralization process in the bone releases both calcium and phosphate into the blood. Although the action of PTH on the bone appears to increase blood phosphate, its action on the kidney, which increases phosphate excretion in the urine, more than compensates for this increase, and the net effect is a decrease in serum phosphate.

The final mechanism of action of PTH involves the activation of vitamin D_3 through the stimulation of 1α-hydroxylase in the kidney. In the gastrointestinal tract, vitamin D_3 is essential for the absorption of calcium. Enhanced absorption of calcium from dietary sources serves to further increase the concentration of calcium in the blood. Many foods, in particular, dairy products, which are rich in calcium, are fortified with vitamin D.

The release of PTH from the parathyroid glands is regulated by plasma calcium levels through negative feedback. A decrease in the level of calcium in the blood stimulates the secretion of PTH within minutes. If the calcium deficiency persists, the parathyroid glands will hypertrophy as much as fivefold or more. An increase in the level of calcium in the blood inhibits secretion. Persistent calcium excess may cause the parathyroid glands to reduce in size.

10.15 Adrenal glands

There are two adrenal glands, one located on the superior surface of each kidney. These glands are composed of two distinct anatomical and, therefore, functional regions:

1. Adrenal medulla
2. Adrenal cortex

10.15.1 Adrenal medulla

The adrenal medulla, derived from neural crest tissue, forms the inner portion of the adrenal gland. It is the site of production of the *catecholamines*, epinephrine and norepinephrine. The catecholamines serve as a circulating counterpart to the sympathetic neurotransmitter, norepinephrine, which is released directly from sympathetic neurons to the tissues. As such, the adrenal

medulla and its hormonal products play an important role in the activity of the sympathetic nervous system. This is fully discussed in Chapter 9.

10.15.2 Adrenal cortex

The adrenal cortex, forming the outer portion of the adrenal gland, accounts for 80% to 90% of the weight of the gland. It is the site of synthesis of many types of steroid hormones, including:

- Mineralocorticoids
- Glucocorticoids
- Adrenal androgens

10.15.2.1 Mineralocorticoids

The primary mineralocorticoid is *aldosterone*. The actions of this hormone include the *stimulation of renal retention of sodium* and the *promotion of renal excretion of potassium*.

Total lack of aldosterone secretion results in death within several days due to the rapid loss of sodium and chloride from the body (as much as 10% to 20% of the total sodium in the body per day). This sodium deficiency results in a decrease in blood volume and, therefore, cardiac output. Death ensues from cardiogenic shock. The loss of sodium is accompanied by the marked accumulation of potassium in the body fluids.

Aldosterone acts on the distal tubule of the nephron to increase sodium reabsorption (Figure 21.4). The mechanism of action involves an increase in the number of sodium-permeable channels on the luminal surface of the distal tubule and an increase in the activity of the Na^+-K^+ ATPase pump on the basilar surface of the tubule. Sodium diffuses down its concentration gradient out of the lumen and into the tubular cells. The pump then actively removes the sodium from the cells of the distal tubule and into the extracellular fluid so that it may diffuse into the surrounding capillaries and return to the circulation. Due to its osmotic effects, the retention of sodium is accompanied by the retention of water. In other words, wherever sodium goes, water follows it. As a result, aldosterone is very important in the regulation of blood volume and blood pressure. The retention of sodium and water expands the blood volume and, consequently, increases mean arterial pressure. The maximum effect is reached only after several hours.

The retention of sodium is coupled to the excretion of potassium. For every three Na^+ ions reabsorbed, two K^+ ions and one H^+ ion are excreted.

The release of aldosterone from the adrenal cortex is regulated primarily by two important factors:

1. Serum potassium levels
2. Renin-angiotensin system

The mechanism by which *potassium* regulates aldosterone secretion is unclear; however, this ion appears to have a direct effect on the adrenal cortex. A small increase in the level of potassium in the blood results in a several-fold increase in the release of aldosterone. The action of aldosterone on the kidney then decreases the level of potassium back to normal.

Angiotensin II is a potent stimulus for the secretion of aldosterone. The formation of angiotensin II occurs by the following process:

Angiotensinogen

↓ renin

Angiotensin I

↓ ACE

Angiotensin II

This multistep process is initiated by the enzyme renin. Angiotensinogen is a precursor peptide molecule released into the circulation from the liver. In the presence of renin, an enzyme produced by specialized cells in the kidney, angiotensinogen is split to form angiotensin I. This prohormone is then acted upon by angiotensin-converting enzyme (ACE) as the blood passes through the lungs to form angiotensin II. Angiotensin II acts directly on the adrenal cortex to promote aldosterone secretion.

Because this process requires renin in order to occur, it is important to understand the following factors involved in its release from the kidney:

- Decrease in blood volume
- Decrease in blood pressure
- Sympathetic stimulation

A decrease in either blood volume or blood pressure may result in a decrease in the blood flow to the kidney. The kidney monitors renal blood flow by way of stretch receptors in the blood vessel walls. A decrease in renal blood flow stimulates the release of renin. The subsequent secretion of aldosterone causes the retention of sodium and water and, therefore, an increase in blood volume and blood pressure back to normal. An increase in renal blood flow tends to cause the opposite effect.

Sympathetic nerve activity causes an increase in blood pressure through many mechanisms, including an increase in cardiac activity and vasoconstriction. Activation of the sympathetic system also causes the stimulation of β_1-adrenergic receptors on the renin-producing cells which promotes renin release.

10.15.2.2 Glucocorticoids

The primary glucocorticoid is *cortisol*. Receptors for the glucocorticoids are found in all tissues. The overall effects of these hormones include an *increase in blood glucose* and an *increase in blood free fatty acids*.

Cortisol increases blood glucose by several mechanisms of action:

- Decrease in glucose utilization by many peripheral tissues (especially muscle and adipose tissue)
- Increase in availability of gluconeogenic substrates
  - Increase in protein catabolism (especially muscle)
  - Decrease in amino acid transport into extrahepatic cells
  - Increase in amino acid transport into hepatic cells
  - Increase in lipolysis
- Increase in hepatic gluconeogenesis (as much as six- to tenfold)

Cortisol-induced lipolysis not only provides substrates for gluconeogenesis (formation of glucose from noncarbohydrate sources), but it also increases the amount of free fatty acids in the blood. As a result, the fatty acids are used by muscle as a source of energy, and glucose is spared for use by the brain to form energy.

The release of cortisol from the adrenal cortex is regulated by several factors:

- Circadian rhythm
- Stress
- Negative-feedback inhibition by cortisol

Corticotropin-releasing hormone (CRH) secreted from the hypothalamus stimulates the release of ACTH from the adenohypophysis. This pituitary hormone then stimulates the release of cortisol from the adrenal cortex. The hormones of this hypothalamic–pituitary–adrenocortical axis exhibit marked diurnal variation. This variation is due to the diurnal secretion of CRH. The resulting secretion of cortisol increases at night and peaks in the early morning, approximately 1 hour after rising (6 A.M. to 8 A.M.). The levels of cortisol then gradually fall during the day to a low point late in the evening, between 8 P.M. and 12 A.M. This rhythm is influenced by many factors including light–dark patterns, sleep–wake patterns, and eating. After an individual changes time zones, it takes about 2 weeks for this rhythm to adjust to the new time schedule, which may account for some aspects of jet lag.

Cortisol is an important component of the body's response to stress, both physical and psychological. Some of the types of stress that stimulate cortisol secretion include trauma, infection, intense heat or cold, surgery, debilitating disease, and physical restraint. Nervous signals regarding stress are transmitted to the hypothalamus, and the release of CRH is stimulated. The resulting increase in cortisol increases the levels of glucose, free fatty acids, and amino acids in the blood, providing the metabolic fuels that enable the individual to cope with the stress.

A potent inhibitor of this system is cortisol. This hormone exerts a negative-feedback effect on both the hypothalamus and the adenohypophysis and inhibits the secretion of CRH and ACTH, respectively.

PHARMACY APPLICATION:
THERAPEUTIC EFFECTS OF CORTICOSTEROIDS

When administered in pharmacological concentrations (greater than physiological), cortisol and its synthetic analogs (hydrocortisone, prednisone) have potent antiinflammatory and immunosuppressive effects. In fact, these steroids inhibit almost every step of the inflammatory response, including decreased release of vasoactive and chemoattractive factors, decreased secretion of lipolytic and proteolytic enzymes, decreased extravasation of leukocytes to areas of injury, and, ultimately, decreased fibrosis. Typically, the inflammatory response is quite beneficial in that it limits the spread of infection. However, there are many clinical conditions, such as rheumatoid arthritis and asthma, where the response becomes a destructive process. Therefore, although the glucocorticoids have no effect on the underlying cause of disease, the suppression of inflammation by these agents is very important clinically.

Corticosteroids also exert inhibitory effects on the overall immune process. These drugs impair the function of the leukocytes responsible for antibody production and destruction of foreign cells. As a result, corticosteroids are used therapeutically in the prevention of organ transplant rejection.

The therapeutic use of corticosteroids should be undertaken with caution. Because they suppress the inflammatory response and the activity of the immune system, the patient is more susceptible to infection. Other adverse effects include the development of gastric ulcers, hypertension, atherosclerosis, and weight gain.

10.15.2.3 Adrenal androgens

The predominant androgens, or male sex hormones, produced by the adrenal cortex are *dehydroepiandrosterone* (DHEA) and *androstenedione*. These steroid hormones are moderately active androgens; however, in the peripheral tissues, they can be converted to more powerful androgens, such as testosterone, or even to estrogens. The quantities of these hormones that are released from the adrenal cortex are very small. Therefore, the contributions of this source of these hormones to androgenic effects in the male are negligible compared to those of the testicular androgens. However, it is likely that a portion of the early development of male sex organs results from the secretion of these hormones during childhood.

The adrenal gland is the major source of androgens in females. These hormones stimulate pubic and axillary (underarm) hair development in pubertal females. In pathological conditions where adrenal androgens are overproduced, masculinization of females may occur.

10.16 Pancreas

The pancreas is both an exocrine gland and an endocrine gland. The exocrine tissue produces a bicarbonate solution and digestive enzymes. These substances are transported to the small intestine where they play a role in the chemical digestion of food. These functions are fully discussed in Chapter 20.

Scattered throughout the pancreas, surrounded by exocrine cells, are small clusters of endocrine cells referred to as the *islets of Langerhans*. These islets make up only 2% to 3% of the mass of the pancreas. However, their blood supply has been modified so that they receive five to ten times more blood than the exocrine pancreas. Furthermore, this blood, carrying the pancreatic hormones, is then transported through the hepatic portal vein and delivered directly to the liver where the hormones, in a relatively high concentration, carry out many of their metabolic effects. The most important hormones produced by the pancreas that regulate glucose metabolism are insulin and glucagon.

10.16.1 Insulin

Insulin is a peptide hormone produced by the β cells of the islets of Langerhans which constitute approximately 60% of all of the cells of the islets. Insulin is an important anabolic hormone secreted at times when the concentration of nutrient molecules in the blood is high, such as periods following a meal. Its overall effects include allowing the body to use carbohydrates as an energy source and to store nutrient molecules. Specifically, insulin exerts its important actions on the following tissues:

- Liver
  - Increase in glucose uptake
  - Increase in glycogenesis (formation of glycogen, the storage form of glucose)
  - Increase in lipogenesis (formation of triglycerides, the storage form of lipids)
- Adipose tissue
  - Increase in glucose uptake
  - Increase in free fatty acid uptake
  - Increase in lipogenesis
- Muscle
  - Increase in glucose uptake

- Increase in glycogenesis
- Increase in amino acid uptake
- Increase in protein synthesis

Insulin is the only hormone that lowers blood glucose (epinephrine, growth hormone, cortisol, and glucagon all increase blood glucose). It does so by stimulating the uptake of glucose from the blood into the liver, adipose tissue, and muscle. This glucose is first used as an energy source and then it is stored in the form of glycogen in the liver and in muscle. Excess glucose is stored as fat in adipose tissue.

Insulin also plays a role in fat metabolism. In humans, most fatty acid synthesis takes place in the liver. The mechanism of action of insulin involves directing excess nutrient molecules toward metabolic pathways leading to fat synthesis. These fatty acids are then transported to storage sites, predominantly adipose tissue. Finally, insulin stimulates the uptake of amino acids into cells where they are incorporated into proteins.

The secretion of insulin from the pancreas is regulated primarily by the circulating concentration of glucose. When serum glucose increases, the secretion of insulin is stimulated; when serum glucose decreases, the secretion if insulin is inhibited. Insulin secretion typically begins to increase within 10 minutes following the ingestion of food and reaches a peak in 30 to 45 minutes. This increased insulin stimulates the uptake of glucose into the body's cells and lowers serum glucose levels back to normal. Other factors affecting insulin secretion include circulating amino acids and free fatty acids; several gastrointestinal hormones, including gastrin, secretin, cholecystokinin, and gastric inhibitory peptide (which is likely the most important of these hormones); and the parasympathetic nervous system. Each of these factors stimulates the secretion of insulin. Sympathetic nervous stimulation inhibits insulin secretion.

10.16.2 Glucagon

Also a peptide hormone, *glucagon* is produced by the α cells of the islets of Langerhans which constitute approximately 25% of the cells of the islets. The overall effects of glucagon include:

- Increase in hepatic glucose production
  - Glycogenolysis
  - Gluconeogenesis
- Stimulation of lipolysis in the liver and in adipose tissue

The effects of glucagon on glucose metabolism are generally opposite to those of insulin. Acting primarily on the liver, glucagon stimulates glycogenolysis (breakdown of glycogen, the storage form of glucose) and gluconeogenesis, which increase blood glucose levels within minutes. This

hormone also stimulates lipolysis, which increases the circulating concentration of free fatty acids. These molecules may then be used as an alternative energy source by muscle or they may serve as gluconeogenic substrates in the liver. Finally, glucagon stimulates the hepatic uptake of amino acids which also serve as substrates for gluconeogenesis.

Factors that stimulate glucagon secretion include: a decrease in blood glucose, an increase in blood amino acids, sympathetic nervous stimulation, stress, and exercise. Factors that inhibit glucagon secretion include insulin and an increase in blood glucose.

Table 10.2 summarizes the major functions of the hormones discussed in this chapter.

Medical terminology

Amplification (ăm″plĭ-fĭ-kā′shŭn): Enlargement, magnification.

Anabolism (ă-năb′ō-lĭzm): Building of body tissue; metabolic reactions in which substances are synthesized.

Androgenic (ăn″drō-jĕn-ĭk): Causing masculinization.

Antagonism (ăn-tăg′ŏn-ĭzm): Contrary or opposing action.

Antidiuretic (ăn″tī-dī-ū-rĕt′ĭk): Decreasing urine formation.

Cardiogenic (kăr″dē-ō-jĕn′ĭk): Originating in the heart.

Catabolism (kă-tăb′ō-lĭzm): Breaking down of body tissue; metabolic reactions in which substances are degraded.

Cervix (sĕr′vĭks): Neck of the uterus.

Diurnal (dī-ŭrn′ăl): Repeating once each 24 hours.

Down-regulate (down-rĕg′ūlāt″): To decrease the number of receptors on a cell membrane.

Extravasation (ĕks-trăv″ă-sā′shŭn): Escape of fluid from the vascular compartment into the tissue spaces.

Follicle (fŏl′ĭ-kl): In the ovary, structure that releases the mature ovum.

Gamete (găm′ēt): Mature female or male reproductive cell (e.g., ovum, sperm).

Gluconeogenesis (gloo″kō-nē″ō-jĕn′ĕ-sĭs): Formation of glucose from precursors other than carbohydrates (e.g., amino acids, lactic acid).

Glycogenesis (glī″kō-jĕn′ĕ-sĭs): Formation of glycogen from glucose molecules.

Glycogenolysis (glī″kō-jĕn-ŏl′ĭ-sĭs): Breakdown of glycogen into its component glucose molecules.

Gonad (gō′năd): Reproductive organ (e.g., ovary in the female; teste in the male).

Half-life (hăf′līf): Time required to eliminate half of a chemical substance.

Hypercalcemia (hī″pĕr-kăl-sē′mē-ă): Presence of excess calcium in the blood.

Hyperglycemia (hī″pĕr-glī-sē′mē-ă): Presence of excess glucose in the blood.

Hypertrophy (hī-pĕr′trŏ-fē): Excessive development or growth of an organ or tissue.

Hypoglycemia (hī″pō-glī-sē′mē-ă): Abnormally low level of glucose in the blood.
Intranasal (ĭn″tră-nā′zl): Within the nasal cavity.
Lipogenesis (līp″ō-jĕn′ĕ-sĭs): Formation of fat.
Lipolysis (līp″-ŏl′ĭ-sĭs): Breakdown of fat to release fatty acids.
Neurosecretory cell (nū″rō-sē′krĕ-tor-ē sĕl): In the hypothalamus, neuron that secretes a hormone from its axon terminal in response to an action potential.
Osteoblast (ŏs′tē-ō-blăst): Cell that causes bone formation.
Osteoclast (ŏs′tē-ō-klăst): Cell that causes bone resorption or bone breakdown.
Ovum (ō′vŭm): Female reproductive cell; egg.
Parturition (păr-tū-rĭsh′ŭn): Process of giving birth to offspring.
Plexus (plĕks′ŭs): Network of nerves or blood vessels.
Polydipsia (pŏl″ē-dĭp′sē-ă): Excessive thirst.
Polyuria (pŏl″ē-ū′rē-ă): Excessive urination.
Portal vein (pŏr′tăl vān): A vein that provides circulatory communication between two capillary beds.
Postmenopausal (pōst″mĕn-ō-paw′zăl): Occurring after permanent cessation of menstruation.
Potent (pō′tĕnt): Highly effective medicinally.
Proteolytic (prō″tē-ō-lĭt′ĭk): Causing the breakdown of proteins.
Reservoir (rĕz′ĕr-vwor): A place for storage.
Resorption (rē-sorp′shŭn): In the bone, demineralization, or the removal of bone tissue.
Sperm (spĕrm): Male reproductive cell.
Sublingual (sŭb-lĭng′gwăl): Beneath the tongue.
Synergism (sĭn′ĕr-jĭzm): Coordinated action between two or more agents so that the combined action is greater than the sum of each acting separately.
Translocate (trănz-lō′kāt): To move from one region to another.
Tremor (trĕm′or): Quivering or involuntary movement of the muscles.
Up-regulate (ŭp-rĕg′yŭ-lāt): To increase the number of receptors on a cell membrane.
Vasoactive (vās″ō-ăk′tĭv): Causing constriction of blood vessels.

Bibliography

Boushey, H. A., Drugs Used in Asthma, in *Basic and Clinical Pharmacology,* 8th ed., Katzung, B. G., Ed., McGraw-Hill, New York, 2001, chap. 20.
Farwell, A. P., and Braverman, L. E., Thyroid and Antithyroid Drugs, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 56.
Fitzgerald, P. A., Hypothalamic and Pituitary Hormones, in *Basic and Clinical Pharmacology,* 8th ed., Katzung, B. G., Ed., McGraw-Hill, New York, 2001, chap. 37.
Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Friedman, P. A., Agents Affecting Mineral Ion Homeostasis and Bone Turnover, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 61.

Golan, D. E., Tashjian, A. H., Armstrong, E. J., Galanter, J. M., Armstrong, A. W., Arnout, R. A., and Rose, H., Integrative Inflammation Pharmacology: Asthma, in *Principles of Pharmacology: The Pathophysiologic Basis of Drug Therapy,* Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 45.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., Elsevier/ Saunders, Philadelphia, PA, 2006.

Iversen, S., Iversen, L., and Saper, C. B., The Autonomic Nervous System and the Hypothalamus, in *Principles of Neural Science*, 4th ed., Kandel, E. R., Schwartz, J. H., and Jessell, T. M., Eds., McGraw-Hill, New York, 2000, chap. 49.

Jackson, E. K., Vasopressin and Other Agents Affecting the Renal Conservation of Water, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 29.

Kettyle, W. M., and Arky, R. A., *Endocrine Pathophysiology*, Lippincott-Raven, Philadelphia, PA, 1998.

Matfin, G., Kuenzi, J. A., and Guven, S., Disorders of Endocrine Control of Growth and Metabolism, in *Pathophysiology, Concepts of Altered Health States*, 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 42.

Page, C. P., Recent Advances in Our Understanding of the Use of Theophylline in the Treatment of Asthma. *J. Clin. Pharmacol.* 39: 237–240, 1999.

Parker, K. L., and Schimmer, B. P., Pituitary Hormones and Their Hypothalamic Releasing Factors, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 55.

Porterfield, S. P., *Endocrine Physiology*, 2nd ed., Mosby, St. Louis, MO, 2001.

Porth, C. M., *Pathophysiology, Concepts of Altered Health States*, 7th ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005.

Schimmer, B. P., and Parker, K. L., Adrenocorticotropic Hormone; Adrenocortical Steroids and Their Synthetic Analogs; Inhibitors of the Synthesis and Actions of Adrenocortical Hormones, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 59.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Undem, B. J., Pharmacotherapy of Asthma, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 27.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter eleven

The reproductive system

Study objectives

- Explain the process and the purpose of gametogenesis
- Describe the location and the function of each of the organs and structures in the male reproductive system including the testes, seminiferous tubules, testicular interstitial tissue, scrotum, epididymides, vas deferens, ejaculatory ducts, penis, prostate gland, seminal vesicles, and bulbourethral glands
- Distinguish between Sertoli cells and Leydig cells
- List the functions of testosterone in the male
- Describe the location and the function of each of the organs and structures in the female reproductive system including the ovaries, fallopian tubes, uterus, endometrium, myometrium, and the vagina
- Distinguish between granulosa cells and thecal cells
- Explain the events of the follicular phase of the ovarian cycle including changes in hormone (follicle-stimulating hormone [FSH], luteinizing hormone [LH], estrogen, and progesterone) secretion, follicular development, and the endometrium
- Explain the events of the luteal phase of the ovarian cycle including changes in hormone (FSH, LH, estrogen, and progesterone) secretion, corpus luteum development, and the endometrium
- Describe how FSH, LH, estrogen, and progesterone regulate the ovarian cycle
- List the other functions of the female sex hormones

11.1 Introduction

Unlike each of the other organ systems discussed in this textbook, the reproductive system does not contribute to the maintenance of homeostasis. However, it is essential for producing *offspring.* Reproduction is the process of passing genetic material from one generation to the next. As such, it involves *gametogenesis*, or the production of sex cells (*gametes*). The male gamete, or *sperm*, contains 23 chromosomes, and the female gamete, or *ovum*, also contains 23 chromosomes. The fertilization of the ovum by the sperm results in the formation of a *zygote* with the full human complement of 46 chromosomes. This chapter will discuss the following:

- Gametogenesis (spermatogenesis and oogenesis)
- Male reproductive system
- Testosterone
- Female reproductive system
- Ovarian cycle

11.2 Gametogenesis

11.2.1 Spermatogenesis

Germ cells in the male are referred to as *spermatogonia*. At puberty, these germ cells, which are found in the *Sertoli cells* of the testes, divide. Some of the cells reproduce themselves by way of *mitosis*. The spermatogonia may also undergo *meiosis* and become *primary spermatocytes*. Meiotic division by each primary spermatocyte eventually leads to the formation of four *spermatozoa*, or *sperm*. The entire development process takes about 64 days.

11.2.2 Oogenesis

Germ cells in the female, which are found in the ovaries, are referred to as *oogonia*. These cells have completed mitotic replication and begin the first stage of meiosis by replicating their DNA to form *primary oocytes* during fetal development. At birth, each ovary holds about one million of these oocytes. Each oocyte is contained within its own hollow ball referred to as an *ovarian follicle*.

Meiotic division resumes at puberty, at which point the number of oocytes and follicles has been reduced to about 400,000. The meiotic process produces two to three *polar bodies* that disintegrate and one egg, or ovum. The mature egg is released from the ovary by the process of *ovulation*, which is discussed in Section 11.6. Only about 400 of the oocytes will ovulate during the female's reproductive years. The rest will die by way of apoptosis. Oogenesis ceases entirely when the female reaches *menopause*, or the period that marks the permanent cessation of menstrual activity. Menopause typically occurs between the ages of 35 and 58 (95% by the age of 55).

11.3 Male reproductive system

The male reproductive system includes the following organs and structures:

- Testes
- Epididymides
- Vas deferens
- Ejaculatory ducts
- Penis
- Prostate gland

- Seminal vesicles
- Bulbourethral glands

11.3.1 Testes

The testes are paired, ovoid structures located in the *scrotum*. A component of the external *genetalia*, the scrotum is a sac-like structure into which the testes migrate during fetal development. One of the functions of the testes is spermatogenesis. The location of the testes outside of the abdominal cavity is necessary for the production of viable sperm that require a temperature 2°F to 3°F lower than the core body temperature.

The testes have two compartments: the *seminiferous tubules* and the *interstitial tissue*. The *seminiferous tubules* account for approximately 90% of the weight of the testes in the adult. Each tubule contains two types of cells:

1. *Spermatogonia* in various stages of development into sperm: The spermatogonia in different regions of the tubule are in different stages of development. In this way, sperm production is staggered and can remain relatively constant at a rate of 200 million sperm per day (about the number of sperm released in a single ejaculation).
2. *Sertoli cells*: Spermatogenesis in the Sertoli cells is stimulated by the hormones FSH and testosterone.

The *interstitial tissue* is found between the seminiferous tubules. *Leydig cells,* which secrete testosterone, are located here. Testosterone secretion by the Leydig cells is stimulated by the hormone LH.

11.3.2 Epididymides

Spermatozoa are drained from each testis by the *epididymis* (plural, epididymides). At this point, the sperm are nonmotile. These gametes undergo maturational changes as they pass through the epididymis such that they become more resistant to changes in pH and temperature. The epididymis also serves as a storage site for sperm between ejaculations.

11.3.3 Vas deferens

Spermatozoa are drained from the epididymis by the *vas deferens* (also referred to as the *ductus deferens*). This structure transports the sperm out of the scrotum and into the pelvic cavity.

11.3.4 Ejaculatory ducts

Spermatozoa are drained from the vas deferens by the *ejaculatory ducts*. These ducts enter the *prostate* and quickly merge with the *urethra*.

11.3.5 Penis

The urethra, which is a common passageway for urine and sperm, is found in the *penis*. *Emission* refers to the movement of semen into the urethra, and *ejaculation* refers to the forcible expulsion of the semen out of the penis. Emission and ejaculation occur as the result of peristaltic contractions of the tubular system, contractions of the seminal vesicles and prostate, and contractions of muscles at the base of the penis. These actions are the result of sympathetic nerve stimulation. *Erection*, which involves increased blood flow into the penis, is mediated by parasympathetic nerve activity and the release of nitric oxide.

Semen consists of the sperm as well as fluid (99%) secreted from the three male accessory sex organs: prostate gland, seminal vesicles, and bulbourethral glands.

PHARMACY APPLICATION: ERECTILE DYSFUNCTION

Erectile dysfunction (ED) is defined by the U.S. National Institutes of Health (NIH) Consensus Development Panel on Impotence as "the inability to achieve and maintain an erection sufficient to permit satisfactory sexual intercourse." Although ED is more common in older males, it is not a natural part of the aging process. In fact, 20% to 46% of men aged 40 to 69 have self-reported moderate or complete erectile dysfunction. Until the late 1990s, there were no truly effective oral medications for ED. More recently, a class of drugs referred to as phosphodiesterase-5 (PDE-5) inhibitors have been developed for the treatment of ED. These medications include sildenafil (Viagra®), vardenafil (Levitra®), and tadalafil (Cialis®). Differences between these drugs involve the onset and duration of their therapeutic effects.

Recall from the previous discussion that erection results from increased blood flow into the penis and is mediated by parasympathetic nerve activity and the release of nitric oxide. Nitric oxide causes vasodilation by increasing the formation of cyclic guanasine monophosphate (cGMP) within vascular smooth muscle cells. Vasodilation causes an increase in blood flow, engorgement, swelling, and stiffening of the penis or, in other words, an erection. The mechanism of action of the PDE-5 inhibitors involves blocking the breakdown of cGMP. This effect enhances the accumulation of cGMP and, therefore, improves erectile function.

11.3.6 Prostate

The doughnut-shaped prostate gland surrounds the neck of the bladder and the urethra. Its ducts open into the prostatic portion of the urethra. The

slightly alkaline fluid produced by the prostate during ejaculation neutralizes the pH of the vagina so that the sperm are fully motile and become capable of fertilizing an ovum.

11.3.7 Seminal vesicles

These large sac-like glands are located below the urinary bladder. The duct from each gland merges with the vas deferens to form the ejaculatory duct. The fluid produced by the seminal vesicles, which accounts for about 60% of the volume of the semen, is rich in fructose and nourishes the sperm.

11.3.8 Bulbourethral glands

The small bulbourethral glands are located below the prostate. The ducts of these glands drain into the urethra just after it leaves the prostate. The fluid produced by the bulbourethral glands contains mucus, which acts as a lubricant, and buffers, which assist in neutralizing the acidic environment of the vagina.

11.4 Testosterone

Testosterone is produced primarily by the Leydig cells of the testes. A small amount (5%) is secreted by the adrenal cortex. In many target cells, testosterone must be converted to *dihydrotestosterone* (DHT) in order to be effective. This derivative is more potent than testosterone.

Testosterone secretion may begin as early as 8 weeks after conception. Secretion reaches a peak at 12 to 14 weeks and then declines to very low levels by about 21 weeks. Testosterone secretion during fetal development plays a very important role in sex differentiation and the masculinizing of the embryonic structures. Secretion of testosterone begins again at puberty and is necessary for spermatogenesis, the development of secondary sex characteristics, and several anabolic effects. Secretion declines gradually and to varying degrees in males over 50 years of age. The mechanism of this decline is not known. A summary of the actions of testosterone is found in Table 11.1.

11.5 Female reproductive system

The female reproductive system includes the following organs and structures:

- Ovaries
- Fallopian tubes
- Uterus
- Vagina

Table 11.1 Actions of Testosterone in the Male

Sex Differentiation

- Development of the fetal *Wolffian ducts* into the epididymides, vas deferens, seminal vesicles, and ejaculatory ducts
- Development of the fetal urogenital sinus into the prostate gland
- Development of the external genitalia (fetal genital tubercle into the penis; fetal labioscrotal swellings into the scrotum)

Initiation and Maintenance of Spermatogenesis in the Sertoli Cells

Secondary Sex Characteristics

- Growth and maintenance of accessory sex organs
- Growth of penis
- Growth of facial and axillary hair
- Body growth
- Inhibition of action of estrogen on breast growth
- Stimulation of sex drive and enhancement of aggressive behavior

Anabolic Effects

- Protein synthesis and muscle growth
- Growth of bones and other organs and tissues (including the larynx)
- Cessation of bone growth
- Secretion of erythropoietin from the kidneys

Miscellaneous

- Inhibition of gonadotropin releasing hormone (GnRH) from the hypothalamus
- Inhibition of LH from the adenohypophysis
- Maintenance of relatively constant testosterone secretion from the testes

11.5.1 Ovaries

The ovaries are paired, almond-sized organs in the upper left and right regions of the pelvic cavity. As with the testes in the male, these female gonads have two primary functions: the production of gametes (ova) and the secretion of sex hormones (estrogen, progesterone).

11.5.2 Fallopian tubes

Ova released from ovarian follicles are transported to the uterus by way of the *fallopian tubes*. Also referred to as *oviducts*, these tubes are 20 to 25 cm in length and about the diameter of a typical pencil. Ciliated epithelial cells lining the fallopian tubes assist with the movement of the ova toward the uterus.

11.5.3 Uterus

This muscular, hollow, pear-shaped organ is located in the center of the pelvic cavity. The uterus is responsible for maintaining the fetus during gestation

and for expelling it at the end of the pregnancy. Prior to the occurrence of a pregnancy, the uterus is relatively small and is about 3 inches long, 2 inches wide, and 1 inch thick. The uterus has three layers:

1. Perimetrium
2. Myometrium
3. Endometrium

The *perimetrium* is the outermost layer of the uterus. This serous layer of connective tissue is continuous with the visceral peritoneum that covers the abdominopelvic organs.

The *myometrium* is the thick middle layer of smooth muscle in the uterus. Contraction of this muscle is responsible for *parturition*, or delivery of the fetus.

The *endometrium* is the innermost layer of the uterus. It is highly vascularized and nutrient rich. Endometrial cells accumulate lipids and glycogen which provide nourishment to a developing embryo prior to the formation of the placenta. This layer includes a stratified squamous epithelium that lines the uterus and alternately proliferates and sloughs off. This layer also includes a basal layer of connective tissue where the blood vessels and *endometrial glands* are located. The secretions from these glands are necessary for the implantation of the embryo in the uterine wall. Furthermore, these glands are a source of nutrients and growth factors during the first trimester of human pregnancy. Support of the fetus is later provided by the placenta which is not fully established until 10 to 12 weeks of pregnancy.

11.5.4 Vagina

This muscular, expandable tube connects the uterus to the external surface of the body. The vagina is lined with stratified squamous epithelium that is resistant to bacterial colonization. This structure serves as a passageway for the insertion of the penis, the reception of semen, the discharge of menstrual flow, and the delivery of the fetus.

11.6 Ovarian cycle

The *ovarian cycle* alternates between two phases: the *follicular phase* and the *luteal phase*. The average cycle lasts about 28 days. However, the length of the ovarian cycle varies among women and among cycles in any particular woman.

11.6.1 Follicular phase

This phase includes the development of the primary follicle into a mature, or *graafian*, follicle as well as ovulation, or the release of the ovum from the follicle. The steps involved in the follicular phase are as follows:

- *Primary oocytes* (46 chromosomes) are contained within *primary follicles* (40 µm): This follicle consists of a single layer of follicular cells, specifically, *granulosa cells*.
- *Development of the secondary follicle*:
  - *Follicular cells proliferate*: Granulosa cells proliferate forming several layers of cells that surround the primary oocyte.
  - *Formation of the zona pellucida*: Granulosa cells secrete a thick gel-like layer of proteins and polysaccharides that lies between the inner layer of these cells and the primary oocyte. Cytoplasmic processes of the inner layer of granulosa cells penetrate the zona pellucida and form gap junctions with the oocyte. Nutrients and chemical messengers (e.g., factors that maintain meiotic arrest such as *oocyte maturation-inhibiting substance*) are delivered to the primary oocyte through these passageways.
  - *Formation of thecal cells*: Concurrent with granulosa cell proliferation is the differentiation of surrounding ovarian connective tissue into thecal cells. The follicular cells of the secondary follicle, which now includes granulosa cells and thecal cells, secrete estrogen.
  - *Formation of the antrum*: A fluid-filled space referred to as the antrum begins to form within the granulosa layer. This fluid is derived from the granulosa cells and begins to accumulate some of the estrogen secreted by these cells. As the antrum increases in size, so does the follicle which reaches a diameter of 12 to 16 mm in the mature follicle.
- *Formation of the mature follicle*: One primary follicle will typically grow more rapidly than other follicles and develop into the mature follicle within approximately 14 days after the onset of follicular development. This follicle is dominated by the antrum. The enclosed oocyte completes the first stage of meiotic division so that it is now a *secondary oocyte* (23 chromosomes). (The second meiotic division is completed after ovulation only if fertilization has occurred.)
- *Ovulation*: The mature follicle bulges on the surface of the ovary. The follicular cells release enzymes (e.g., collagenase) that digest the thin layer of connective tissue in the wall of the ovary which significantly weakens it. Following the LH surge and the rupture of the follicle, the ovum is swept into the fallopian tube by the antral fluid. Any other follicles that had been developing simultaneously in the ovary will undergo *atresia*, or degeneration. If the released ovum is not fertilized within a few days of its release, it will also degenerate.

11.6.2 Luteal phase

The follicular cells left behind in the ruptured follicle are transformed into a *corpus luteum* (*corpus*, body; *luteum*, yellow). It is so named for the abundant storage of cholesterol within the granulosa cells which will be used to form

sex hormones. Copious quantities of progesterone and smaller amounts of estrogen are secreted from the corpus luteum. If fertilization and implantation occur, then the corpus luteum grows larger and continues to secrete the hormones necessary to maintain the pregnancy until the placenta is fully developed. If fertilization and implantation do not occur, then the corpus luteum will degenerate within 14 days of its formation.

11.6.3 Hormonal regulation of the ovarian cycle

The ovarian cycle is regulated by the complex interaction of hypothalamic, pituitary, and ovarian hormones:

- *GnRH*: Gonadotropin releasing hormone; hypothalamus
- *FSH and LH*: Follicle-stimulating hormone and luteinizing hormone (gonadotropins); adenohypophysis
- *Estrogen and progesterone*: Ovary

The secretion of these hormones and their effects during the cycle are summarized in Figure 11.1, Table 11.2, and in the following discussion. Day 1 of the ovarian cycle is the first day of *menstruation*, or the bleeding caused by the sloughing off of the lining of the uterus. This day was chosen because the menstrual discharge provides an easily observed sign.

During the *follicular phase*, hormonal secretion and actions include the following:

- *Gonadotropin* secretion increases during the end of the previous luteal phase and the beginning of the new follicular phase. *FSH*, in particular, stimulates the development of several follicles within the ovary.
- As the primary follicle grows, *FSH* and *LH* stimulate the secretion of estrogen from the granulosa cells and the thecal cells, respectively.
- *Estrogen* secreted from the follicle during the follicular phase:
  - Inhibits GnRH secretion from the hypothalamus and FSH and LH secretion from the adenohypophysis, which inhibits the development of new follicles and ovulation during the follicular phase of that particular cycle.
  - Enhances estrogen secretion from the granulosa cells, which allows hormone secretion to continue even as levels of FSH decrease.
  - Causes the endometrium to thicken, or the *uterine proliferative phase*.
  - Induces the production of progesterone receptors in the endometrium.
- Estrogen secretion peaks near the end of the follicular phase. The high levels of estrogen at this point elicit a surge in LH production at mid-cycle. Ovulation occurs about 16 to 24 hours after the LH surge.

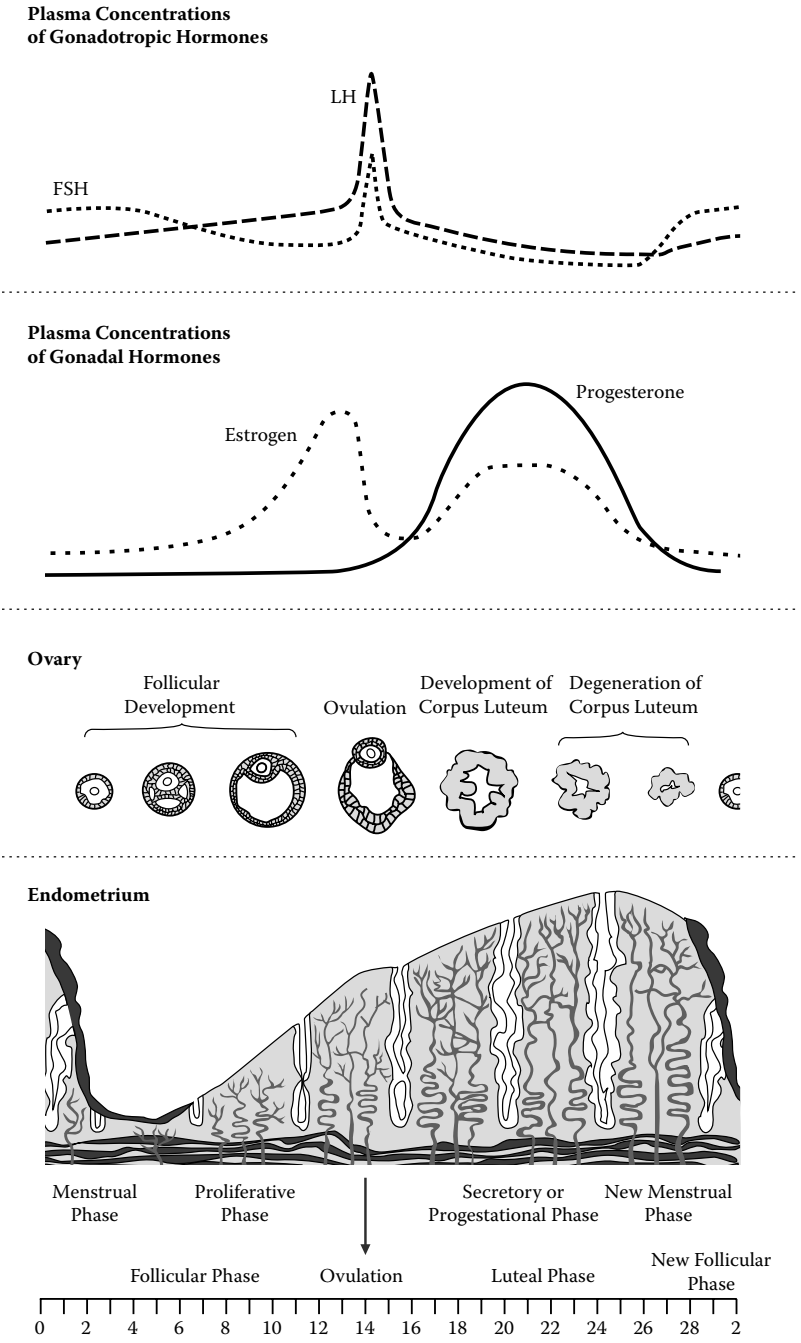


Figure 11.1 The correlation between hormone levels and changes in the ovary and the endometrium during the ovarian cycle.

Table 11.2 Actions of Sex Hormones in the Female

GnRH, Gonadotropin-Releasing Hormone

- Stimulates the secretion of FSH and LH from the adenohypophysis

FSH, Follicle-Stimulating Hormone

- Stimulates the development of ovarian follicles
- Stimulates the secretion of estrogen from the granulosa cells of the primary follicle

LH, Luteinizing Hormone

- Stimulates thecal cells to secrete androgen that is converted to estrogen in the granulosa cells
- LH surge
  - Inhibits estrogen synthesis by follicular cells
  - Reinitiates meiosis in the oocyte by inhibiting the release of oocyte maturation-inhibiting substance
  - Elicits the production of prostaglandins that induce ovulation
  - Causes differentiation of follicular cells into luteal cells
  - Stimulates steroid hormone secretion from the corpus luteum

Estrogen, during the Follicular Phase of the Ovarian Cycle

- Inhibits the secretion of GnRH, FSH, and LH
- Enhances its own secretion from the granulosa cells
- Causes the endometrium to thicken
- Induces the production of progesterone receptors in the myometrium
- Triggers the surge in LH secretion at mid-cycle

Estrogen, during Pregnancy

- Stimulates the growth of the myometrium and increases uterine strength which is necessary for parturition
- Helps to prepare the mammary glands for lactation following parturition by promoting the development of the ducts through which milk will be ejected
- Inhibits the effects of prolactin during the last half of pregnancy and, thereby, prevents milk secretion prior to parturition

Estrogen, Days before Parturition

- Promotes the formation of gap junctions between uterine smooth muscle cells so that the myometrium may contract in a coordinated fashion during parturition
- Markedly increases the number of receptors for oxytocin in uterine smooth muscle cells

Estrogen, Miscellaneous

- Stimulates the growth and maintenance of the entire reproductive tract
- Enhances the transport of sperm to the fallopian tubes by stimulating upward contractions of the uterus

continued

Table 11.2 (continued) Actions of Sex Hormones in the Female

Estrogen, Miscellaneous (continued)
• Promotes fat deposition
• Increases bone density
• Closes epiphyseal plates
Progesterone
• Inhibits the secretion of GnRH, FSH, and LH during the luteal phase of the ovarian cycle
• Elicits the secretory phase in the endometrium and provides a suitable, nurturing environment for an implanted embryo
• Promotes the formation of a mucus plug in the cervix
• Stimulates breast development during pregnancy
• Inhibits the effects of prolactin during pregnancy
• Prevents miscarriage by inhibiting uterine contractions during pregnancy

During the *luteal phase*, hormonal secretion and actions include the following:

- *LH* promotes:
  - The production of locally acting prostaglandins that cause the rupture of the follicle.
  - Transformation of granulosa cells and thecal cells into the luteal cells of the corpus luteum, a process referred to as *luteinization*.
- The corpus luteum secretes abundant *progesterone* as well as some *estrogen*. Although low levels of estrogen tend to stimulate the secretion of GnRH from the hypothalamus and FSH and LH from the adenohypophysis, these effects are overwhelmed by the actions of progesterone which inhibit the secretion of these hormones. As a result, new follicular maturation and ovulation during the luteal phase of that particular cycle are prevented.
- *Progesterone* from the corpus luteum acts on the endometrium to produce vascular and secretory changes that will provide a suitable and nurturing environment for an implanted embryo. This is referred to as the *secretory*, or *progestational*, *phase*. Specifically, progesterone:
  - Stimulates blood vessel growth in the connective tissue layer.
  - Elicits endometrial gland growth and coiling.
  - Stimulates endometrial cells to accumulate lipids and glycogen within their cytoplasm.
  - Causes cervical mucus to thicken, forming a plug that blocks the opening of the uterus and prevents the admission of bacteria and sperm.

- Degeneration of the corpus luteum causes *progesterone* and *estrogen* secretion to decrease. As a result:
  - Inhibition of GnRH, FSH, and LH secretion is removed, which leads to a new follicular phase.
  - Maintenance of a secretory endometrium ceases. Blood vessels in the surface layer of the endometrium constrict. The resulting decrease in the delivery of oxygen and nutrients causes the cells of this layer to die. About two days after the corpus luteum ceases to function, the surface layer of the endometrium sloughs off and menstruation begins. Menstrual discharge consists of about 80 ml of blood, serous fluid, and cellular debris. It typically lasts 3 to 7 days, well into the follicular phase of the next ovarian cycle.

PHARMACY APPLICATION: ORAL CONTRACEPTIVES

Oral contraceptives are among the most widely used class of drugs in the United States and throughout the world. They are convenient, accessible, and reliable. The most commonly used agents contain a combination of both estrogen-like and progestin-like steroids. The preparations come in 28-day packs. The first 21 pills contain the active steroids and the last 7 days contain only inert ingredients.

The mechanism of action of the steroids involves the inhibition of GnRH, FSH, and LH secretion. As a result, follicular development and ovulation do not occur. The endometrium thickens and develops secretory capacity as it would normally do under the influence of endogenously produced estrogen and progesterone. Then during the last 7 days when the steroids are withdrawn, the endometrium sloughs, and menstruation occurs as it would normally upon degeneration of the corpus luteum.

Medical terminology

Antrum (ăn′trŭm): In the reproductive system, the fluid-filled cavity of a developing ovarian follicle.

Atresia (ă-trē′zē-ă): In the reproductive system, the normal death of an ovarian follicle.

Corpus luteum (kor′pŭs lū′tŭm): Yellow structure that forms within a ruptured ovarian follicle and secretes progesterone and estrogen.

Ejaculation (ē-jăk″ū-lā′shŭn): Ejection of semen from the male urethra.

Emission (ē-mĭsh′ŭn): Movement of semen into the urethra.

Endometrium (ĕn-dō-mē′trē-ŭm): Lining of the uterus where implantation of the embryo takes place.

Erection (ē-rĕk′shŭn): State of swelling and stiffness of the penis due to its engorgement with blood.

Estrogen (ĕs′trō-jĕn): Female sex hormone.
Fertilization (fĕr-tĭl-ĭ-zā′shŭn): Fusion of the sperm with an ovum.
Gamete (găm′ēt): Mature male (sperm) or female (ovum) reproductive cell.
Gametogenesis (găm″ĕt-ō-jĕn′ĕ-sĭs): Formation of gametes.
Genitalia (jĕn-ĭ-tāl′ē-ă): Reproductive organs.
Graafian follicle (grăf′ē-ăn): Mature ovarian follicle.
Granulosa cell (grăn″ū-lōsă): Estrogen-producing cell within the ovarian follicle.
Implantation (ĭm″plăn-tā′shŭn): Embedding of the embryo within the endometrium.
Leydig cell (lī′dĭg): Testosterone-producing cell of the testes.
Luteinization (lū″tē-ĭn-ī-zā′shŭn): Formation of the corpus luteum within a ruptured follicle.
Meiosis (mī-ō′sĭs): Process where oogonia or spermatogonia undergo cell division to produce the gametes, ova, and sperm.
Menopause (mĕn′ō-pawz): Permanent cessation of menstruation.
Menstruation (mĕn-stroo-ā′shŭn): Sloughing of the endometrium.
Mitosis (mī-tō′sĭs): Process where cells divide in order to reproduce themselves.
Myometrium (mī″ō-mē′trē-ūm): Smooth muscle layer of the uterine wall.
Oogenesis (ō″ō-jĕn′ĕ-sĭs): Formation of a mature ovum.
Oogonium (ō″ō-gō′nē-ŭm): Stem cell from which an oocyte originates.
Ovulation (ŏv″ū-lā′shŭn): Rupture of the follicle and discharge of the ovum.
Parturition (păr-tū-rĭsh′ŭn): Process of giving birth to offspring.
Placenta (plă-sĕn′tă): Uterine structure that provides nourishment and oxygen to a developing fetus.
Progesterone (prō-jĕs′tĕr-ōn): Female sex hormone.
Puberty (pū′bĕr-tē): Stage in life when an individual becomes capable of reproduction.
Scrotum (skrō′tŭm): Pouch that contains the testes external to the pelvic cavity.
Semen (sē′mĕn): Fluid discharge from the male containing sperm.
Sertoli cell (sĕr-tō′lē): Sperm-producing cell of the testes.
Spermatogenesis (spĕr″măt-ō-jĕn′ĕ-sĭs): Formation of mature sperm.
Spermatogonium (spĕr″măt-ō-gō′nē-ŭm): Stem cell from which sperm originate.
Spermatozoa (spĕr″măt-ō-zō′ă): Sperm.
Testosterone (tĕs-tŏs′tĕr-ōn): Male sex hormone.
Thecal cell (thē′kăl): Androgen-producing cell in the ovarian follicle.
Zona pellucida (zō′nă pĕl-lū′sĭă): Thick layer surrounding the oocyte.
Zygote (zī′gōt): Fertilized ovum.

Bibliography

Ansong, K. S., Lewis, C., Jenkins, P., and Bell, J., Epidemiology of erectile dysfunction: a community-based study in rural New York state. *Annals of Epidemiology,* 10: 293–296, 2000.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, New York, 2006.

Gray, C. A., Taylor, K. M., Ramsey, W. S., Hill, J. R., Bazer, F. W., Bartol, F. F., and Spencer, T. E., Endometrial glands are required for preimplantation conceptus, elongation and survival, *Biology of Reproduction,* 64(6): 1608–1613, 2001.

Hempstock, J., Cindrova-Davies, T., Jauniaux, E., and Burton, G., Endometrial glands as a source of nutrients, growth factors and cytokines during the first trimester of human pregnancy: A morphological and immunohistochemical study, *Reproductive Biology and Endocrinology,* 2: 58, 2004.

Kantor, J., Bilker, W. B., Glasser, D. B., and Margolis, D. J., Prevalence of erectile dysfunction and active depression: An analytic cross-sectional study of general medical patients. *American Journal of Epidemiology,* 156: 1035–1042, 2002.

Loose, D. S., and Stancel, G. M., Estrogens and Progestins, in *Goodman and Gilman's The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 57.

Matfin, G., Disorders of the Male Genitourinary System, in *Pathophysiology, Concepts of Altered Health and Disease*, 7th ed., Porth, C. M., Ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.

Michel, T., Treatment of Myocardial Ischemia, in *Goodman and Gilman's The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 31.

NIH Consensus Development Panel on Impotence, NIH Consensus Conference: Impotence, *JAMA,* 270: 83-90, 1993.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2001.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

www.nlm.nih.gov/medlineplus/erectiledysfunction.html.

chapter twelve

Skeletal muscle

Study objectives

- Discuss the functions of skeletal muscle
- Distinguish between origin and insertion; flexion and extension; and agonist, antagonist, and synergist muscles
- Distinguish between isometric and isotonic contractions
- Distinguish between concentric and eccentric contractions
- Describe the components of the thick filaments and the thin filaments
- Explain the functions of the following: myosin cross-bridges, troponin, tropomyosin, sarcomeres, Z lines, neuromuscular junction, transverse tubules, and sarcoplasmic reticulum
- Describe the Sliding Filament Theory of skeletal muscle contraction
- Explain how creatine phosphate, oxidative phosphorylation, and glycolysis provide energy for skeletal muscle contraction
- List the factors that influence the onset of muscle fatigue
- Describe the factors that lead to the development of muscle fatigue
- Describe the metabolic processes that lead to oxygen debt
- Distinguish between the three types of muscle fibers: slow-twitch oxidative, fast-twitch oxidative, and fast-twitch glycolytic
- Explain how the percentage of a muscle fiber type in a given muscle is determined
- Describe the factors that influence the strength of skeletal muscle contraction, including multiple motor unit summation, asynchronous motor unit summation, frequency of nerve stimulation, the length–tension relationship, and the diameter of the muscle fiber

12.1 Introduction

Skeletal muscle composes the largest group of tissues in the human body and accounts for up to 40% of the total body weight. This type of muscle, which is innervated by the somatic nervous system, is under voluntary control. Skeletal muscle performs many important functions in the body, including:

- Movement of body parts
- Heat production
- Respiration
- Vocalization

Most skeletal muscles are attached to bones, which enables them to *control body movements*, such as walking, making facial expressions, chewing, and swallowing. These muscles are also responsible for the manipulation of objects, such as writing with a pencil or eating with a fork. Furthermore, movement of the eyes is carried out by several pairs of skeletal muscles. Finally, the contractions of certain groups of muscles, referred to as "antigravity" muscles, are needed to maintain posture and provide body support.

Only about 20% to 30% of the nutrient energy consumed during skeletal muscle activity is actually converted into purposeful work. The remaining 70% to 80% of the nutrient energy is given off as *heat*. Therefore, because of its large mass, skeletal muscle is the tissue most responsible for maintaining and increasing body temperature.

Although it typically occurs subconsciously, *breathing* is a voluntary activity. The diaphragm and the other muscles of inspiration and expiration are skeletal muscles. As such, breathing can be voluntarily controlled to some extent (see Chapter 19). Finally, *speaking* and other forms of vocalization depend upon the coordinated contraction of skeletal muscles.

12.2 Muscle tension and movement

When muscle contracts, it develops *tension*. This tension is what enables the muscle to perform work such as body movement. Skeletal muscle is typically attached by way of *tendons* to at least two different bones across a joint. The proximal end of the muscle that is attached to a relatively stationary part of the skeleton is referred to as the *origin*. The opposite end of the muscle which is attached to a part of the skeleton that moves more freely is referred to as the *insertion*. When the muscle develops tension and shortens, it pulls the insertion toward the origin.

Movement that decreases the angle of a joint, or bends the joint, and brings the bones toward each other is referred to as *flexion*. Conversely, movement that increases the angle of the joint and straightens the joint is referred to as *extension*. For example, the biceps brachii muscle is located on the anterior surface of the arm. Its origins are found on two portions of the scapula (shoulder), and its insertion is found on the radius (one of the bones of the forearm). This muscle crosses the elbow joint. When it develops tension and shortens, it *flexes* the forearm and pulls it toward the shoulder. In contrast, the triceps brachii muscle is located on the posterior surface of the arm. Its origins are found on the scapula and the upper portion of the humerus (arm), and its insertion is found on the ulna (the other bone of the forearm). This muscle also crosses the elbow joint. However, when it develops tension and shortens, it *extends* the forearm and straightens the elbow joint. The triceps brachii muscle that causes a movement opposite of that of the biceps brachii muscle is referred to as the *antagonist* muscle. The *agonist* muscle, or *prime mover*, provides the force for a specific movement. In this case, the agonist muscle is the biceps brachii. *Synergist muscles* work *with* the prime movers

to achieve the movement. In this case, the synergist muscle is the brachialis muscle of the arm.

12.3 Isometric versus isotonic contraction

There are two primary types of muscle contraction: *isometric contraction* and *isotonic contraction*.

Isometric contraction occurs when the muscle develops tension and exerts force on an object but does not shorten. In other words, it refers to muscle contraction during which the length of the muscle remains constant. For example, supporting an object in a fixed position, such as carrying a book or a backpack, requires isometric contraction. This type of contraction also occurs when attempting to move an object that is too heavy to shift or reposition. In this case, the muscle may exert maximal force against the object; however, because the object does not move, the length of the contracting muscle does not change. Finally, the antigravity muscles of the back and the legs perform submaximal isometric contractions while maintaining posture and for body support.

Isotonic contraction occurs when the muscle changes length under a constant load. For example, when lifting an object, the muscle contracts and becomes shorter while the weight of the object remains constant. Because shortening of the muscle occurs, this is referred to as *concentric contraction*. When placing the object back down, once again, the muscle is generating tension. However, in this case, the muscle is lengthening. This is referred to as *eccentric contraction*. In addition to moving external objects, isotonic contractions are performed for movement of the body, such as moving the legs when walking.

Many activities require both types of contractions by the muscles. An example is running. When one of the legs hits the ground, isometric contraction of the muscles within this limb keeps it stiff and helps to maintain body support. At the same time, isotonic contractions in the opposite leg move it forward to take the next stride.

12.4 Structure of skeletal muscle

A whole muscle is composed of muscle cells, or *muscle fibers*. Muscle fibers are elongated, cylindrical cells. Due to the fusion of many smaller fibers during embryonic development, muscle fibers are the largest cells in the body with several nuclei near their surface. Muscle fibers lie parallel to each other and extend along the entire length of the muscle. These fibers may be a few millimeters in length (muscles of the eyes) or up to two or more feet in length (muscles of the legs).

Muscle fibers are incapable of mitosis. In fact, the number of muscle fibers per muscle is likely determined by the second trimester of fetal development. Therefore, enlargement of a whole muscle is not due to an increase in the

number of fibers in the muscle. Instead, it is due to the hypertrophy of the existing fibers.

There are no gap junctions between muscle fibers so that electrical activity cannot spread from one cell to the next. Therefore, each muscle fiber must be innervated by a branch of an alpha motor neuron. A *motor unit* is defined as an alpha motor neuron and all of the muscle fibers it innervates.

Internally, muscle fibers are highly organized. Each fiber contains numerous *myofibrils*. These cylindrical structures also lie parallel to the long axis of the muscle. The myofibrils are composed of *thick filaments* and *thin filaments*. The arrangement of these filaments creates the alternating light and dark bands observed microscopically along the muscle fiber. Hence, skeletal muscle is also referred to as *striated muscle*.

12.4.1 Sarcomeres

The thick filaments and thin filaments are organized into repeating segments referred to as *sarcomeres*, which are the functional units of skeletal muscle. (In other words, the sarcomere is the smallest contractile unit within skeletal muscle.) A myofibril is composed of hundreds or thousands of sarcomeres in series along its length. When a muscle is stimulated, each of these sarcomeres contracts and becomes shorter. As a result, the entire muscle contracts and becomes shorter. Therefore, the function of the sarcomere determines whole muscle function.

A sarcomere is the area between two *Z lines* (see Figure 12.1c). The function of the Z line is to anchor the thin filaments in place at either end of the sarcomere. The thick filaments are found in the central region of the sarcomere. Ultimately, the interaction between the thick filaments and the thin filaments causes shortening of the sarcomere.

12.4.2 Thick filaments

Each thick filament contains 200 to 300 *myosin* molecules. Each myosin molecule is made up of two identical subunits that are shaped like golf clubs: two long shafts wound together with a *myosin head*, or *cross-bridge*, on the end of each of them. These molecules are arranged so that the shafts are bundled together and oriented toward the center of the thick filament. The myosin heads project outward from either end of the thick filament (see Figure 12.1a).

12.4.3 Thin filaments

The thin filaments are composed of three types of proteins:

1. Actin
2. Tropomyosin
3. Troponin

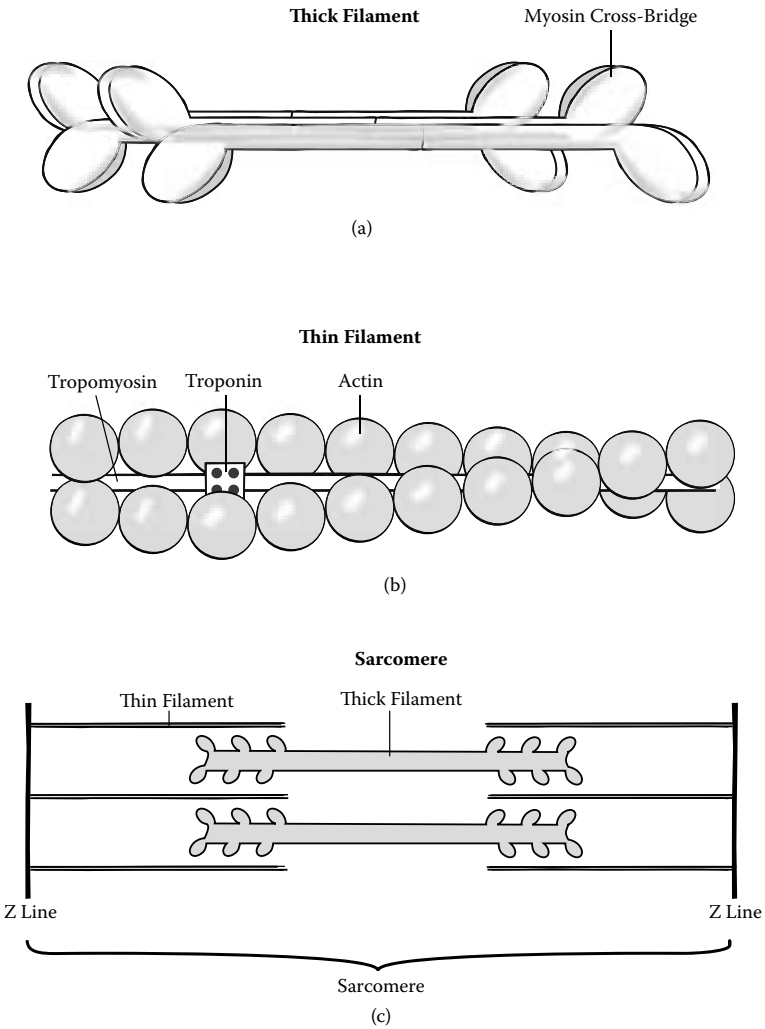


Figure 12.1 Components of the sarcomere. (a) Thick filament. The thick filament is composed of myosin molecules. These molecules are shaped like golf clubs and consist of a long shaft with a globular portion at one end. The myosin is arranged so that the shafts are in the center of the thick filament and the globular portions, or myosin cross-bridges, protrude from each end of the thick filament. The myosin cross-bridges bind to the actin of the thin filament. (b) Thin filament. The thin filament consists of three types of proteins. Globular actin molecules join together to form two strands of fibrous actin that twist around each other. Tropomyosin is a filamentous protein found on the surface of the actin, physically covering the binding sites for the myosin cross-bridges. Troponin molecules stabilize the tropomyosin filaments in position on the actin. (c) Sarcomere. The thick filaments and thin filaments are highly organized. They are arranged to form a sarcomere, which is the functional unit of skeletal muscle. The sarcomere is the region between two Z lines.

PHARMACY APPLICATION: MYASTHENIA GRAVIS

Myasthenia gravis is an autoimmune disorder that involves the neuromuscular junction. The peak incidence occurs in patients between 20 and 30 years of age, and it is approximately three times more common in women than it is in men.

Myasthenia gravis is characterized by progressive muscle weakness. It is caused by an antibody-mediated loss of acetylcholine receptors in the neuromuscular junction. As the number of receptors decreases, then it becomes less likely that the acetylcholine released from the alpha motor neuron will bind to and stimulate a functioning receptor. The decrease in receptor stimulation leads to a decrease in the electrical activity in the muscle fiber and weaker muscle contractions.

A common approach in the pharmacotherapy of myasthenia gravis is to increase the concentration of acetylcholine in the neuromuscular junction. As the concentration of neurotransmitter increases, then the likelihood that functioning receptors will be located and stimulated also increases. As a result, muscle function is improved. Drugs of choice for this method of treatment include pyridostigmine and neostigmine. These reversible anticholinesterase drugs inhibit the breakdown of acetylcholine.

The predominant protein, *actin*, consists of spherical subunits (globular actin) arranged into two chains twisted around each other (fibrous actin). (Imagine two strands of pearls twisted around each other.) *Tropomyosin* is a long, thread-like protein found on the outer surface of the actin chain. Each tropomyosin molecule is associated with 6 to 7 actin subunits. The function of tropomyosin is to cover the binding sites for myosin on the actin subunits when the muscle is in the resting state. This prevents the interaction between actin and myosin that causes muscle contraction. *Troponin* is a smaller protein consisting of three subunits. One subunit binds to actin, another binds to tropomyosin, and the third binds with calcium. When the muscle is relaxed, troponin holds the tropomyosin in its blocking position on the surface of the actin (see Figure 12.1b).

A summary of the structural organization in skeletal muscle is as follows:

Muscle → Muscle Fiber → Myofibril → Thick filaments (Myosin)
and Thin filaments (Actin, Troponin, Tropomyosin)

12.5 Neuromuscular junction

Each muscle fiber is innervated by a branch of an alpha motor neuron. The synapse between the somatic motor neuron and the muscle fiber is referred

to as the *neuromuscular junction*. Action potentials in the motor neuron cause the release of the neurotransmitter, *acetylcholine*. The binding of acetylcholine to its receptors on the muscle fiber causes an increase in the permeability to Na^+ ions and K^+ ions. The ensuing depolarization generates an action potential that travels along the surface of the muscle fiber in either direction. This is referred to as a *propagated action potential*. This action potential elicits the intracellular events that lead to muscle contraction.

12.6 Mechanism of contraction

As mentioned previously, skeletal muscle fibers are very large cells with a wide diameter. The action potential is readily propagated, or transmitted, along the surface of the muscle fiber. However, a mechanism is needed to transmit the electrical impulse into the central region of the muscle fiber as well. The *transverse tubules* (T tubules) are invaginations of the cell membrane that penetrate deep into the muscle fiber and surround each myofibril. (Imagine poking one's fingers into an inflated balloon.) As the action potential travels along the surface of the fiber, it is also transmitted into the T tubules. As a result, all regions of the muscle fiber are stimulated by the action potential.

All types of muscle require *calcium* for contraction. In skeletal muscle, Ca^{++} ions are stored within an extensive membranous network referred to as the *sarcoplasmic reticulum*. This network is found throughout the muscle fiber and surrounds each myofibril. Furthermore, segments of the sarcoplasmic reticulum lie adjacent to each T tubule. The T tubule with a segment of sarcoplasmic reticulum on either side of it is referred to as a *triad*. As the action potential is transmitted along the T tubule, it stimulates the release of Ca^{++} ions from the sarcoplasmic reticulum through the Ca^{++}-release channels. The sarcoplasmic reticulum is the only source of calcium in skeletal muscle.

The mechanism of skeletal muscle contraction is described by the *Sliding Filament Theory* (see Figure 12.2). This mechanism begins with the "*priming*" of the myosin cross-bridge, a process that requires energy. This energy is supplied by adenosine triphosphate (ATP). Each myosin cross-bridge contains an enzyme, *myosin ATPase*. When ATP attaches to its binding site on the myosin cross-bridge, it is split by the myosin ATPase to yield adenosine diphosphate (ADP) and inorganic phosphate (P_i). The ADP and P_i remain tightly bound to the myosin cross-bridge. The energy released by this process causes the myosin cross-bridge to swivel outward toward the end of the thick filament. When the myosin cross-bridge is in this conformation, it is "primed" and is referred to as the *high-energy form of myosin*. The high-energy form of myosin is capable of binding to actin. However, this interaction is prevented by tropomyosin, which physically covers the binding sites for myosin on the actin subunits. In order to uncover these binding sites, calcium is needed.

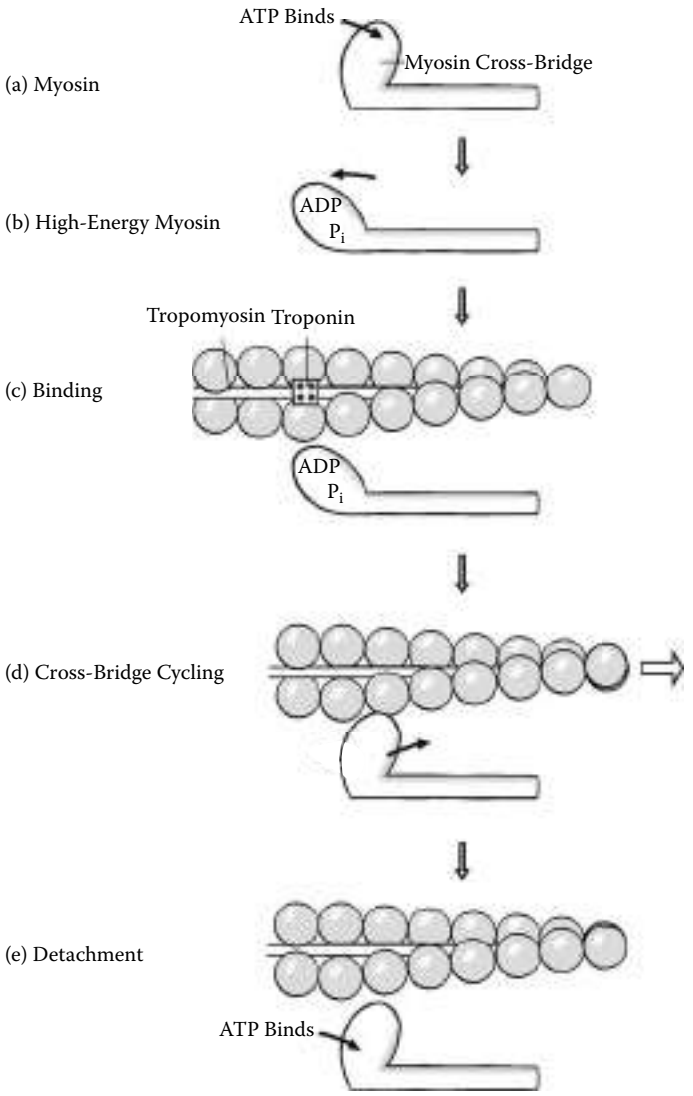


Figure 12.2 Mechanism of skeletal muscle contraction. (a) Myosin. The myosin cross-bridge has a binding site for ATP. (b) High-energy myosin. Within the cross-bridge, myosin ATPase splits ATP into ADP and P_i. As a result, the cross-bridge swivels outward and energy is stored. (c) Binding. In the presence of calcium, which binds to troponin, tropomyosin is repositioned into the groove between the two strands of actin. As a result, the binding sites for myosin on the actin are uncovered and the cross-bridges attach to actin. The ADP and P_i are released. (d) Cross-bridge cycling. The energy stored within the myosin cross-bridge is released and the cross-bridge swivels inward, pulling the actin inward. (e) Detachment. Binding of a new molecule of ATP to the myosin cross-bridge allows the myosin to detach from the actin and the process begins again.

In a stimulated muscle fiber, Ca^{++} ions are released from the sarcoplasmic reticulum. These Ca^{++} ions bind to troponin. As a result, the troponin–actin linkage is weakened allowing the troponin and, therefore, the tropomyosin to be repositioned. The troponin–tropomyosin complex moves away from the surface of the actin chains such that the myosin binding sites are uncovered. The primed myosin cross-bridges now bind to the actin. Binding to actin causes the energy previously stored within the myosin to be discharged, and the cross-bridge swivels inward toward the center of the thick filament. This process is referred to as *cross-bridge cycling* or the *power stroke*. As the myosin cross-bridge swivels inward, it pulls the actin inward as well. It is important to note that the interaction between the actin and the myosin causes the thin filaments to slide inward over the thick filaments toward the center of the sarcomere. Consequently, the sarcomeres shorten, and therefore, the whole muscle shortens or contracts. It is for this reason that this process is referred to as the sliding filament theory of muscle contraction.

When the myosin cross-bridge binds with the actin, the ADP and P_i are released from the myosin. This opens the binding site to another molecule of ATP. In fact, the myosin remains attached to the actin until another ATP molecule binds to the myosin. Binding of a new ATP causes the myosin to release the actin. This ATP is split by the ATPase, and the myosin cross-bridge swivels outward once again, returning the myosin to its high-energy state. As long as Ca^{++} ions are present and the binding sites on the actin are uncovered, cross-bridge cycling continues. The cross-bridges of the thick filament pull the thin filaments inward incrementally such that the sarcomeres become shorter and the muscle contracts further. The major events of muscle contraction are summarized in Table 12.1.

Interestingly, the myosin cross-bridges do not all cycle at the same time. At any given moment, some cross-bridges remain attached to the actin, and others are in the process of releasing the actin in order to cycle once again. In other words, myosin cross-bridge cycling is staggered. This process maintains the shortening of the sarcomere and prevents the thin filaments from slipping back to their original positions in between cycles.

In the absence of ATP, the myosin cross-bridges are unable to release the actin. As a result, the sarcomeres and, therefore, the muscle, remain contracted. This phenomenon is referred to as *rigor mortis*. Following death, the concentration of intracellular calcium increases. This calcium allows the contractile process between the previously formed high-energy myosin and the actin to take place. However, the muscle stores of ATP are rapidly depleted, the myosin remains attached to the actin, and stiffness ensues. Rigor mortis begins 3 to 4 hours after death and becomes complete in about 12 hours. It subsides during the next several days as the contractile proteins begin to degrade.

When the action potentials in the alpha motor neuron cease, stimulation of the muscle fiber is ended. The Ca^{++} ions are actively pumped back into the sarcoplasmic reticulum, and troponin and tropomyosin return to

Table 12.1 Major Events of Muscle Contraction

ATP binds to the myosin cross-bridge and is split by myosin ATPase.
Energy released by this reaction causes the myosin cross-bridge to swivel outward toward the end of the thick filament forming the "primed" or high-energy form of myosin which is capable of binding with actin.
Nerve impulse is generated by the alpha motor neuron.
Acetylcholine is released from the axon terminal into the neuromuscular junction and binds to receptors on the muscle fiber.
The resulting increased permeability to Na^+ ions and K^+ ions elicits an action potential in the muscle fiber.
The action potential is propagated along the surface of the muscle fiber as well as into the transverse tubules.
The action potential in the transverse tubules triggers the release of calcium from the sarcoplasmic reticulum.
Calcium binds to troponin.
The troponin linkage with actin is weakened.
Troponin and tropomyosin are repositioned such that the binding sites on actin for the myosin cross-bridges are uncovered.
Myosin cross-bridges bind to actin and swivel inward (cross-bridge cycling or the power stroke).
Cross-bridge cycling pulls actin inward toward the center of the sarcomere such that the sarcomere becomes shorter and tension is developed.
ATP binds to the myosin cross-bridge and is split by myosin ATPase.
The energy released allows the cross-bridges to release actin and swivel outward.
As long as calcium is present and the binding sites are uncovered, cross-bridges continue binding to actin and continue cycling so that the sarcomere becomes shorter and tension is maintained.

their original positions. As a result, the myosin binding sites on the actin are covered once again. The thin filaments return passively to their original positions, resulting in muscle relaxation.

12.7 Sources of adenosine triphosphate (ATP) for muscle contraction

Skeletal muscle uses only ATP as a source of energy for contraction. However, intracellular stores of ATP are quite limited. In fact, the amount of ATP normally found in skeletal muscle is enough to sustain only a few seconds of contraction. Therefore, metabolic pathways to form additional ATP are needed. These pathways include:

- Creatine phosphate
- Oxidative phosphorylation
- Glycolysis

Energy may be transferred from *creatine phosphate* to ADP by way of the following reaction:

$$\text{Creatine phosphate} + \text{ADP} \overset{\text{CK}}{\leftrightarrow} \text{Creatine} + \text{ATP}$$

The enzyme creatine kinase (CK) facilitates the transfer of phosphate and energy to a molecule of ADP to form ATP. There are sufficient stores of creatine phosphate to sustain approximately 15 more seconds of muscle contraction. Because this is a single-step process, it provides ATP very rapidly, and it is the first pathway for the formation of ATP to be accessed.

The second pathway to be utilized in the formation of ATP is *oxidative phosphorylation*. This process involves the metabolic breakdown of glucose and fatty acids. Because it requires oxygen (*aerobic* metabolism), oxidative phosphorylation provides energy at rest and under conditions of mild (walking) to moderate (jogging) exercise. This pathway is advantageous because it produces a large amount of energy (36 molecules of ATP) from each molecule of glucose. However, oxidative phosphorylation is comparatively slow due to the number of steps involved. Finally, this process requires enhanced blood flow to the active muscles for the continuous delivery of oxygen as well as nutrient molecules. Although glucose may be obtained by way of glycogenolysis within the skeletal muscle fibers, these glycogen stores are limited. Glycogenolysis in the liver and lipolysis in the adipose tissues yield additional molecules of glucose and fatty acids for energy formation. As exercise is sustained, skeletal muscle relies more upon fatty acids as a source of fuel for the oxidative phosphorylation pathway. In this way, glucose is spared for the brain.

During intense exercise, when the oxygen supply cannot keep pace with the oxygen demand, skeletal muscle produces ATP *anaerobically* by way of *glycolysis*. Although this pathway provides ATP more rapidly, it produces much less energy (two molecules of ATP) from each molecule of glucose. Furthermore, glycolysis results in the production of lactic acid in the muscle tissue. The accumulation of lactic acid may lead to both pain as well as muscle fatigue.

12.8 Muscle fatigue

Muscle fatigue is defined as the inability of a muscle to maintain a particular degree of contraction over time. The onset of fatigue is quite variable and is influenced by several factors:

- Intensity and duration of contractile activity.
- Utilization of aerobic versus anaerobic metabolism for energy.
- Composition of the muscle (muscle fiber type, discussed in Section 12.10).
- Fitness level of the individual.

Although the exact mechanisms leading to muscle fatigue remain somewhat unclear, several factors have been implicated:

- Depletion of energy reserves.
- Conduction failure.
- Accumulation of lactic acid.
- Increase in inorganic phosphate.

Depletion of glycogen stores within the contracting skeletal muscles fibers is associated with the onset of fatigue. Interestingly, this occurs even though the muscle is utilizing fatty acids as its primary energy source.

The buildup of potassium ions in the T tubules as the result of repetitive stimulation may result in *conduction failure*. Excess K^+ ions in the fluid within the T tubules leads to a persistent depolarization and, consequently, the inactivation of sodium channels. As a result, action potentials are no longer conducted into the muscle fiber along the T tubules, and the release of Ca^{++} ions from the sarcoplasmic reticulum is interrupted. Recovery occurs rapidly upon the removal of the K^+ ions by way of the Na^+-K^+ pump.

The *accumulation of lactic acid* lowers the pH within the muscle. The change in pH may ultimately alter the activity of enzymes involved with energy production as well as cross-bridge cycling.

The breakdown of creatine phosphate causes an *increase in the concentration of inorganic phosphate*. Fatigue associated with elevated inorganic phosphate may be due to slowed release of P_i from myosin and, therefore, a decreased rate of cross-bridge cycling. It may also involve a decreased sensitivity of the contractile proteins to calcium, which would also impair cross-bridge cycling.

12.9 Oxygen debt

Hyperventilation persists for a period of time following the cessation of exercise. This hyperventilation is due to the *oxygen debt* incurred during the exercise. Specifically, oxygen is needed for the following metabolic processes:

- Restoration of creatine phosphate reserves.
- Metabolism of lactic acid.
- Replacement of glycogen stores.

During the recovery period from exercise, ATP, newly produced by way of oxidative phosphorylation, is needed to replace the *creatine phosphate reserves*. This process may be completed within a few minutes. Second, the *lactic acid*

produced during glycolysis must be metabolized. In the muscle, lactic acid is converted into pyruvic acid. Some of this pyruvic acid is then used as a substrate in the oxidative phosphorylation pathway to produce ATP. The remainder of the pyruvic acid is converted into glucose in the liver. This glucose is then stored in the form of *glycogen* in the liver and in the skeletal muscles. These later metabolic processes require several hours for completion.

12.10 Types of muscle fibers

There are two major differences between the types of muscle fibers: the *speed of contraction* and the *metabolic pathway used to form ATP*. As such, there are three types of muscle fibers (see Table 12.2):

1. Slow-twitch oxidative
2. Fast-twitch oxidative
3. Fast-twitch glycolytic

In humans, most skeletal muscles contain a mixture of all three types of muscle fibers. The dominant type of fiber in a given muscle is determined largely by the type of activity for which the muscle is specialized.

Fast-twitch muscle fibers develop tension two to three times faster than slow-twitch muscle fibers. This is due to the more rapid splitting of ATP by *myosin ATPase*. This enables the myosin cross-bridges to cycle more rapidly. Another factor influencing the speed of contraction involves the rate of

Table 12.2 Features of Skeletal Muscle Fiber Types

Feature	Slow-Twitch Oxidative	Fast-Twitch Oxidative	Fast-Twitch Glycolytic
Myosin ATPase Activity	Slow	Fast	Fast
Speed of Contraction	Slow	Fast	Fast
Removal of Calcium	Slow	Fast	Fast
Duration of Contraction	Long	Short	Short
Mitochondria	Many	Many	Few
Capillaries	Many	Many	Few
Myoglobin Content	High	High	Low
Color of Fiber	Dark red	Red	Pale
Diameter of Fiber	Small	Intermediate	Large
Source of ATP	Oxidative phosphorylation	Oxidative phosphorylation; glycolysis	Glycolysis
Glycogen Content	Low	Intermediate	High
Onset of Fatigue	Delayed	Intermediate	Rapid

removal of calcium from the cytoplasm. Muscle fibers remove the Ca^{++} ions by pumping them back into the sarcoplasmic reticulum. Fast-twitch muscle fibers remove Ca^{++} ions more rapidly than slow-twitch muscle fibers. This results in quicker twitches that are useful in fast, precise movements. The contractions generated in slow-twitch muscle fibers may last up to ten times longer than those of fast-twitch muscle fibers. Therefore, these twitches are useful in sustained, more powerful movements.

Muscle fibers also differ in their ability to resist fatigue. Slow-twitch muscle fibers rely primarily on oxidative phosphorylation for the production of ATP. Accordingly, these muscle fibers have a greater number of *mitochondria*, the organelles where these metabolic processes are carried out. These fibers also have an *extensive capillary network* for the delivery of oxygen and nutrient molecules. Furthermore, the *high myoglobin content* within slow-twitch muscle fibers facilitates the diffusion of oxygen into the cells from the extracellular fluid. The myoglobin imparts the characteristic red color to these muscle fibers. Hence, these muscles are referred to as *red muscles.*

Finally, slow-twitch muscle fibers have a *small diameter*. This facilitates the diffusion of oxygen through the fiber to the mitochondria where it is utilized. Taken together, each of these characteristics enhances the ability of these fibers to utilize oxygen. Therefore, in slow-twitch oxidative muscle fibers, *oxidative phosphorylation predominates* and *fatigue is delayed*. Examples of these muscles are the antigravity muscles of the back which are active and maintain posture for much of the day.

Fast-twitch muscle fibers fall into two categories. Fast-twitch glycolytic muscle fibers have fewer mitochondria, fewer capillaries, less myoglobin, and larger diameters. As a result, these fibers rely primarily on glycolysis for the production of ATP. The resulting accumulation of lactic acid and decrease in pH hastens the onset of fatigue. Because this type of muscle fiber has less myoglobin, it has a much paler appearance than the slow-twitch oxidative muscle fibers. Therefore, these muscles are referred to as "*white muscles.*" Examples of these muscles are the muscles of the hands and the eyes.

Fast-twitch oxidative muscle fibers have many mitochondria, many capillaries, a significant amount of myoglobin, and intermediate-sized diameters. These muscle fibers utilize a combination of oxidative and glycolytic metabolism to produce ATP. As a result, fast-twitch oxidative muscle fibers are more fatigue-resistant than fast-twitch glycolytic muscle fibers. Examples of these muscles are the muscles that move the limbs.

The percentage of the different types of muscle fibers in a given muscle may also be influenced by heredity. An individual with a high percentage of fast-twitch glycolytic muscle fibers would be better suited for power and sprint events. The individual with a high percentage of slow-twitch oxidative muscle fibers would be better suited for endurance events. For example, elite sprinters may have less than 20% slow-twitch muscle fibers in their quadriceps femoris muscles of the thighs, where marathon runners may have as high as 95% slow-twitch muscle fibers in these same muscles.

Finally, exercise training may cause fast-twitch muscle fibers of one type to convert to the other type. For example, fast-twitch glycolytic fibers may be converted to fast-twitch oxidative muscle fibers as the result of regular endurance training such as running. On the other hand, fast-twitch oxidative muscle fibers may be converted to fast-twitch glycolytic muscle fibers by resistance training such as weight lifting. Interestingly, fast-twitch and slow-twitch muscle fibers are not interconvertible. Whether a muscle fiber is fast-twitch or slow-twitch is determined by its nerve supply. Fast-twitch muscle fibers are innervated by motor neurons that exhibit intermittent, rapid bursts of electrical activity. Conversely, slow-twitch muscle fibers are innervated by motor neurons that exhibit a low-frequency pattern of electrical activity. It is important to note that all of the muscle fibers within a given motor unit are of the same type.

12.11 Muscle mechanics

A *muscle twitch* is a brief, weak contraction produced in a muscle fiber in response to a single action potential. Although the action potential lasts 1 to 2 msec, the resulting muscle twitch lasts approximately 100 msec. However, a muscle twitch in a single muscle fiber is too brief and too weak to be useful or to perform any meaningful work. In fact, hundreds or thousands of muscle fibers are organized into whole muscles. In this way, the fibers may work together to produce muscle contractions that are strong enough and of sufficient duration to be productive. Furthermore, muscles must be able to generate contractions of variable strengths. Different tasks require different degrees of contraction or tension development within the whole muscle. The *strength of skeletal muscle contraction* depends on two major factors: the *number of muscle fibers contracting* and the *amount of tension developed by each contracting muscle fiber.*

12.11.1 Number of muscle fibers contracting

As the *number of contracting muscle fibers* increases, then the strength of skeletal muscle contraction increases. Two major factors determine the number of muscle fibers that are activated at any given moment: *multiple motor unit summation* and *asynchronous motor unit summation*.

A *motor unit* is defined as an alpha motor neuron and all of the skeletal muscle fibers it innervates. The number of muscle fibers innervated by an alpha motor neuron varies considerably, depending upon the function of the muscle. For example, the muscles of the eyes and the hands have very small motor units. In other words, each alpha motor neuron associated with these muscles synapses with only a few muscle fibers. As a result, each of these muscles is innervated by a comparatively large number of alpha motor neurons. Densely innervated muscles are capable of carrying out more precise, complex motor activities. On the other hand, antigravity muscles

have very large motor units. For example, the gastrocnemius muscle of the calf has about 2000 muscle fibers in each motor unit. Muscles with large motor units tend to be more powerful and more coarsely controlled.

Multiple motor unit summation involves the recruitment of motor units. As the number of motor units stimulated at any given moment increases, then the strength of contraction increases.

Asynchronous motor unit summation refers to the condition where motor unit activation within a muscle is alternated. In other words, at one moment, some of the motor units within the muscle are activated, and other motor units are relaxed. This is followed by the relaxation of previously activated motor units and the activation of previously relaxed motor units. Consequently, only a fraction of the motor units within the muscle generate tension at any given moment. Therefore, this type of summation may generate submaximal contractions only.

An advantage of asynchronous motor unit summation is that the onset of muscle fatigue is significantly delayed. This is because each motor unit has alternating periods of relaxation where there is time for the restoration of energy supplies. The antigravity muscles of the back and the legs employ asynchronous motor unit summation. These muscles are required to generate sustained, submaximal contractions in order to maintain posture and body support over the course of the day.

12.11.2 *Amount of tension developed by each contracting muscle fiber*

As the amount of tension developed by each individual muscle fiber increases, then the overall strength of skeletal muscle contraction increases. Three major factors determine the amount of tension that may be developed by a contracting muscle fiber:

1. Frequency of nerve stimulation
2. Length of the muscle fiber at the onset of contraction
3. Diameter of the muscle fiber

As mentioned previously, a single action potential lasting only 2 msec causes a muscle twitch that lasts approximately 100 msec. If the muscle fiber has adequate time to completely relax before it is stimulated by another action potential, the subsequent muscle twitch will be of the same magnitude as the first. However, if the muscle fiber is restimulated before it has completely relaxed, then the tension generated during the second muscle twitch is added to that of the first (see Figure 12.3). In fact, the *frequency of nerve impulses* to a muscle fiber may be so rapid that there is no time for relaxation in between stimuli. In this case, the muscle fiber attains a state of smooth, sustained, maximal contraction referred to as *tetanus*.

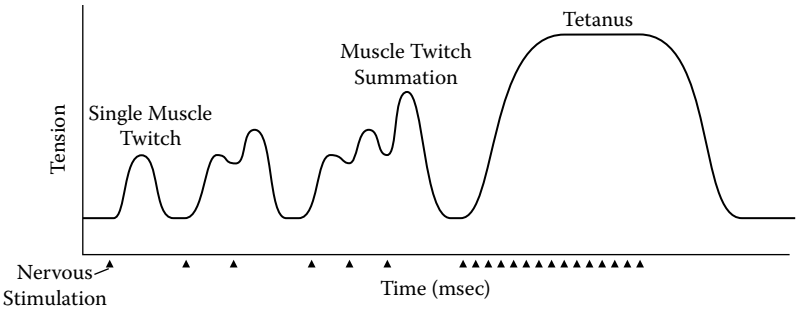


Figure 12.3 Muscle twitch summation and tetanus. A single action potential (represented by ▲) generates a muscle twitch. Because the duration of the action potential is so short, subsequent action potentials may restimulate the muscle fiber before it has completely relaxed. This leads to muscle twitch summation and greater tension development. When the frequency of stimulation becomes so rapid that there is no relaxation between stimuli, tetanus occurs. Tetanus is a smooth, sustained, maximal contraction.

The amount of tension developed by a muscle fiber during tetanic contraction can be as much as three to four times greater than that of a single muscle twitch. The mechanism involved with this increased strength of contraction involves the concentration of cytosolic calcium. Each time the muscle fiber is stimulated by an action potential, Ca^{++} ions are released from the sarcoplasmic reticulum. However, as soon as the Ca^{++} ions are released, a continuously active calcium pump begins returning the Ca^{++} ions to the sarcoplasmic reticulum. Consequently, fewer Ca^{++} ions are available to bind with troponin, and only a portion of the binding sites on the actin become available to the myosin cross-bridges. Each subsequent stimulation of the muscle fiber results in the release of more Ca^{++} ions from the sarcoplasmic reticulum. In other words, as the frequency of nerve stimulation increases, the rate of Ca^{++} ion release exceeds the rate of Ca^{++} ion removal. Therefore, the cytosolic concentration of calcium remains elevated. A greater number of Ca^{++} ions bind with troponin resulting in a greater number of binding sites on the actin that become available to the myosin cross-bridges. As the number of cycling cross-bridges increases, then the amount of tension developed increases.

The amount of tension developed by a stimulated muscle fiber is highly dependent upon the *length of the muscle fiber at the onset of contraction*. This association between the resting length of the muscle fiber and tension development is referred to as the *length–tension relationship*. The sarcomere length at which maximal tension can be developed is termed the *optimal length* (L_o). In skeletal muscle, optimal length is between 2.0 and 2.2 μ. At this point, the actin filaments have overlapped all of the myosin cross-bridges on the thick filaments (see Figure 12.4a). In other words, the potential for cross-bridge cycling and tension development upon stimulation has been maximized.

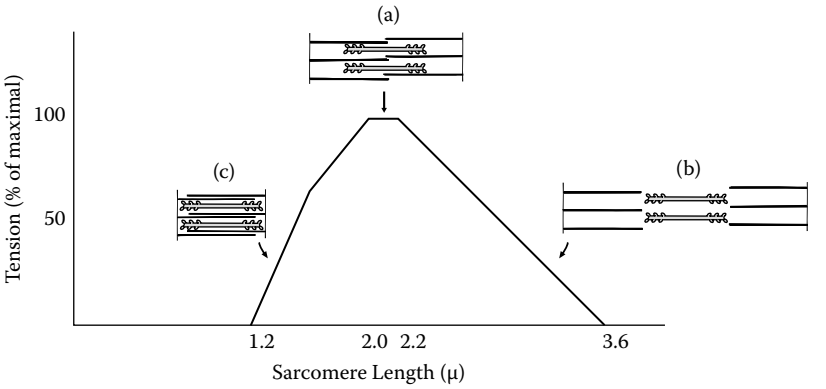


Figure 12.4 The length–tension relationship. The length of the sarcomere prior to stimulation influences the amount of tension that may be developed in the muscle fiber. (a) The optimal length of the sarcomere is between 2.0 and 2.2 μ. At this length, actin overlaps all of the myosin cross-bridges. The potential for cross-bridge cycling and the tension that may be developed upon stimulation of the muscle fiber are maximized. (b) When the sarcomere is overstretched so that the actin does not overlap the myosin cross-bridges, then cross-bridge cycling cannot take place and tension cannot be developed in the muscle fiber. (c) When the sarcomere is shortened prior to stimulation, then the thin filaments overlap each other and the thick filaments abut the Z lines. Further shortening and tension development upon stimulation is markedly impaired.

If the muscle fiber is stretched prior to stimulation such that the actin filaments have been pulled out to the end of the thick filaments, then there is no overlap between actin and the myosin cross-bridges (see Figure 12.4b). In this case, there is no cross-bridge cycling, and tension development is zero.

Tension development is also impaired when the muscle fiber is allowed to shorten prior to stimulation (see Figure 12.4c). If the actin filaments overlap each other, then there are fewer binding sites available for the myosin cross-bridges. Second, if the thick filaments are forced up against the Z lines, then further shortening cannot take place.

Interestingly, the range of resting sarcomere lengths is limited by the attachment of the skeletal muscles to the bones. Because of this fixed orientation, skeletal muscles cannot overstretch or overshorten prior to stimulation. Typically, these muscles are within 70% to 130% of their optimal length. In other words, attachment to the bones ensures that the overlap of actin and myosin is such that cross-bridge cycling approaches the maximum at all times. As the number of cycling cross-bridges increases, then the strength of muscle contraction increases.

The *diameter of the muscle fiber* is influenced by two major factors: *resistance training* and *testosterone*. Repeated bouts of anaerobic, high-intensity *resistance training*, such as weight lifting, cause muscle hypertrophy and an increase in

the diameter of the muscle fiber. This form of training promotes the synthesis of actin and myosin filaments. As a result, the number of cross-bridges available to cycle and develop tension is increased. Hence, larger muscles are capable of developing more powerful contractions.

Muscle fibers in males are thicker than those found in females. Therefore, their muscles are larger and stronger, even without the benefit of resistance training. This enlargement is due to the effects of *testosterone*, a sex hormone found primarily in males. Testosterone promotes the synthesis of actin and myosin filaments in muscle fibers.

Medical terminology

Agonist (ăg′ŏn-ĭst): Prime mover; muscle directly engaged in contraction and causing a movement.

Antagonist (ăn-tăg′ō-nĭst): Muscle that opposes the action of the prime mover.

Extension (ĕks-tĕn′shŭn): The act of straightening a joint.

Flexion (flĕk′shŭn): The act of bending a joint.

Insertion (ĭn-sĕr′shŭn): The more movable attachment at the distal end of a muscle.

Isometric (ī″sō-mĕ′trĭk): Regarding the maintenance of a constant length during muscular contraction.

Isotonic (ī″sō-tŏn′ĭk): Regarding the maintenance of a constant force during muscular contraction.

Motor unit: Somatic motor neuron and all of the muscle fibers it innervates.

Myoglobin (mī″ō-glō′bĭn): Iron-containing protein within muscle cells that binds with and stores oxygen.

Origin (or′ĭ-jĭn): The relatively fixed attachment of a muscle.

Oxygen debt (ŏk′sĭjĕn dĕt): The oxygen required during the recovery period following strenuous physical activity.

Rigor mortis (rĭg′or mŏr′tĭs): Stiffness that occurs in a dead body.

Sarcomere (săr′kō-mēr): Unit of contraction in muscle consisting of thick and thin filaments.

Synergist (sĭn′ĕr-jĭst): Muscle that assists the prime mover in producing a movement.

Tendon (tĕn′dŭn): Fibrous connective tissue that attaches muscles to bones.

Tension (tĕn′shŭn): Force; muscle contraction that performs work and produces heat.

Tetanus (tĕt′ă-nŭs): Smooth, sustained muscle contraction.

Bibliography

Fox, S. I., *Human Physiology*, 9th ed., McGraw-Hill, Boston, MA, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., Elsevier/Saunders, Philadelphia, PA, 2006.

Marieb, E. N., and Hoehn, K., *Anatomy and Physiology,* 3rd ed., Pearson Benjamin Cummings, San Francisco, CA, 2008.

McArdle, W. D., Katch, F. I., and Katch, V. L., *Exercise Physiology, Energy, Nutrition and Human Performance,* 5th ed., Williams & Wilkins, Baltimore, MD, 2001.

Porth, C. M., and Curtis, R. L., Disorders of Motor Function, in *Pathophysiology, Concepts of Altered Health States*, 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005.

Robergs, R. A., and Roberts, S. O., *Exercise Physiology, Exercise, Performance, and Clinical Applications,* Mosby, St. Louis, MO, 1997.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter thirteen

Smooth muscle

Study objectives

- Describe the morphological differences between skeletal muscle and smooth muscle
- Explain how contraction of smooth muscle occurs
- Explain how relaxation of smooth muscle occurs
- Explain why smooth muscle contraction is slow and prolonged
- Describe the latch state condition
- Distinguish between multiunit smooth muscle and single-unit smooth muscle
- Compare and contrast pacemaker potentials and slow-wave potentials
- List the factors that may alter smooth muscle contractile activity
- Explain how intracellular calcium concentration may be increased
- Describe the stress–relaxation response
- Describe the length–tension relationship in smooth muscle
- Discuss hyperplasia in smooth muscle

13.1 Introduction

Although skeletal muscle composes the bulk of the muscle tissue in the body, *smooth muscle* is far more important in terms of homeostasis. Most smooth muscle is found in layers or sheets in the walls of tubes and hollow organs. Contraction and relaxation of the smooth muscle in these tissues regulates the movement of substances within them. For example, contraction of the smooth muscle in the wall of a blood vessel narrows the diameter of the vessel and leads to a decrease in the flow of blood through it. Contraction of the smooth muscle in the wall of the stomach exerts pressure on its contents and pushes these substances forward into the small intestine.

Smooth muscle functions at a subconscious level and is *involuntary*. It is innervated by the *autonomic nervous system*, which regulates its activity.

13.2 Structure of smooth muscle

Smooth muscle cells are small and spindle shaped (thin and elongated) (see Table 13.1). Typically, these cells have a diameter of 5 to 10 µm and are 30 to 200 µm in length. This is in contrast to skeletal muscle fibers that may be ten times wider and thousands of times longer.

Table 13.1 Comparison of Skeletal and Smooth Muscle

Feature	Skeletal Muscle	Multiunit Smooth Muscle	Single-Unit Smooth Muscle
Location	Attached to bones; openings of some hollow organs (sphincters)	Large blood vessels; eyes; hair follicles	Walls of hollow organs of digestive, reproductive, and urinary tracts; small blood vessels
Thick Filaments	Myosin	Myosin	Myosin
Thin Filaments	Actin, troponin, tropomyosin	Actin, tropomyosin	Actin, tropomyosin
Filament Arrangement	Sarcomeres	Diamond-shaped lattice	Diamond-shaped lattice
Microscopic Appearance	Striated	Smooth	Smooth
Control	Voluntary	Involuntary	Involuntary
Innervation	Somatic nervous system	Autonomic nervous system	Autonomic nervous system
Contraction	Neurogenic	Neurogenic	Myogenic
Role of Nervous System	Initiate contraction	Initiate contraction	Modify contraction
Morphology	Large, cylindrical	Small, spindle shaped	Small, spindle shaped
Transverse Tubules	Yes	No	No
Sarcoplasmic Reticulum	Well developed	Very little	Very little
Source of Calcium	Sarcoplasmic reticulum	Extracellular fluid (most); sarcoplasmic reticulum (some)	Extracellular fluid (most); sarcoplasmic reticulum (some)
Site of Calcium Binding	Troponin	Calmodulin	Calmodulin
Function of Calcium	Reposition troponin/tropomyosin to uncover myosin binding sites on actin	Phosphorylate and activate myosin to bind with actin	Phosphorylate and activate myosin to bind with actin
Regulation of Tension Development	Alter number of contracting motor units; frequency of nerve stimulation	Alter number of contracting cells; alter intracellular Ca++ concentration	Alter intracellular Ca++ concentration
Length–Tension Development	Narrow	Broad	Broad

Similar to skeletal muscle, the contractile apparatus in smooth muscle consists of thick filaments composed of *myosin* and thin filaments composed of *actin*. However, in contrast to skeletal muscle, these filaments are not organized into sarcomeres. As such, there are no striations in this muscle, resulting in a "smooth" appearance.

Because there are no sarcomeres in smooth muscle, there are no Z lines. Instead, the actin filaments are attached to *dense bodies*. These structures, which contain the same protein as the Z lines, are positioned throughout the cytoplasm of the smooth muscle cell as well as attached to the internal surface of the plasma membrane. Myosin filaments are associated with the actin filaments forming contractile bundles that are oriented in a diagonal manner. This arrangement forms a *diamond-shaped lattice* of contractile elements throughout the cytoplasm. Consequently, the interaction of actin and myosin during contraction causes the cell to become shorter and wider.

The action potential easily penetrates all regions of these small cells. Therefore, smooth muscle does not have *transverse tubules*. Furthermore, there is very little *sarcoplasmic reticulum* in smooth muscle cells. Consequently, the intracellular storage of calcium is limited. Instead, the calcium needed for contraction is obtained primarily from the extracellular fluid. The cell membrane has multiple pouch-like infoldings referred to as *caveolae*. These caveolae are filled with extracellular fluid containing a high concentration of calcium. The influx of Ca^{++} ions through their channels in the cell membrane stimulates the release of a small amount of Ca^{++} ions from the sarcoplasmic reticulum.

13.3 Calcium and the mechanism of contraction

In skeletal muscle, calcium binds to troponin and causes the repositioning of tropomyosin. As a result, the myosin binding sites on the actin become uncovered and cross-bridge cycling takes place. Although an increase in cytosolic calcium is also needed in smooth muscle, its role in the mechanism of contraction is very different. There are three major steps involved in *smooth muscle contraction*:

1. Calcium binding with calmodulin.
2. Activation of myosin light-chain kinase.
3. Phosphorylation of myosin.

Upon entering the smooth muscle cell, Ca^{++} ions bind with *calmodulin*, an intracellular protein with a chemical structure similar to that of troponin. The resulting Ca^{++}–calmodulin complex binds to and activates *myosin light-chain kinase*. This activated enzyme then *phosphorylates myosin*. Cross-bridge cycling in smooth muscle may take place only when myosin has been phosphorylated.

Relaxation of smooth muscle involves two steps: the *removal of calcium ions* and the *dephosphorylation of myosin*.

Calcium ions are actively pumped back into the extracellular fluid as well as the sarcoplasmic reticulum by Ca^{++}-ATPase. When the concentration of calcium falls below a certain level, steps one and two of the contractile process are reversed. Calcium no longer binds with calmodulin, and myosin light-chain kinase is no longer activated.

The dephosphorylation of the myosin requires the activity of *myosin phosphatase*. This enzyme, which is located in the cytoplasm of the smooth muscle cell, splits the phosphate group from the myosin. Dephosphorylated myosin is inactive. Cross-bridge cycling no longer takes place, and the muscle relaxes.

PHARMACY APPLICATION: ACHALASIA

Achalasia is a condition where the lower esophageal sphincter leading into the stomach fails to relax. This esophageal spasm is also referred to as "nutcracker" esophagus. Achalasia results in considerable difficulty in swallowing, or *dysphagia*. Swallowed food remains lodged in the distended esophagus and passes into the stomach slowly over time. Of significant concern is the possibility of aspiration, or inhalation, into the airways of the esophageal contents when the patient lies down.

Pharmacotherapy may involve the administration of calcium channel antagonists. These medications facilitate the relaxation of the lower esophageal sphincter and the passage of food into the stomach.

13.4 Smooth muscle contraction is slow and prolonged

Contraction of smooth muscle is significantly slower than that of skeletal muscle. Furthermore, contraction in smooth muscle is quite prolonged (3000 msec) compared to that in skeletal muscle (100 msec). The slow onset of contraction as well as its sustained nature is due to the slowness of attachment and detachment of the myosin cross-bridges with the actin. Two factors are involved: *myosin ATPase activity* and the *rate of calcium removal*.

In smooth muscle, the myosin cross-bridges have less *myosin ATPase activity* than those of skeletal muscle. As a result, the splitting of adenosine triphosphate (ATP) that provides the energy to "prime" the cross-bridges, preparing them to interact with actin, is markedly reduced. Consequently, the rate of cross-bridge cycling is slower and the rate of tension development is slower. Furthermore, a slower *rate of calcium removal* causes the muscle to relax more slowly.

Interestingly, the reduction in myosin ATPase activity causes smooth muscle to be more *economical*. In other words, smooth muscle can maintain contraction with significantly less ATP consumption. Specifically, it can

maintain the same degree of tension for prolonged periods with only 1% of the energy that would be required by skeletal muscle. This is of benefit in many tissues, such as the blood vessels, that maintain tonic contraction all day, every day with little energy consumption and without developing fatigue. Furthermore, the prolonged attachment of the myosin cross-bridges to the actin results in an equal, if not greater, *force of contraction*. Smooth muscle is capable of developing a force of 4 to 6 kg/cm^2 cross-sectional area compared to 3 to 4 kg/cm^2 in skeletal muscle.

In some smooth muscles, when stimulation is sustained and the cytosolic calcium levels remain elevated, the rate of myosin ATPase activity in the cross-bridges declines; however, tension is maintained. This condition is referred to as the *latch state*. It occurs when a phosphorylated myosin cross-bridge becomes dephosphorylated while it is still attached to the actin. Dissociation of dephosphorylated cross-bridges from the actin occurs at a much slower rate than does the dissociation of phosphorylated cross-bridges. As a result, the muscle can maintain tension for long periods of time with very little ATP consumption.

13.5 Types of smooth muscle

There are two major types of smooth muscle (although it should be noted that many smooth muscles exhibit properties of each type): *multiunit smooth muscle* and *single-unit smooth muscle*.

Multiunit smooth muscle is located in the large blood vessels, the eyes (iris and ciliary muscle of the lens), and the piloerector muscles at the base of the hair follicles. This type of muscle consists of discrete smooth muscle cells or units that function independently. Each of these units is innervated by the autonomic nervous system. In fact, like skeletal muscle, this type of smooth muscle must be stimulated by these nerves in order to initiate contraction. Therefore, this muscle is referred to as *neurogenic*. Interestingly, nerve stimulation elicits graded potentials only. Action potentials do not occur in this muscle. The amount of ion flux that occurs in a single smooth muscle cell is inadequate to depolarize the cell to threshold. However, the graded potentials are sufficient to cause smooth muscle contraction. The contractile response of the whole muscle results from the sum of the responses of the multiple individual units.

Most smooth muscle is *single-unit smooth muscle*. Also referred to as *visceral smooth muscle*, it is found in the walls of tubes and hollow organs in the digestive system, the reproductive system, and the urinary system, as well as in the walls of small blood vessels. The cells of this type of smooth muscle are connected electrically by *gap junctions*. In this way, electrical activity can spread from one cell to the next, forming a *functional syncytium*. Any change in electrical activity in one region of the muscle quickly spreads throughout the muscle layer, such that the cells of the muscle function as one, or as a "single unit."

Action potentials are generated in single-unit smooth muscle. The simultaneous depolarization of 30 to 40 smooth muscle cells is required to generate a propagated action potential. The presence of the gap junctions allows this to occur readily.

Single-unit smooth muscle is *self-excitable* and is capable of generating action potentials without any input from the autonomic nervous system. Therefore, this type of smooth muscle is referred to as *myogenic*. In this muscle, the function of the autonomic nervous system is to modify contractile activity only. Input is not needed to elicit contraction.

The ability to spontaneously depolarize is related to the unstable resting membrane potentials in single-unit smooth muscle. There are two types of spontaneous depolarizations that may occur: *pacemaker potentials* and *slow-wave potentials*.

A *pacemaker potential* involves the gradual depolarization of the cell membrane to threshold. The subsequent generation of an action potential causes smooth muscle contraction. This type of spontaneous depolarization is referred to as a "pacemaker potential" because it creates a regular rhythm of contraction.

Slow-wave potentials also involve the gradual depolarization of the cell membrane. However, these depolarizations do not necessarily reach threshold. Therefore, the depolarization may simply be followed by a repolarization back to the initial membrane potential. These slow "wave-like" potentials occur rhythmically and do not lead to smooth muscle contraction. The peak-to-peak amplitude of the slow-wave potential is in the range of 15 to 30 mV. Therefore, under the appropriate conditions, the depolarization phase of the slow-wave potential may, in fact, reach threshold. When this occurs, a burst of action potentials is generated, resulting in muscle contraction.

The mechanism of the slow-wave potential is unclear. One hypothesis is that the rate at which sodium ions are actively transported out of the cell rhythmically increases and decreases. A decrease in the outward movement of Na^+ ions allows positive charges to accumulate along the internal surface of the cell membrane, and depolarization takes place. This is followed by an increase in the outward movement of Na^+ ions which causes the internal surface of the cell membrane to become more negative and repolarization takes place.

13.6 *Factors influencing the contractile activity of smooth muscle*

Many factors influence the contractile activity of smooth muscle. The strength of contraction of multiunit smooth muscle may be enhanced by the *stimulation of a greater number of cells*, or contractile units. This mechanism is directly comparable to motor-unit recruitment employed by skeletal muscle. As the number of contracting muscle cells increases, so does the strength of

contraction. However, this mechanism is of no value in single-unit smooth muscle. Due to the presence of gap junctions, all of the muscle cells in the tissue are activated at once.

The following are other factors that influence contractile activity:

- Autonomic nervous system
- Hormones and blood-borne substances
- Locally produced substances
- Intracellular calcium concentration

The autonomic nervous system (ANS) modifies the contractile activity of both types of smooth muscle. As discussed in Chapter 9, the ANS innervates the smooth muscle layer in a very diffuse manner, such that neurotransmitter is released over a wide area of the muscle. Typically, the effects of sympathetic stimulation and parasympathetic stimulation in a given tissue oppose each other; one system enhances contractile activity and the other system inhibits it. The specific effects (excitatory or inhibitory) that the two divisions of the ANS have on a given smooth muscle depend upon its location.

Many *hormones and other blood-borne substances* (including drugs) also alter the contractile activity of smooth muscle. Some of the more important substances include epinephrine, norepinephrine, angiotensin II, vasopressin, oxytocin, and histamine. *Locally produced substances* that may alter contraction in the tissue in which they are synthesized include nitric oxide, prostaglandins, leukotrienes, carbon dioxide, and hydrogen ion. It is important to note that smooth muscle cells are typically exposed to more than one of these substances at a time. Therefore, the magnitude of contraction at any given moment is determined by the sum of the effects of these agents on the muscle.

All of these factors (ANS stimulation, blood-borne and locally produced substances) alter smooth muscle contractile activity by altering the *intracellular concentration of calcium*. An increase in cytosolic calcium leads to an increase in cross-bridge cycling and, therefore, an increase in tension development. There are several mechanisms by which the concentration of calcium within the cytoplasm of the smooth muscle cell may be increased:

- Voltage-gated Ca^{++} channels
- Ligand-gated Ca^{++} channels
- Inositol triphosphate (IP_3)-gated Ca^{++} channels
- Stretch-activated Ca^{++} channels

Voltage-gated Ca^{++} channels open when the smooth muscle cell is depolarized. Calcium then enters the cell down its electrochemical gradient. *Ligand-gated Ca^{++} channels* are associated with various hormone or neurotransmitter receptors. Binding of a given substance to its receptor causes the ligand-gated Ca^{++} channel to open and, once again, Ca^{++} ions enter the cell. This process,

which occurs without a significant change in membrane potential (due to a simultaneous increase in Na^+ ion removal from the cell), is referred to as *pharmacomechanical coupling*.

Inositol triphosphate (IP_3)-gated channels are also associated with membrane-bound receptors for hormones and neurotransmitters. In this case, binding of a given substance to its receptor causes the activation of another membrane-bound protein, phospholipase C. This enzyme promotes the hydrolysis of phosphatidylinositol 4,5-diphosphate (PIP_2) to IP_3. The IP_3 then diffuses to the sarcoplasmic reticulum and opens its calcium channels to release Ca^{++} ions from this intracellular storage site.

Finally, an increase in volume or pressure within a tube or hollow organ causes stretch or distortion of the smooth muscle in the organ wall. This may cause activation of *stretch-activated Ca^{++} channels*. The subsequent influx of calcium initiates contraction of the smooth muscle. This process is referred to as *myogenic contraction*. Interestingly, the increased tension is only temporary as the smooth muscle cells adapt to their new length and relax. This is referred to as the *stress–relaxation response*. It allows a hollow organ to fill or expand slowly to accommodate a greater volume without developing strong contractions that would expel its contents. This is quite beneficial in organs such as the stomach and the urinary bladder whose functions include the temporary storage of food and urine, respectively.

13.7 Length–tension relationship

The length of the smooth muscle prior to stimulation has little effect on subsequent tension development. This is in marked contrast to skeletal muscle that exhibits a strong length–tension relationship. As discussed in Chapter 12, the influence of the resting muscle length on the tension developed in skeletal muscle is based upon the arrangement of the thick and thin filaments into sarcomeres. Any change in muscle length alters the degree of overlap of these filaments and, therefore, the number of cross-bridges cycling and the amount of tension developed.

The contractile elements in smooth muscle are not organized into sarcomeres. Furthermore, the resting length of smooth muscle is much shorter than its optimal length. In other words, this muscle can be significantly stretched, and the amount of tension developed may actually increase as the muscle becomes closer to its optimal length. Finally, the thick filaments are longer in smooth muscle than they are in skeletal muscle. As a result, there is still overlap of the thick and thin filaments, even when the muscle has been markedly stretched.

This very broad length–tension relationship in smooth muscle is physiologically advantageous. Tubes and hollow organs may be stretched considerably as substances pass through them. For example, in the urinary bladder, the smooth muscle cells may be stretched up to 2.5 times their resting length. By the end of pregnancy, the smooth muscle cells of the uterus may

be stretched up to eight times their resting length. Regardless, the smooth muscle must retain its ability to contract forcefully and regulate the movement of these through the organs.

13.8 Hyperplasia

All types of muscles may become larger by way of *hypertrophy* or an increase in cell size. This is the only mechanism by which skeletal muscles enlarge, as skeletal muscle fibers are incapable of mitosis. However, certain smooth muscles may also become larger by way of *hyperplasia* or an increase in cell numbers. For example, uterine smooth muscle is responsive to estrogen. At puberty, when a female's estrogen levels rise, the synthesis of smooth muscle is stimulated, which enables the uterus to grow to adult size. Furthermore, during pregnancy, the high levels of estrogen result in both hypertrophy and hyperplasia of the uterine smooth muscle so that the uterus may accommodate the growing fetus.

Medical terminology

Caveola (kăv-ē-ō′lă): Small pit or pouch formed on the cell surface.
Hyperplasia (hī″pĕr-plā′zē-ă): Increase in the number of cells in an organ or tissue.
Hypertrophy (hī-pĕr′trŏ-fē): Increase in size of an organ, tissue, or cell.
Ligand (lī′gănd, lĭg′ănd): Any chemical that binds to a specific receptor.
Myogenic (mī-ō-jĕn′ĭk): Originating in muscle.
Neurogenic (nū-rō-jĕn′ĭk): Resulting from nerve impulses.
Phosphorylation (fŏs″for-ĭ-lā′shŭn): The addition of a phosphate to an organic compound.
Syncytium (sĭn-sĭt′ē-ŭm): Group of cells in which the cytoplasm of one cell is continuous with that of the adjacent cells.

Bibliography

Costanzo, L., *Physiology*, 3rd ed., W.B. Saunders, Philadelphia, PA, 2006.
Fox, S. I., *Human Physiology*, 9th ed., McGraw-Hill, Boston, MA, 2006.
Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., Elsevier/Saunders, Philadelphia, PA, 2006.
Lombard, J. H., and Rusch, N. J., Cells, Nerves and Muscles, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 1.
Marieb, E. N., and Hoehn, K., *Anatomy and Physiology*, 3rd ed., Pearson Benjamin Cummings, San Francisco, CA, 2008.
Pasricha, P. J., Prokinetic Agents, Antiemetics, and Agents Used in Irritable Bowel Syndrome, in *Goodman & Gilman's: The Pharmacological Basis of Therapeutics*, 10th ed., Hardman, J. G., and Limbird, L. E., Eds., McGraw-Hill, New York, 2001.

Porth, C. M., *Pathophysiology, Concepts of Altered Health States*, 7th ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.
Rhoades, R., and Pflanzer, R., *Human Physiology,* 4th ed., Thomson Learning, Pacific Grove, CA, 2003.
Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.
Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.
Sperelakis, N., Pancreatic Secretion, in *Essentials of Physiology*, 2nd ed., Sperelakis, N., and Banks, R. O., Eds., Little, Brown, Boston, MA, 1996, chap. 48.
Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.
Weisbrodt, N. W., Swallowing, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 3.
Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter fourteen

Cardiac physiology

Study objectives

- Compare the functions of the right side of the heart and the left side of the heart
- Compare the functions of the arterial system and the venous system
- Describe the functions of the three major mechanical components of the heart: atria, ventricles, and valves
- Describe the route of blood flow through the heart
- Discuss the functions of the chordae tendinae and the papillary muscles
- Explain why the thickness of the myocardium varies between the different heart chambers
- Compare and contrast the functional and structural features of cardiac muscle and skeletal muscle
- Discuss the sources of calcium for myocardial contraction
- Explain the mechanism of action of cardiac glycosides in congestive heart failure
- Understand the physiological importance of the myocardial syncytium
- Describe the components of the specialized electrical conduction system of the heart
- Explain how the pacemaker of the heart initiates the heartbeat
- Understand the physiological importance of the atrioventricular (AV) nodal delay
- Describe the mechanism and the physiological significance of the rapid electrical conduction through the Purkinje fibers
- Compare and contrast the action potentials generated by the sinoatrial (SA) node and the ventricular muscle cells
- Distinguish between the types of sodium channels and calcium channels involved in the initiation and the conduction of the electrical impulse through the heart
- Discuss the mechanism and the physiological significance of the effective refractory period
- List the types of information obtained from an electrocardiogram
- Describe each of the components of the electrocardiogram
- Define *tachycardia* and *bradycardia*
- Understand how arrhythmias may be treated pharmacologically
- Define *systole* and *diastole*
- Describe the mechanical events, the status of the valves, and the pressure changes that take place during each phase of the cardiac cycle

14.1 Introduction

The cardiovascular system includes the *heart*, which serves as a pump for the blood, the *blood vessels*, which transport the blood throughout the body, and the *blood.* Under normal conditions, this system is a continuous, closed circuit, meaning that the blood is found only in the heart and the blood vessels. The heart, in particular, is discussed in this chapter and in Chapter 15. The blood vessels and the circulatory system, in general, are considered in Chapter 16. Finally, the blood and hemostasis are addressed in Chapter 17.

The heart actually consists of two separate pumps. The right side of the heart pumps blood to the lungs through the pulmonary circulation so that gas exchange, the uptake of oxygen, and the elimination of carbon dioxide, can take place. The left side of the heart pumps blood to the rest of the tissues of the body through the systemic circulation. In this way, oxygen and nutrients are delivered to the tissues to sustain their activities, and carbon dioxide and other metabolic waste products are removed from the tissues. In both circulations, blood vessels of the *arterial system*, arteries and arterioles, carry blood away from the heart and toward the tissues. The arterioles deliver blood to the *capillaries* where the exchange of substances between the blood and the tissues takes place. From the capillaries, blood flows into the vessels of the *venous system*, veins and venules, which carry blood back to the heart.

The human heart begins pumping approximately 3 weeks after conception, and it must continue this activity without interruption all day, every day for an entire lifetime. In a typical individual, this means the heart pumps over 100,000 times per day and propels about 2,000 gallons of blood through almost 65,000 miles of blood vessels. This function of the heart will be discussed in this chapter as well as in Chapter 15.

14.2 Functional anatomy of the heart

The heart is located in the center of the thoracic cavity. It sits directly above the muscles of the diaphragm, which separates the thorax from the abdomen, and lies beneath the sternum between the two lungs. The outermost layer of the heart, the *epicardium*, consists of a thin, fibrous membrane. The heart is enclosed and anchored in place by another fibrous membrane, or sac, referred to as the *pericardium*. The pericardium produces a small amount of *pericardial fluid* that fills the very narrow space between the two membranes. This fluid minimizes the friction produced by the movement of the heart when it beats.

To function mechanically as a pump, the heart must have:

- Receiving chambers
- Delivery chambers
- Valves

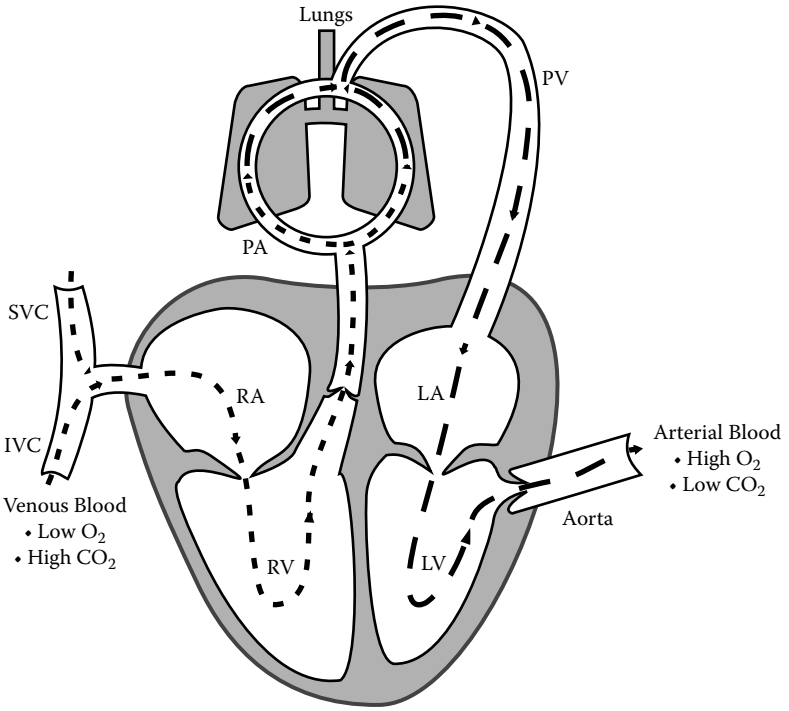


Figure 14.1 Route of blood flow through the heart. Systemic blood returns to the heart by way of the superior vena cava (SVC) and the inferior vena cava (IVC). This blood, which is low in oxygen and high in carbon dioxide, enters the right atrium (RA) and then the right ventricle (RV). The right ventricle pumps the blood through the pulmonary artery (PA) to the pulmonary circulation. It is within the lungs that gas exchange takes place. Next, this blood, which is now high in oxygen and low in carbon dioxide, returns to the heart by way of the pulmonary veins (PV). The blood enters the left atrium (LA) and then the left ventricle (LV). The left ventricle then pumps the blood through the aorta to the systemic circulation and the peripheral tissues.

The *atria* (singular: *atrium*) are the chambers that receive the blood returning to the heart through the veins. The blood then moves to the *ventricles* of the heart, or the delivery chambers. The powerful contractions of the ventricles generate a force sufficient to propel the blood through either the systemic or the pulmonary circulations. *Valves* ensure the one-way, or forward, flow of the blood.

The route of blood flow through the heart begins with the venae cavae, which return blood from the peripheral tissues to the right side of the heart (see Figure 14.1). The *superior vena cava* returns blood from the head and arms to the heart, and the *inferior vena cava* returns blood from the trunk of the body and the legs to the heart. Because this blood has already passed through the tissues of the body, it is low in oxygen. Blood from the venae

cavae first enters the *right atrium* and then the *right ventricle*. Contraction of the right ventricle propels this blood to the lungs through the pulmonary circulation by way of the *pulmonary artery*. As it flows through the lungs, the blood becomes enriched with oxygen and eliminates carbon dioxide to the atmosphere. Blood then returns to the heart through the *pulmonary veins*. The blood first enters the *left atrium* and then the *left ventricle*. Contraction of the left ventricle propels the now oxygen-rich blood back to the peripheral tissues through the systemic circulation by way of the *aorta*, the largest arterial vessel.

In summary, the heart is a single organ consisting of two pumps: the right heart delivers blood to the lungs and the left heart delivers blood to the rest of the body. Both pumps work simultaneously. The atria fill with blood and then contract at the same time, and the ventricles fill with blood and then contract at the same time. Contraction of the atria occurs prior to contraction of the ventricles in order to ensure proper filling of the ventricles with blood.

There are two sets of valves in the heart to maintain the one-way flow of blood as it passes through the heart chambers: the *atrioventricular* (AV) *valves* and the *semilunar valves*.

Each of these valves consists of thin flaps of flexible but tough fibrous tissue. Movement of the flaps is passive. In other words, they open and close in response to pressure changes within the heart chambers. The *AV valves* are found between the atria and the ventricles. The right AV valve is a tricuspid valve and has three cusps or leaflets. The left AV valve is a bicuspid valve as it has two cusps. This valve is also referred to as the *mitral valve*. When the ventricles contract, the pressure within them increases substantially, creating a pressure gradient for blood flow from the ventricles back into the atria where the pressure is very low. Closure of the AV valves prevents this potential backward flow of blood. However, what prevents this increased ventricular pressure from causing eversion of the valves, or the opening of the valves in the opposite direction, which would also allow blood to flow backward into the atria? There are strong fibrous ligaments, the *chordae tendineae*, attached to the flaps of the valves (see Figure 14.2). The chordae tendineae arise from cone-shaped *papillary muscles* that protrude into the ventricles. These muscles are continuous with the ventricular muscle. Therefore, when the ventricles are stimulated to contract, the papillary muscles also contract, pulling downward on the chordae tendineae. In this way, the flaps of the valves are not pushed open into the atria (a condition referred to as *prolapse*), but instead are held in place in the closed position. Blood is now forced to continue its forward progression and move from the ventricles into their respective arteries. (It is important to note that the papillary muscles do not open or close the AV valves. Their function is to limit their movement and prevent the valves from being pushed backward into the atria.)

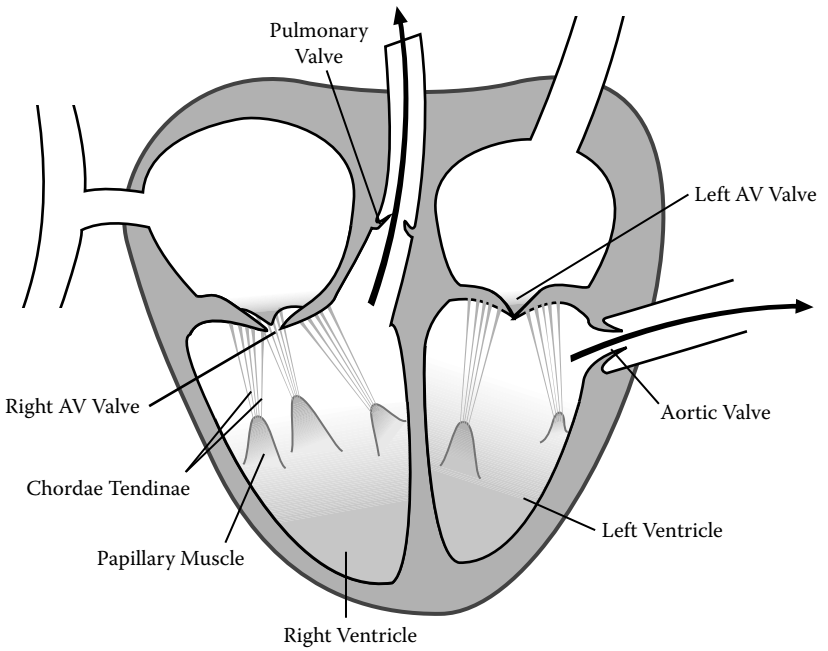


Figure 14.2 Papillary muscles and chordae tendineae. Papillary muscles, which protrude into the ventricles, are continuous with the ventricular muscle. Chordae tendineae extend from the papillary muscles to the flaps of the AV valves. Contraction of the ventricular muscle increases the pressure within the ventricles. This increased pressure closes the AV valves and opens the semilunar valves. In this way, blood is forced out of the semilunar valves to the pulmonary and systemic circulations. Theoretically, this increased pressure could also cause eversion, or the backward opening, of the AV valves into the atria. However, contraction of the papillary muscles pulls downward on the chordae tendineae. This action opposes eversion and holds the flaps of the AV valves in the closed position.

The *semilunar valves* separate the ventricles from their associated arteries. The *pulmonary valve* is found between the right ventricle and the pulmonary artery, and the *aortic valve* is found between the left ventricle and the aorta. These valves prevent the backward flow of blood from the pulmonary artery or the aorta into their preceding ventricles when the ventricles relax. The semilunar valves also have three cusps. There are no valves between the venae cavae or the pulmonary veins and the atria into which they deliver blood.

The closure of the valves causes the "lub-dub" associated with the heartbeat. The *first heart sound*, or the "lub," occurs when the ventricles contract and the AV valves close. The *second heart sound*, or the "dub," occurs when the ventricles relax and the semilunar valves close.

The heart valves are supported structurally by fibrous connective tissue that forms rings around each valve. These rings are referred to as *annuli fibrosi*.

14.3 Myocardial wall

The wall of the heart has three layers:

1. Epicardium
2. Endocardium
3. Myocardium

The outermost layer, the *epicardium*, is the thin, fibrous membrane on the external surface of the heart. The innermost layer, the *endocardium*, consists of a thin delicate layer of cells lining the chambers of the heart and the valve leaflets. The endocardium is continuous with the *endothelium*, which lines the blood vessels.

The middle layer is the *myocardium*, which is the muscular layer of the heart. This is the thickest layer although the thickness varies from one chamber to the next. Thickness of the myocardium is related to the amount of work that a given chamber must perform when pumping the blood. The atria, which serve primarily as receiving chambers, perform little pumping action. Under normal resting conditions, most of the blood (75%) moves passively along a pressure gradient (higher pressure to lower pressure) from the veins into the atria and into the ventricles where the pressure is close to zero. Therefore, it follows that the atria have relatively thin layers of myocardium, as powerful contractions are not necessary. On the other hand, when the ventricles contract, they must develop enough pressure to force open the semilunar valves and propel the blood through the entire pulmonary or systemic circulations. Under normal resting conditions, between heartbeats, the pressure in the pulmonary artery is approximately 8 mmHg, and the pressure in the aorta is approximately 80 mmHg. Therefore, in order to eject blood into the pulmonary artery, the right ventricle must generate a pressure greater than 8 mmHg. Furthermore, in order to eject blood into the aorta, the left ventricle must generate a pressure greater than 80 mmHg. Because the left ventricle performs significantly more work, its wall is much thicker than that of the right ventricle.

Cardiac muscle has many structural and functional similarities with skeletal muscle (Chapter 12) (see Table 14.1). The contractile elements, composed of thin actin filaments and thick myosin filaments, are organized into *sarcomeres*. Therefore, as with skeletal muscle, tension development within the myocardium occurs by way of the *sliding filament mechanism*. As the action potential travels along the surface of the muscle cell membrane, the impulse also spreads into the interior of the cell along the *transverse* (T) *tubules*. This stimulates the release of calcium from the *sarcoplasmic reticulum*. Calcium promotes the interaction of actin and myosin resulting in cross-bridge cycling and muscle shortening.

Unlike skeletal muscle whose only source of calcium is the sarcoplasmic reticulum, cardiac muscle also obtains calcium from the T tubules, which

Table 14.1 Distinguishing Features of Cardiac Muscle and Skeletal Muscle

Cardiac Muscle	Skeletal Muscle
Organized into sarcomeres	Organized into sarcomeres
Sliding-filament mechanism of contraction	Sliding-filament mechanism of contraction
Source of calcium:	Source of calcium:
Sarcoplasmic reticulum	Sarcoplasmic reticulum
Tranverse tubules	
Resting length of sarcomere Less than optimal length	Resting length of sarcomere Equal to optimal length
Gap junctions provide electrical communication between cells forming a functional syncytium	No gap junctions
Myogenic	Neurogenic
Contraction modified by autonomic nervous system	Contraction elicited by somatic nervous system

are filled with extracellular fluid. This calcium, which enters the myocardial cells through L-type Ca^{++} channels, binds to calcium receptors on the external surface of the sarcoplasmic reticulum. Receptor binding then stimulates the release of calcium from the sarcoplasmic reticulum. This mechanism is referred to as *calcium-induced calcium release*. Although most of the calcium utilized in the contractile process is obtained from the sarcoplasmic reticulum (90%), the release is dependent upon the movement of extracellular calcium into the muscle cells.

The amount of calcium that enters the cytoplasm determines the strength of contraction. Any physiological factor or pharmacological agent that increases cytosolic calcium will increase the strength of contraction and, therefore, increase the volume of blood pumped by the heart per minute. Physiological factors that influence cytosolic calcium concentration are discussed in Chapter 15.

Contraction ends when the calcium is returned to the sarcoplasmic reticulum and the extracellular fluid by way of Ca^{++}-ATPase pumps. As mentioned previously, calcium is also returned to the extracellular fluid by way of Na^{+}-Ca^{++} exchangers.

The arrangement of the myofilaments into sarcomeres renders the cardiac muscle subject to the *length–tension relationship*. When the resting sarcomere length is altered, the amount of tension developed by the myocardium upon stimulation is altered as well. In the heart, the resting sarcomere length is determined by the volume of blood within the ventricle immediately prior to contraction. This length–tension relationship is described by the *Frank-Starling mechanism* and is discussed in more detail in Chapter 15.

PHARMACY APPLICATION: CARDIAC GLYCOSIDES AND MYOCARDIAL CONTRACTILITY

A patient is considered to be in heart failure when cardiac output, or the volume of blood pumped by the heart per minute, is insufficient to meet the metabolic demands of the body. One way to improve cardiac output is to enhance myocardial contractility which will increase stroke volume, or the volume of blood pumped per beat. Contractility may be enhanced by medications that increase cytosolic calcium.

Digoxin is a cardiac glycoside and it may be used in the treatment of severe heart failure. This drug binds to and inactivates the Na^+-K^+ ATPase on the myocardial cell membrane. (Recall that for each adenosine triphosphate [ATP] expended, three Na^+ ions are pumped out of the cell and two K^+ ions are pumped into the cell.) Inhibition of the Na^+-K^+ ATPase results in the following:

- Decreased extrusion of Na^+ ions from the myocardial cell.
- Accumulation of Na^+ ions within the myocardial cell.
- Decreased concentration difference for Na^+ ions between the extracellular fluid and the intracellular fluid.
- Decreased diffusion of Na^+ ions into the myocardial cell.

An important mechanism of calcium removal from the myocardial cell between heartbeats involves the Na^+-Ca^{++} exchangers. The activity of the Na^+-Ca^{++} exchanger relies on the inward diffusion of sodium. As sodium enters the cell by way of the exchanger, calcium leaves the cell, returning to the extracellular fluid. Therefore, inhibition of the Na^+-K^+ ATPase with digoxin ultimately interrupts the activity of the Na^+-Ca^{++} exchangers (decreased Na^+ ion influx and decreased Ca^{++} ion efflux). As a result, calcium accumulates within the myocardial cell and contractility is increased.

There are also important differences between skeletal muscle and cardiac muscle. Skeletal muscle cells are elongated and run the length of the entire muscle. Furthermore, there is no electrical communication between these cells. Cardiac muscle cells, on the other hand, are considerably shorter than skeletal muscle fibers, and they branch and interconnect with each other. Intercellular junctions found where adjoining cells meet end-to-end are referred to as *intercalated discs*. There are two types of cell-to-cell junctions within these discs. *Desmosomes* hold the muscle cells together and provide the structural support needed when the heart beats and exerts a mechanical stress that would tend to pull the cells apart. *Gap junctions* are areas of very low electrical resistance (1/400 the resistance of the outside membrane) that

allow free diffusion of ions. It is through the gap junctions that the electrical impulse, or heartbeat, spreads rapidly from one cell to another. As a result, the myocardium is a *syncytium* where the initiation of a heartbeat in one region of the heart results in the stimulation and contraction of all of the cardiac muscle cells at essentially the same time. The heart is actually composed of two syncytiums: the atrial syncytium and the ventricular syncytium. In each case, but particularly in the ventricles, the simultaneous stimulation of all of the muscle cells results in a more powerful contraction, facilitating the pumping of the blood.

Skeletal muscle is neurogenic and requires stimulation from the somatic nervous system to initiate contraction. Because there is no electrical communication between these cells, each muscle fiber is innervated by a branch of an alpha motor neuron. Cardiac muscle, however, is *myogenic*, or self-excitatory. This muscle spontaneously depolarizes to threshold and generates action potentials without external stimulation. The region of the heart with the fastest rate of inherent depolarization initiates the heartbeat and determines the heart rhythm. In normal hearts, this "*pacemaker*" region is the sinoatrial node.

The heart is richly innervated by the autonomic nervous system (ANS). The sympathetic division, which innervates the entire heart, is excitatory and increases heart rate and contractility. The parasympathetic division, contained in the vagus nerves, primarily innervates the atria. Vagal activity inhibits the heart and decreases heart rate.

Skeletal muscle contracts only when needed. During intense contractions when oxygen demand is greater than oxygen supply, skeletal muscle may form ATP by way of glycolysis. The drawback to this mechanism of ATP production is the accumulation of lactic acid which leads to pain and muscle fatigue. In contrast, cardiac muscle must beat all day, every day. Furthermore, it must obtain all of its ATP by way of oxidative-phosphorylation which is energetically more efficient and avoids fatigue. In order to meet this energy demand, the heart must have a substantial blood flow for the delivery of oxygen and nutrients, as well as an abundance of mitochondria where oxidative-phosphorylation takes place.

The blood that flows through the chambers of the heart does not supply the heart muscle with oxygen and nutrients. Instead, the myocardium is supplied with blood from the *coronary arteries*. These arteries originate at the very beginning of the aorta and lead to a branching network of vessels (arterioles, capillaries, veins) similar to that found in other tissues. When one of these arteries becomes occluded with an atherosclerotic plaque or a thrombus, the patient suffers a *myocardial infarction*, or a "heart attack." On average, under resting conditions, the heart muscle receives approximately 5% of the cardiac output, or about 250 ml/min.

Mitochondria occupy approximately 30% of the volume of a cardiac muscle cell. These organelles convert oxygen and nutrients into the ATP needed for

muscle contraction. Cardiac muscle consumes, or extracts, 70% to 80% of the oxygen delivered to it each minute. This is more than twice the amount consumed by other tissues in the body. Once again, this illustrates the need for continuous blood flow to the heart muscle as there is very little oxygen reserve. During exercise when cardiac workload is increased, the muscle cannot meet its oxygen needs by extracting more oxygen from the blood. Instead, more blood must be delivered to the myocardium. Blood flow may increase to four to five times the resting rate, or to as much as 1250 ml/min.

14.4 Electrical activity of the heart

The specialized excitation and electrical conduction system in the heart consists of the following:

- Sinoatrial node
- Interatrial pathway
- Internodal pathway
- Atrioventricular node
- Bundle of His
- Bundle branches
- Purkinje fibers

The *sinoatrial* (SA) *node* is located in the wall of the right atrium near the entrance of the superior vena cava. The specialized cells of the SA node spontaneously depolarize to threshold and generate 70 to 75 heartbeats/minute. Because the SA node has the fastest rate of spontaneous depolarization of any tissue in the heart, it is considered the *pacemaker* of the heart. In other words, the SA node determines the heart rate.

The "resting" membrane potential, or *pacemaker potential*, is different from that of neurons, which were discussed in Chapter 3. First, this potential is approximately –55 mV, which is less negative than that found in neurons (–70 mV) (see Figure 14.3a). Second, the pacemaker potential is unstable and slowly depolarizes toward threshold (*Phase 4*). Three important ion currents contribute to this slow depolarization and involve the following types of channels:

1. Potassium channels
2. F-type sodium channels
3. T-type calcium channels

Throughout the pacemaker potential, there is a progressive decrease in the permeability to potassium. As such, positive charges do not diffuse out of the cell.

Pacemaker cells have unique sodium channels referred to as F-type sodium channels (F = funny). Interestingly, these voltage-gated channels

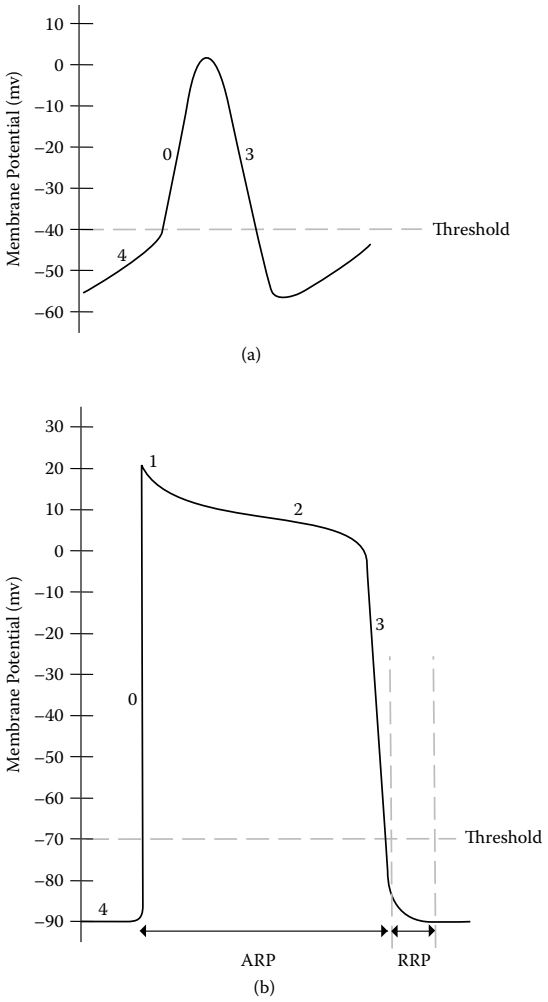


Figure 14.3 Cardiac action potentials. (a) Sinoatrial (SA) node: During Phase 4, the pacemaker potential, the cells of the SA node depolarize toward threshold due to the influx of Na^+ ions and Ca^{++} ions. The upward swing of the action potential, Phase 0, results from the influx of calcium through slow Ca^{++} channels. Repolarization, Phase 3, is due to the efflux of K^+ ions. (b) Ventricular muscle: The resting membrane potential, Phase 4, is very negative due to the high permeability of the K^+ channels. The upward swing of the action potential, Phase 0, results from the rapid influx of sodium through fast Na^+ channels. The brief repolarization that occurs during Phase 1 is due to the abrupt closure of these channels. The plateau of the action potential, Phase 2, results from the influx of calcium through slow Ca^{++} channels. Finally, repolarization, Phase 3, is due to the efflux of K^+ ions. The absolute, or effective, refractory period (ARP) persists until the fast Na^+ channels return to their resting state (–70 mV). No new action potentials may be generated during this period. This is followed by the relative refractory period (RRP).

open when the membrane potential becomes *negative*. The resulting influx of sodium and, therefore, positive charges contributes to the depolarization of the pacemaker cells during this phase. It is important to note that these F-type sodium channels differ from the fast Na^+ channels that cause rapid depolarization in other types of excitable cells such as skeletal muscle and neurons. In fact, pacemaker cells are essentially devoid of fast sodium channels.

Toward the end of Phase 4, T-type Ca^{++} channels (T = transient) start to become activated allowing Ca^{++} ion influx. The diffusion of additional positive charges into the pacemaker cells continues to depolarize the membrane toward threshold.

Phase 0 begins when the membrane potential reaches threshold (–40 mV). Recall that the upstroke of the action potential in neurons is due to increased permeability of fast Na^+ channels resulting in a steep, rapid depolarization. However, in the SA node, the action potential develops more slowly because the fast Na^+ channels do not play a role. Whenever the membrane potential is less negative than –60 mV for more than a few milliseconds, these channels become inactivated. With a resting membrane potential of –55 mV, this is clearly the case in the SA node. Instead, when the membrane potential reaches threshold in this tissue, many L-type Ca^{++} channels (L = long-lasting) open, resulting in the depolarization phase of the action potential. These channels open slowly, and therefore, the slope of this depolarization is less steep than that of neurons.

Phase 3 begins at the peak of the action potential. At this point, the L-type Ca^{++} channels close and K^+ channels open. The resulting efflux of K^+ ions causes the repolarization phase of the action potential.

Because cardiac muscle is myogenic, nervous stimulation is not necessary to elicit the heartbeat. However, the heart rate is modulated by input from the autonomic nervous system. Both the sympathetic and the parasympathetic systems innervate the SA node. Sympathetic stimulation causes an increase in heart rate or an increased number of beats per minute. Norepinephrine, which stimulates β_1-adrenergic receptors, increases the rate of pacemaker depolarization by increasing the permeability to both Na^+ ions and Ca^{++} ions. If the heartbeat is generated more rapidly, then there will be more beats per minute.

Parasympathetic stimulation causes a decrease in heart rate. Acetylcholine, which stimulates muscarinic receptors, increases the permeability to potassium. Enhanced K^+ ion efflux has a twofold effect. First, the cells become hyperpolarized, and therefore, the membrane potential is farther away from threshold. Second, the rate of pacemaker depolarization is decreased because the outward movement of K^+ ions opposes the effect of the inward movement of Na^+ and Ca^{++} ions. The result of these two effects of potassium efflux is that it takes longer for the SA node to reach threshold and generate an action potential. If the heartbeat is generated more slowly, then there will be fewer beats per minute.

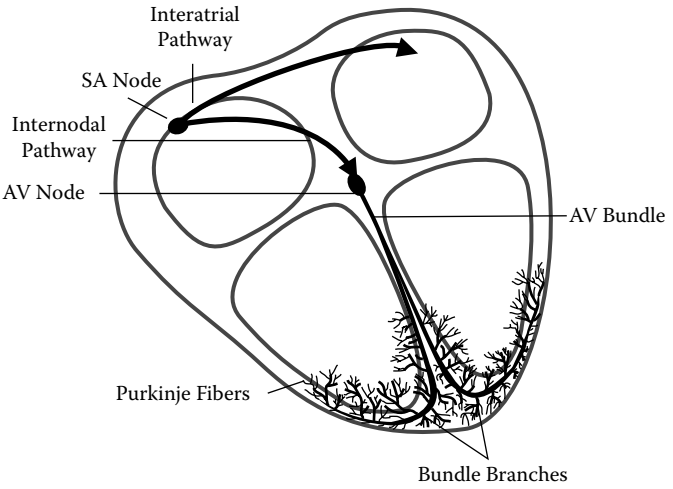


Figure 14.4 Route of excitation and conduction in the heart. The heartbeat is initiated in the sinoatrial (SA) node, or the pacemaker, in the right atrium of the heart. The electrical impulse is transmitted to the left atrium through the interatrial conduction pathway and to the atrioventricular (AV) node through the internodal pathway. From the AV node, the electrical impulse enters the ventricles and is conducted through the AV bundle, the left and right bundle branches, and, finally, the Purkinje fibers, which terminate on the true cardiac muscle cells of the ventricles.

From the SA node, the heartbeat spreads rapidly throughout both atria by way of the gap junctions. As mentioned previously, the atria are stimulated to contract simultaneously. An *interatrial conduction pathway* extends from the SA node in the right atrium across to the left atrium. Its function is to facilitate the conduction of the impulse through the right and left atria simultaneously, creating the atrial syncytium (see Figure 14.4).

An *internodal conduction pathway* (or *Bachmann's bundle*) also extends from the SA node and transmits the impulse directly to the *AV node*. This node is located at the base of the right atrium near the interventricular septum, which is the wall of myocardium separating the two ventricles. The atria and the ventricles are separated from each other by fibrous connective tissue referred to as the *fibrous skeleton* of the heart. Therefore, the electrical impulse cannot spread directly to the ventricles. Instead, the AV node serves as the only pathway through which the impulse can be transmitted to the ventricles. The speed of conduction through the AV node is slowed, resulting in a slight delay (0.1 seconds). The cause of this *AV nodal delay* is partly due to the smaller fibers of the AV node. More importantly, however, there are fewer gap junctions between the cells of the node, which increases the resistance to current flow. The physiological advantage of the AV nodal delay is that it allows the atria to complete their contraction before ventricular

contraction begins. This timing ensures proper filling of the ventricles prior to contraction.

From the AV node, the electrical impulse spreads through the *AV bundle* or the *bundle of His* (see Figure 14.4). This portion of the conduction system penetrates the fibrous tissue separating the atria from the ventricles, and it enters the interventricular septum where it divides into the *left and right bundle branches*. The bundle branches travel down the septum toward the apex of the heart and then reverse direction traveling back toward the atria along the outer walls of the ventricles. This route of conduction of the impulse facilitates the ejection of blood from the ventricles. If the impulse were to be conducted directly from the atria to the ventricles, then the ventricular contraction would begin at the top of the chambers and proceed downward toward the apex. This would trap the blood at the bottom of the chambers. Instead, the wave of ventricular electrical stimulation and contraction moves from the apex of the heart toward the top of the chambers where the semilunar valves are located and ejection takes place.

Not surprisingly, the first regions of the ventricles to be stimulated and, therefore, to contract, are the interventricular septum and the papillary muscles. Contraction of the septum makes it more rigid, and as a result, it provides physical support for the subsequent contraction of the rest of the ventricular myocardium. As discussed previously, contraction of the papillary muscles prevents potential eversion of the AV valves during ventricular systole.

The final portion of the specialized conduction system consists of the *Purkinje fibers,* which extend from the bundle branches (see Figure 14.4). These fibers, which spread throughout the myocardium, terminate on the true cardiac muscle cells of the ventricles. The rate of conduction of the impulse through the Purkinje fibers is very rapid, resulting in the functional syncytium of the ventricles discussed earlier. The entire ventricular myocardium is stimulated almost simultaneously, which strengthens its pumping action. The increased rate of conduction (six times the rate of other ventricular muscle cells) is due in part to the large diameter of the Purkinje fibers. Typical myocardial fibers are 10 to 15 μm in diameter, where Purkinje fibers are 70 to 80 μm in diameter. Furthermore, there is a very high level of permeability of the gap junctions, which decreases the resistance to current flow. It is estimated that Purkinje fibers conduct impulses at a velocity of 1 to 4 m/sec.

The action potential generated in the ventricular muscle is very different from that originating in the SA node (see Figure 14.3). The resting membrane potential is not only stable, it is much more negative than that of the SA node. Second, the slope of the depolarization phase of the action potential is much steeper. Finally, there is a lengthy plateau phase of the action potential where the muscle cells remain depolarized for approximately 250 to 300 msec. The physiological significance of this sustained depolarization is that it leads to sustained contraction (also about 300 msec) that facilitates the ejection of the

blood. These disparities in the action potentials are explained by differences in ion channel activity in ventricular muscle compared to the SA node.

At rest, the permeability to K^+ ions in ventricular muscle cells is significantly greater than that of Na^+ ions. Cardiac muscle possesses a unique subtype of potassium channel that is especially leaky at negative membrane potentials. As a result, ventricular muscle cells have a stable resting membrane potential that approaches the equilibrium potential for K^+ of –90 mV (*Phase 4*) (see Figure 14.3b). Upon stimulation by an electrical impulse, the voltage-gated fast Na^+ channels open causing a marked increase in the permeability to Na^+ ions and a rapid and profound depolarization of the membrane potential to almost 30 mV (*Phase 0*). These voltage-gated Na^+ channels remain open very briefly, and within 1 msec they are inactivated. The resulting decrease in sodium permeability causes a small repolarization (*Phase 1*).

The ventricular muscle cells do not completely repolarize immediately as do neurons and skeletal muscle cells, however. Instead, there is a plateau phase of the action potential (*Phase 2*). During this phase, there is a decrease in the permeability to K^+ ions (through the subtype of channels mentioned above) and a marked increase in the permeability to Ca^{++} ions. Like the voltage-gated Na^+ channels, the voltage-gated, L-type Ca^{++} channels are also activated by depolarization; however, they open much more slowly. The combination of decreased K^+ ion efflux and increased Ca^{++} ion influx causes the prolonged depolarization.

Repolarization (*Phase 3*) occurs when the Ca^{++} channels close and the K^+ channels open allowing for the rapid efflux of K^+ ions and a return to the resting membrane potential. These potassium channels, which are involved with repolarization, are similar to those that repolarize neurons and skeletal muscle cells. In other words, they open in response to depolarization, and they close in response to a return to a negative membrane potential.

As in neurons, cardiac muscle cells undergo an absolute or *effective refractory period* where at the peak of the action potential the voltage-gated fast Na^+ channels become inactivated and incapable of opening regardless of further stimulation. Therefore, the fast Na^+ channels cannot reopen, Na^+ ions cannot enter the cell, and another action potential cannot be generated. These channels do not return to their resting position and become capable of opening in sufficient numbers to generate a new action potential until the cardiac muscle cell has repolarized to approximately –70 mV. As a result, the absolute refractory period lasts almost as long as the duration of the associated contraction, about 250 to 300 msec. The physiological significance of this phenomenon is that it prevents the development of tetanus or spasm of the ventricular myocardium. By the time the cardiac muscle cell can be stimulated to generate another action potential, the contraction from the previous action potential is over. Therefore, the tension from sequential action potentials cannot accumulate

and become sustained. This is in contrast to skeletal muscle where tetanic contractions readily occur in order to produce maximal strength (Chapter 12). The pumping action of the heart, however, requires alternating contraction and relaxation so that the chambers can fill with blood. Sustained contraction or tetanus would preclude ventricular filling.

The effective refractory period is followed by a *relative refractory period* that lasts for the remaining 50 msec of the ventricular action potential. During this period, action potentials may be generated; however, the myocardium is more difficult than normal to excite.

14.5 Electrocardiogram

A portion of the electrical current generated by the heartbeat flows away from the heart through the surrounding tissues and reaches the body surface. Using electrodes placed on the skin, this current can be measured and used to produce a recording referred to as the *electrocardiogram* (ECG). (The ECG is also referred to as the EKG, which is the abbreviation for the German *elektrokardiogramm*.) An important point to remember regarding the ECG is that it represents the sum of all electrical activity throughout the heart at any given moment, not individual action potentials. Therefore, an upward deflection of the recording does not necessarily represent depolarization, nor does a downward deflection represent repolarization. Furthermore, a recording is made only when current is flowing through the heart during the actual process of depolarization or repolarization. No recording is made when the heart is completely depolarized (during the plateau phase of the ventricular action potential) or completely repolarized (between heartbeats). The ECG provides information concerning the following:

- The relative size of the heart chambers.
- Various disturbances of rhythm and electrical conduction.
- The extent and location of ischemic damage to the myocardium.
- The effects of altered electrolyte concentrations.
- The influence of certain drugs (e.g., digoxin and antiarrhythmic drugs).

The ECG provides no information concerning contractility or, in other words, the mechanical performance of the heart as a pump.

The normal ECG is composed of the following (Figure 14.5):

- *P wave*: Caused by atrial depolarization.
- *QRS complex*: Caused by ventricular depolarization.
- *T wave*: Caused by ventricular repolarization.

There are several noteworthy characteristics of the ECG. First, the firing of the SA node, which initiates the heartbeat, precedes atrial depolarization.

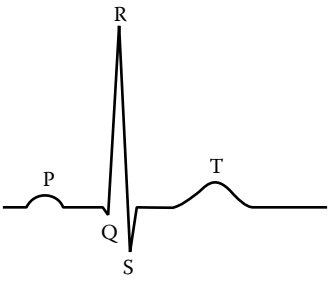


Figure 14.5 Electrocardiogram. The electrocardiogram (ECG) is a measure of the overall electrical activity of the heart. The P wave is caused by atrial depolarization, the QRS complex is caused by ventricular depolarization, and the T wave is caused by ventricular repolarization.

Therefore, it should be apparent immediately prior to the P wave. However, due to its small size, it does not generate enough electrical activity to spread to the surface of the body and be detected by the electrodes. Therefore, there is no recording of the depolarization of the SA node.

Second, the area under the curve of the P wave is small compared to that of the QRS complex. This is related to the muscle mass of the chambers. The ventricles have significantly more muscle than the atria and, therefore, generate more electrical activity. Furthermore, although it may not appear to be the case given the spike-like nature of the QRS complex, the areas under the QRS complex and the T wave are approximately the same. This is because these recordings both represent electrical activity of the ventricles even though one is caused by depolarization and the other is caused by repolarization. Either way, the muscle mass involved is the same.

Third, there is no recording during the PR segment. At this time, the electrical impulse is being conducted through the AV node. As with the SA node, there is not enough tissue involved to generate sufficient electrical activity to be detected by the electrodes. The length of the PR segment is determined by the duration of the AV nodal delay.

Finally, there is no recording during the ST segment. This is the period between ventricular depolarization and ventricular repolarization. In other words, the ventricles are completely depolarized and the muscle cells are in the plateau phase of the action potential. As mentioned earlier, unless current is actually flowing through the myocardium, there is no recording.

Using the ECG, the heart rate may be determined by calculating the time from the beginning of one P wave to the beginning of the next P wave, or from peak to peak of the QRS complexes. A normal resting heart rate in adults is approximately 70 beats/minute. A heart rate of less than 60 beats/minute is referred to as *bradycardia,* and a heart rate of more than 100 beats/minute is referred to as *tachycardia*.

PHARMACY APPLICATION: ANTIARRHYTHMIC DRUGS

Normal cardiac contraction depends on the conduction of the electrical impulses through the myocardium in a highly coordinated fashion. Any abnormality of the initiation or propagation of the impulse is referred to as an *arrhythmia*. These disorders are the most common clinical problem encountered by a cardiologist. There is a wide range of types of arrhythmias with multiple etiologies and a variety of symptoms. Two types of cardiac tachyarrhythmias are discussed here. The most common treatment for these conditions is drug therapy.

Atrial or supraventricular tachycardia is an arrhythmia whose pathophysiology originates above the bifurcation of the bundle of His. Verapamil (Class IV antiarrhythmic drug) is an effective agent for this type of arrhythmia. A Ca^{++} channel blocker, it is most potent in tissues where the action potentials depend on calcium currents. These include slow-response tissues such as the SA node and the AV node. The effects of verapamil include a decrease in heart rate and a decrease in the conduction velocity of the electrical impulse through the AV node. The resulting increase in duration of the AV nodal delay, which is illustrated by a lengthening of the PR segment in the ECG, reduces the number of impulses permitted to penetrate to the ventricles to cause contraction.

Due to the potential for interfering with the pumping action of the heart, ventricular arrhythmias are considered to be more serious than those occurring in the atria. Ventricular tachycardia is an arrhythmia whose pathophysiology originates distal to the bifurcation of the bundle of His. Procainamide (Class IA antiarrhythmic drug) is an effective agent for this type of arrhythmia. Its mechanism of action involves the blockade of the fast Na^{+} channels responsible for Phase 0 in the fast response tissue of the ventricles. Therefore, its effect is most pronounced in the Purkinje fibers. The effects of this drug's activity include a decrease in the excitability of the myocardial cells and a decrease in conduction velocity. Therefore, a decrease in the rate of the Phase 0 upstroke and a prolonged repolarization are observed. As a result, the duration of the action potential and the associated refractory period are prolonged, and the heart rate is reduced. These effects are illustrated by an increase in the duration of the QRS complex.

14.6 Cardiac cycle

The *cardiac cycle* is the period of time from the beginning of one heartbeat to the beginning of the next. As such, it consists of two alternating phases:

1. *Systole*: Where the chambers contract and eject the blood out of them.
2. *Diastole*: Where the chambers relax, allowing them to fill with blood.

Table 14.2 Systole and Diastole in the Atria versus the Ventricles

	Systole	Diastole	Total
Atria	0.1	0.7	0.8
Ventricles	0.3	0.5	0.8

Note: In a resting adult with a heart rate of 75 beats/minute, the cardiac cycle lasts 0.8 seconds. The atria, which are the receiving chambers, have a relatively longer period of diastole (0.7 seconds). The ventricles, which are the delivery chambers, have a relatively longer period of systole (0.3 seconds).

Table 14.3 Summary of Events Occurring during the Cardiac Cycle

	Filling	Isovolumetric Contraction	Ejection	Isovolumetric Relaxation
Period	Diastole	Systole	Systole	Diastole
Pressures	$P_A > P_V < P_{aorta}$ P_A, $P_V \approx$ 0–10 mmHg $P_{aorta} \approx$ 80 mmHg	$P_A < P_V < P_{aorta}$ P_V increases toward 80 mmHg	$P_A < P_V > P_{aorta}$ P_V, $P_{aorta} \approx$ 120 mmHg	$P_A < P_V < P_{aorta}$ P_V decreases toward 0 mmHg
AV Valve	Open	Closed	Closed	Closed
Aortic Valve	Closed	Closed	Open	Closed
Ventricular Volume	Increases from 60 ml (ESV) to 130 ml (EDV)	No change	Decreases from 130 ml (EDV) to 60 ml (ESV)	No change
ECG	TP segment P Wave PR segment	QRS complex	ST segment	T wave
Heart Sounds	None	First heart sound	None	Second heart sound

Both the atria and the ventricles undergo phases of systole and diastole; however, the duration of each phase in the chambers differs (see Table 14.2). In the atria, whose primary function is to receive blood returning to the heart from the veins, diastole is the predominant phase lasting for almost 90% of each cardiac cycle at rest. In the ventricles, whose primary function is to develop enough force to eject the blood into the pulmonary or systemic circulations, systole is much longer lasting and accounts for almost 40% of each cycle at rest.

A discussion of the cardiac cycle requires the correlation of pressure changes, ventricular volume changes, valve activity, and heart sounds. In this section, the focus will be on the left side of the heart (see Table 14.3). Identical events occur simultaneously on the right side of the heart; however, the pressures are lower.

There are four separate events that occur during the cardiac cycle:

1. Ventricular filling (diastole)
2. Isovolumetric contraction (systole)
3. Ejection (systole)
4. Isovolumetric relaxation (diastole)

14.6.1 Ventricular filling

This process occurs during ventricular diastole. When the ventricle has completely relaxed and pressure in the ventricle is lower than the pressure in the atrium, the AV valve opens. The pressure in the atrium at this time is greater than that of the ventricle due to the continuous return of blood from the veins. The initial phase of filling is rapid as blood had accumulated in the atrium prior to the opening of the AV valve. Once this valve opens, the accumulated blood rushes in. The second phase of filling is slower as blood continues to flow from the veins into the atrium and then into the ventricle. This phase of filling is referred to as *diastasis*. Up to this point, ventricular filling has occurred *passively*, and at rest, approximately 75% of the blood entering the ventricle does so in this manner. The third phase of ventricular filling results from *atrial contraction*. At this time, the remaining 25% of the blood is forced into the ventricle by this *active* process. The volume of blood in the ventricles at the end of the filling period is referred to as the *end-diastolic volume* (EDV) and is approximately 120 to 130 ml at rest. Note that during the entire diastolic filling period, the aortic (semilunar) valve is closed. The ventricular pressure during the filling phase is very low (0 to 10 mmHg). The pressure in the aorta during diastole is approximately 80 mmHg. Therefore, the aortic valve remains closed to prevent the backward flow of blood from the aorta into the ventricle during ventricular diastole.

14.6.2 Isovolumetric contraction

This process occurs during ventricular systole. When the ventricular myocardium begins to contract and squeeze down on the blood within the chamber, the pressure begins to increase rapidly. In fact, ventricular pressure is almost instantly greater than atrial pressure. As a result, the AV valve closes to prevent the backward flow of blood from the ventricle into the atrium during ventricular systole. The closure of this valve results in the first heart sound ("lub"). The ventricle continues its contraction and its buildup of pressure. However, there is a period of several milliseconds where ventricular pressure is climbing toward that of the aorta (from less than 10 mmHg up toward 80 mmHg). Until ventricular pressure exceeds aortic pressure, the aortic valve remains closed. As a result, both valves leading into and out of the chamber are closed, and this period is referred to as *isovolumetric contraction*.

During this phase, there is neither filling of the ventricle nor ejection of blood from the ventricle, so there is no change in blood volume.

14.6.3 Ejection

Eventually, the buildup of ventricular pressure overtakes the aortic pressure and the aortic valve is pushed open. At this point, *ejection*, or ventricular emptying, takes place. It is important to note that the chamber does not eject all of the blood within it. Some blood remains in the ventricle following contraction, and this volume, referred to as *end-systolic volume* (ESV), is approximately 50 to 60 ml at rest. Therefore, the volume of blood pumped out of each ventricle per beat, or the *stroke volume (SV)*, is about 70 ml in a healthy adult heart at rest.

14.6.4 Isovolumetric relaxation

After systole, the ventricles abruptly relax and the ventricular pressure decreases rapidly. Pressure in the aorta, which has peaked at 120 mmHg during systole, remains above 100 mmHg at this point. Therefore, the blood in the distended artery is immediately pushed back toward the ventricle down the pressure gradient. The backward movement of blood snaps the aortic valve shut. The closure of this valve results in the second heart sound ("dub"). During this portion of ventricular diastole, there is a period of several milliseconds where ventricular pressure is dissipating and falling back toward zero. Because atrial pressure is close to zero, the AV valve remains closed. Therefore, during this phase of *isovolumetric relaxation*, both valves leading into and out of the chamber are closed. As with isovolumetric contraction, there is no change in the blood volume of the ventricle during this phase of isovolumetric relaxation. When the ventricular pressure falls to a point where it is once again exceeded by atrial pressure, the AV valve opens, ventricular filling occurs, and the cardiac cycle begins again.

Due to the alternating phases of systole and diastole, the heart pumps blood intermittently. It contracts to pump the blood into the arteries and then it relaxes so it can once again fill with blood. However, capillary blood flow is not interrupted by this cycle as blood flow to the tissues is continuous. This steady blood flow is due to the elastic properties of the arterial walls. When the stroke volume is ejected into the arterial system, some of the blood is pushed forward toward the tissues. The remainder of the stroke volume is retained in the arteries. These large blood vessels are characterized by an abundance of collagen fibers and elastin fibers. These connective tissue fibers allow the arteries to be quite strong, capable of withstanding high pressures, but also reasonably distensible. The rapid addition of the stroke volume causes arterial distension, or stretch, resulting in the "storage" of a portion of this blood in these vessels. During diastole, when the heart relaxes, the arteries recoil and regain their original shape. This recoil squeezes down on

the stored blood and pushes it forward toward the tissues. Therefore, blood flow through the circulation is continuous during both ventricular systole and diastole.

Medical terminology

Annulus (ăn′ū-lŭs): Ring.
Arrhythmia (ā-rĭth′mē-ă): Irregular heartbeat.
Atherosclerosis (ăth″ĕr-ō-sklĕ-rō′sĭs): Common form of arteriosclerosis characterized by cholesterol-lipid-calcium deposits in the walls of arteries causing an occlusion or obstruction to blood flow.
Atrium (ā′trē-ŭm): Chamber of the heart that receives venous blood from either the systemic circulation (right atrium) or the pulmonary circulation (left atrium).
Automaticity (aw-tō-mă-tĭs′ĭ-tē): Refers to the ability to be self-excitable and generate a heartbeat without nervous stimulation.
Bradycardia (brăd″ē-kăr′dē-ă): Heart rate less than 60 beats per minute.
Bundle of His (bŭn′dĕl of hĭs): AV bundle, portion of the specialized electrical conduction system found in the interventricular septum.
Cardiac output (kăr-dē′ak owt′put): Volume of blood pumped by the heart per minute.
Chordae tendineae (kor′dă tĕnd-ĭn-ē′ă): Tendinous cords that connect the flaps of the atrioventricular valves to the papillary muscles.
Cusp (kŭsp): In the heart, a flap of a valve.
Desmosome (dĕs′mō-sōm): Intercellular junction that provides attachment and structural support between the cells.
Diastasis (dī-ăs′tă-sĭs): Following the period of rapid ventricular filling, the period of slow inflow of blood from the atria to the ventricles.
Diastole (dī-ăs′tō-lē): Period of relaxation and filling with blood in the heart.
End-diastolic volume (ĕnd-dī-ă-stŏl′ĭk vŏl′yŭm): Volume of blood in the ventricle of the heart just prior to contraction.
Endocardium (ĕn″dō-kăr′dē-ŭm): Endothelial membrane that lines the chambers of the heart.
Endothelium (ĕn″dō-thă′lē- ŭm): Squamous epithelium that lines blood and lymphatic vessels.
End-systolic volume (ĕnd-sĭ-stŏl′ĭk vŏl′yŭm): Volume of blood remaining in the ventricle of the heart following ejection.
Epicardium (ĕp″ĭ-kăr′dē-ŭm): Membrane on the surface of the heart.
Infarct (ĭn′fărkt): Region of necrosis in a tissue resulting from inadequate blood flow or ischemia.
Intercalated disc (ĭn-tĕr′kă-lā-tĕd dĭsk): Occurring at the junction of two myocardial cells, site of transmission of the electrical impulse.
Interventricular (ĭn″tĕr-vĕn-trĭk′ū-lăr): Between the two ventricles.
Isovolumetric (ī″sō-vŏl-yū-mĕt′rĭk): Occurring without a change in volume.

Myocardium (mī-ō-kăr′dē-ŭm): Middle layer of the wall of the heart, composed of cardiac muscle.

Myogenic (mī′ō-jĕn-ĭk): Originating in muscle; controlled by inherent properties of the muscle instead of nerves; describes the self-excitation of cardiac muscle and single-unit smooth muscle cells.

Neurogenic (nū-rō-jĕn′ĭk): Originating in nervous tissue; controlled by nervous factors.

Papillary muscle (păp′ĭ-lăr-ē mŭs′ĕl): Cone-shaped muscle arising from the floor of each ventricle and attaching to the chordae tendineae.

Pericardium (pĕr″ĭ-kăr-dē-ŭm): Serous membrane enclosing the heart.

Prolapse (prō′lăps): Dropping down of an organ or body part.

Purkinje fiber (pŭr-kĭn′jē fī′bĕr): Specialized cardiac muscle cell forming the last portion of the electrical conducting system.

Septum (sĕp′tŭm): Wall dividing two cavities.

Stroke volume (strōk vŏl′yŭm): Volume of blood ejected from each ventricle of the heart per beat.

Supraventricular (soo″pră-vĕn-trĭk′ū-lăr): Located above the heart ventricles, specifically, in the atria.

Syncytium (sĭn-sĭt′ē-ŭm): Cells of a tissue interconnected electrically by way of gap junctions that function together as a single unit.

Systole (sĭs′tō-lē): Period of contraction and ejection of blood from the heart.

Tachycardia (tăk″ē-kăr′dē-ă): Heart rate greater than 100 beats per minute.

Tetanus (tĕt′ă-nŭs): Smooth, sustained maximal contraction of a muscle.

Ventricle (vĕn′trĭk-l): Chamber of the heart that delivers blood to either the systemic circulation (left ventricle) or the pulmonary circulation (right ventricle).

Bibliography

Du Pont Pharma, *Introduction to Nuclear Cardiology*, 3rd ed., Du Pont Pharma Radiopharmaceuticals, North Billerica, MA, 1993.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Golan, D. E., Tashjian, A. H., Armstrong, E. J., Galanter, J. M., Armstrong, A. W., Arnout, R. A., and Rose, H., Pharmacology of Cardiac Rhythm, in *Principles of Pharmacology, The Pathophysiologic Basis of Drug Therapy,* Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 17.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology*, 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hondeghem, L. M., and Roden, D. M., Agents Used in Cardiac Arrhythmias, in *Basic and Clinical Pharmacology,* 8th ed., Katzung, B. G., Ed., McGraw-Hill, New York, 2001, chap. 14.

Katzung, B. G., and Parmley, W. W., Cardiac Glycosides and Other Drugs Used in Congestive Heart Failure, in *Basic and Clinical Pharmacology,* 8th ed., Katzung, B. G., Ed., McGraw-Hill, New York, 2001, chap. 13.

Levy, M. N., and Pappano, A. J., *Cardiovascular Physiology*, 9th ed., Mosby, St. Louis, MO, 2007.

Lilly, L. S., *Pathophysiology of Heart Disease*, Lea & Febiger, Malvern, PA, 1993, chaps. 11 and 17.

Michel, T., Treatment of Myocardial Ischemia, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 31.

Rhoades, R., and Pflanzer, R., *Human Physiology*, 4th ed., Thomson Learning, Pacific Grove, CA, 2003.

Rocco, T. P., and Fang, J. C., Pharmacotherapy of Congestive Heart Failure, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 33.

Roden, D. M., Antiarrhythmic Drugs, in *Goodman and Gilman's, The Pharmacological Basis of Therapeutics*, 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 34.

Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Smith, E. R., Pathophysiology of Cardiac Electrical Disturbances, in *Physiopathology of the Cardiovascular System*, Alpert, J. S., Ed., Little, Brown, Boston, MA, 1984, chap. 14.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

White Winters, J. M., Cardiac Conduction and Rhythm Disorders, in *Pathophysiology, Concepts of Altered Health States*, 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 27.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter fifteen

Cardiac output

Study objectives

- Describe the factors that determine cardiac output
- Distinguish between cardiac output, cardiac reserve, and cardiac index
- Discuss the factors that control heart rate
- Distinguish between the terms *chronotropic* and *inotropic*
- Discuss the factors that control stroke volume
- Distinguish between preload and afterload
- Describe the Frank-Starling Law of the Heart
- Understand how the cardiac function curve is generated
- Explain the mechanism of action of diuretics in congestive heart failure and hypertension
- Define *ejection fraction*
- Describe how cardiac output varies in a sedentary individual versus an endurance-trained athlete

15.1 Introduction

The primary function of the heart is to deliver a sufficient volume of blood (oxygen and nutrients, and so forth) to the tissues so that they may carry out their functions effectively. As the metabolic activity of a tissue varies, so will its need for blood. An important factor involved in meeting this demand is *cardiac output* (CO) or the volume of blood pumped into the aorta per minute. Cardiac output is determined by heart rate multiplied by stroke volume:

$$\text{Cardiac output (CO)} = \text{Heart rate (HR)} \times \text{Stroke volume (SV)}$$

An average adult at rest may have a heart rate of 70 beats per minute and a stroke volume of 70 milliliters per beat. In this case, the cardiac output would be

$$\begin{aligned} \text{CO} &= 70 \text{ beats/min} \times 70 \text{ ml/beat} \\ &= 4900 \text{ ml/min} \\ &\approx 5 \text{ L/min} \end{aligned}$$

Table 15.1 Factors That Affect Cardiac Output

Heart Rate
• Autonomic Nervous System
• Catecholamines (Epinephrine and Norepinephrine)
• Body Temperature
Stroke Volume
• Length of Diastole
• Venous Return (Preload)
• Contractility of the Myocardium
• Afterload
• Heart Rate
Miscellaneous
• Level of Activity
• Body Size
• Age

This is approximately equal to the total volume of blood in the body. Therefore, the entire blood volume is pumped by the heart each minute.

Many factors are involved in determining the cardiac output in a given individual (see Table 15.1):

- Level of activity
- Size of the body
- Age

A primary determinant is the *level of activity* of the body. During intense exercise in an average sedentary person, cardiac output may increase to 18 to 20 L/min. In a trained athlete, the increase in cardiac output is even greater and may be as much as 30 to 35 L/min. *Cardiac reserve* is the difference between the cardiac output at rest and the maximum volume of blood that the heart is capable of pumping per minute. The effect of endurance training is to significantly increase cardiac reserve so that a greater volume of blood can be pumped to the working muscles. In this way, exercise performance is maximized and fatigue of the muscles is delayed. On the other hand, patients with heart conditions, such as congestive heart failure or mitral valve stenosis, are not able to increase their cardiac output as much, if at all, during exercise. Therefore, these patients are forced to limit their level of exertion as the disease process progresses.

The *size of the body* is another factor that determines cardiac output. Healthy, young men have a cardiac output of about 5.5 to 6.0 L/min, where the cardiac output in women averages 4.5 to 5.0 L/min. This difference does

not involve gender, per se, but instead involves the mass of body tissue that must be perfused with blood. *Cardiac index* normalizes cardiac output for body size and is calculated by the cardiac output per square meter of body surface area. An average person weighing 70 kilograms has a body surface area of approximately 1.7 square meters. Therefore,

$$\text{Cardiac index} = 5\text{L/min} \div 1.7\ \text{m}^2$$

$$= 3\ \text{L/min/m}^2$$

Cardiac output also varies with *age*. Expressed as cardiac index, it rapidly rises to a peak of more than 4 L/min/m^2 at age 10 and then steadily declines to about 2.4 L/min/m^2 at the age of 80. This decrease in cardiac output is a function of overall metabolic activity and is, therefore, indicative of declining activity with age.

15.2 Control of heart rate

Heart rate varies considerably depending upon a number of variables. In normal adults at rest, the typical average heart rate is about 70 beats per minute; however, in children the resting heart rate is much greater. Heart rate will increase substantially (greater than 100 beats per minute) during emotional excitement and exercise, and it will decrease by 10 to 20 beats per minute during sleep. In endurance athletes, the resting heart rate may be 50 beats per minute or lower. This condition, referred to as *training-induced bradycardia*, is beneficial as it reduces the workload of the heart.

Mechanisms that alter heart rate have a *chronotropic effect* (*chrono* = time). A factor that increases heart rate is said to have a *positive chronotropic effect,* and a factor that decreases heart rate is said to have a *negative chronotropic effect*. In this section, three factors that control heart rate will be discussed:

1. Autonomic nervous system
2. Catecholamines
3. Body temperature

The *autonomic nervous system* exerts the primary control on heart rate. Because the sympathetic system and the parasympathetic system have antagonistic effects on the heart, the heart rate at any given moment results from the balance or the sum of their inputs. The sinoatrial (SA) node, which is the pacemaker of the heart, determining the rate of spontaneous depolarization, and the atrioventricular (AV) node are innervated by both the sympathetic and the parasympathetic systems. The specialized ventricular conduction pathway and ventricular muscle are innervated by the sympathetic system only.

Sympathetic stimulation increases heart rate (see Table 15.2). Norepinephrine, the neurotransmitter released from sympathetic nerves, binds to the β-adrenergic receptors in the heart and causes the following effects:

- Increased rate of discharge of the SA node.
- Increased rate of conduction through the AV node.
- Increased rate of conduction through the bundle of His and the Purkinje fibers.

The mechanism of these effects involves the enhanced rate of depolarization of the cells in these tissues due to increased sodium current through the F-type sodium channels (see Chapter 14). With more Na^+ ions entering the cell, the inside of the cell becomes less negative and approaches threshold more rapidly. In this way, action potentials are generated faster, and they travel through the conduction pathway more quickly so that the heart can generate more heartbeats per minute (see Figure 15.1).

Parasympathetic stimulation decreases heart rate (see Table 15.2). Acetylcholine is the neurotransmitter released from the vagus nerves, which are the parasympathetic nerves to the heart. Acetylcholine binds to muscarinic receptors and causes a decreased rate of discharge of the SA node and a decreased rate of conduction through the AV node.

The mechanism of these effects involves the increased permeability to potassium. The enhanced efflux of K^+ ions has two effects on the action potential of the SA node. First, the cells become hyperpolarized so that the membrane potential is farther away from threshold (from a normal resting potential of –55 mV down toward –65 mV). As a result, greater depolarization is now needed to reach threshold and generate an action potential. Second, the rate of depolarization during the pacemaker potential is reduced. The outward movement of positively charged K^+ ions opposes the depolarizing effect of Na^+ ion and Ca^{++} ion influx. In this way, action potentials are developed more slowly and fewer heartbeats are generated per minute (see Figure 15.1).

At rest, the parasympathetic system exerts the predominant effect on the SA node and, therefore, on the heart rate. In a denervated heart, such

Table 15.2 Effects of Autonomic Nerves on the Heart

Heart Tissue	Sympathetic Nerves	Parasympathetic Nerves
Sinoatrial (SA) node	↑ Heart rate (primarily via right sympathetic nerves)	↓ Heart rate (primarily via right vagus nerve)
Atrioventricular (AV) node	↑ Conduction rate	↓ Conduction rate (primarily via left vagus nerve)
Atrial muscle	↑ Contractility	↓ Contractility
Ventricular muscle	↑ Contractility (primarily via left sympathetic nerves)	No significant effect

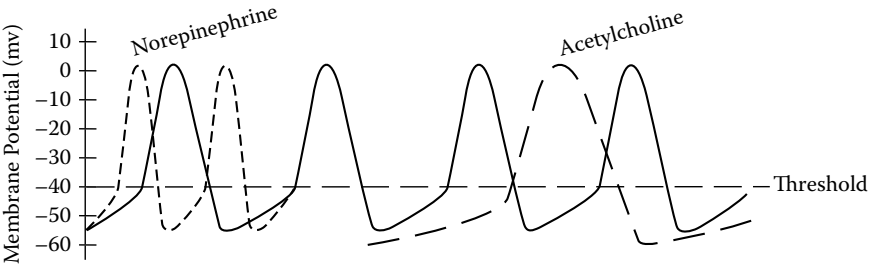


Figure 15.1 Effect of autonomic nervous system stimulation on action potentials of the sinoatrial (SA) node. A normal action potential generated by the SA node under resting conditions is represented by the solid line. The positive chronotropic effect (increased heart rate) of norepinephrine released from sympathetic nerve fibers is illustrated by the short dashed line; and the negative chronotropic effect (decreased heart rate) of acetylcholine released from parasympathetic nerve fibers is illustrated by the long dashed line.

as a transplanted heart, the resting heart rate is 100 beats per minute. This indicates that the SA node, without any input from the autonomic nervous system, has an inherent rate of depolarization of 100 beats per minute. However, the intact or fully innervated heart generates only 70 beats per minute. Therefore, it is evident that the rate of spontaneous discharge by the SA node is suppressed by the influence of the parasympathetic system. In contrast, the sympathetic system dominates during exercise.

It is generally accepted that the maximum heart rate an individual may achieve during exercise can be calculated by the following formula:

$$HR_{max} = 220 - age$$

Therefore, the range of heart rates from 185 to 195 beats/minute may be attained in individuals between 25 and 35 years of age. The maximum heart rate that may be attained in an individual 55 years of age is significantly reduced and is about 165 beats/minute. Many factors may affect maximum heart rate including level of fitness and various health considerations. Average values for maximum heart rate may be found in Table 15.3.

The activity of the autonomic nerves to the heart is regulated by the vasomotor center in the brainstem. Inputs to the vasomotor center, which subsequently influence autonomic nervous system function, are discussed in detail in Chapter 16.

The second factor that exerts control on heart rate is the release of the *catecholamines*, epinephrine and norepinephrine, from the adrenal medulla. The circulating catecholamines have the same effect on heart rate as direct sympathetic stimulation, which is to increase heart rate. In fact, in the intact (innervated) heart, the effect of the catecholamines serves to supplement this direct effect. In a denervated heart, the circulating catecholamines serve to

Table 15.3 Average Maximum Heart Rates and Target Heart Rates

Age	Average Maximum Heart Rate (100%)	Target Heart Rate Zone (50% to 70% max)
20 years old	200 beats/minute	100–150 beats/minute
30 years old	190 beats/minute	95–142 beats/minute
40 years old	180 beats/minute	90–135 beats/minute
50 years old	170 beats/minute	85–127 beats/minute
60 years old	160 beats/minute	80–120 beats/minute
70 years old	150 beats/minute	75–113 beats/minute

replace the effect of direct sympathetic stimulation. In this way, patients who have had a heart transplant may still increase their heart rate during exercise.

Body temperature also affects heart rate by altering the rate of discharge of the SA node. An increase of 1°F in body temperature results in an increase in heart rate of about 10 beats/minute. Therefore, the increase in body temperature during a fever or that which accompanies exercise serves to increase heart rate and, as a result, cardiac output. This enhanced pumping action of the heart delivers more blood to the tissues and supports the increased metabolic activity associated with these conditions.

15.3 Control of stroke volume

Many factors contribute to the regulation of stroke volume. Factors discussed in this section include

- Length of diastole
- Venous return (preload)
- Contractility of the myocardium
- Afterload
- Heart rate

Two important concepts to keep in mind throughout this discussion are the following: "the heart can only pump what it gets" and "a healthy heart pumps all of the blood returned to it." The SA node may generate a heartbeat and cause the ventricles to contract; however, these chambers must be properly filled with blood in order for this activity to be effective. On the other hand, the volume of blood that returns to the heart per minute may vary considerably. The heart has an intrinsic ability to alter its strength of contraction in order to accommodate these changes in volume.

Diastole is the period in the cardiac cycle in which relaxation of the myocardium and ventricular filling take place. In an individual with a resting heart rate of 75 beats/minute, the length of the cardiac cycle is 0.8 seconds and the length of ventricular diastole is 0.5 seconds. As mentioned in the

previous chapter, the end-diastolic volume is approximately 130 ml, and the resulting stroke volume is about 70 ml. Consider the case where the heart rate is increased. Given that cardiac output is determined by heart rate multiplied by stroke volume, an increase in either of these variables should result in an increase in cardiac output. In general, this is quite true. However, the effect of increased heart rate on cardiac output is limited by its effect on the length of diastole. As the heart rate increases, then the length of the cardiac cycle and, therefore, the length of diastole or the time for filling will decrease. At very high heart rates, this may result in a decrease in ventricular filling or end-diastolic volume, a decrease in stroke volume, and a decrease in cardiac output. Once again, "the heart can only pump what it gets." If there is inadequate time for filling, then despite an increase in heart rate, cardiac output will actually decrease. This explains why the maximal heart rate during exercise is about 185 to 195 beats per minute in all individuals. Beyond this rate, the ventricles are unable to properly fill with blood, and the positive effect of increased heart rate on cardiac output is lost.

Venous return is defined as the volume of blood returned to the right atrium per minute. Assuming a constant heart rate and, therefore, a constant length of diastole, then an increase in venous return, or the rate of blood flow into the heart, will increase ventricular filling, increase end-diastolic volume, increase stroke volume, and increase cardiac output. Ventricular blood volume prior to contraction is also referred to as *preload.* Once again, "a healthy heart pumps all of the blood returned to it." In fact, the cardiac output is equal to the venous return. The heart has an inherent, self-regulating mechanism by which it can alter its force of contraction based upon the volume of blood that flows into it. This *intrinsic mechanism*, the *Frank-Starling Law of the Heart*, states that when ventricular filling is increased, then this increased volume of blood stretches the walls of these chambers and, as a result, the ventricles contract more forcefully. The stronger contraction results in a larger stroke volume. This concept is illustrated by the *cardiac function curve* (see Figure 15.2). As end-diastolic volume increases, then stroke volume and, therefore, cardiac output increase. This phenomenon is based on the length–tension relationship of cardiac muscle. Recall in the discussion of skeletal muscle in Chapter 12 that the resting length of the sarcomere determines the amount of tension generated by the muscle upon stimulation. Due to their attachments to the bones, the resting lengths of the skeletal muscles do not vary greatly. Therefore, the sarcomeres are normally at their optimal length of 2.2 μ, which results in maximal tension development. At this point, the overlap of the actin and myosin filaments in the sarcomere is such that the greatest amount of cross-bridge cycling to generate tension takes place.

The myocardium of the heart, however, is not limited by attachment to any bones. Therefore, the resting length of the sarcomeres of these muscle cells may vary substantially due to changes in ventricular filling. Interestingly, at a normal resting end-diastolic volume of 130 ml, the cardiac muscle fiber length is *less than the optimal length.* Therefore, as filling increases, the

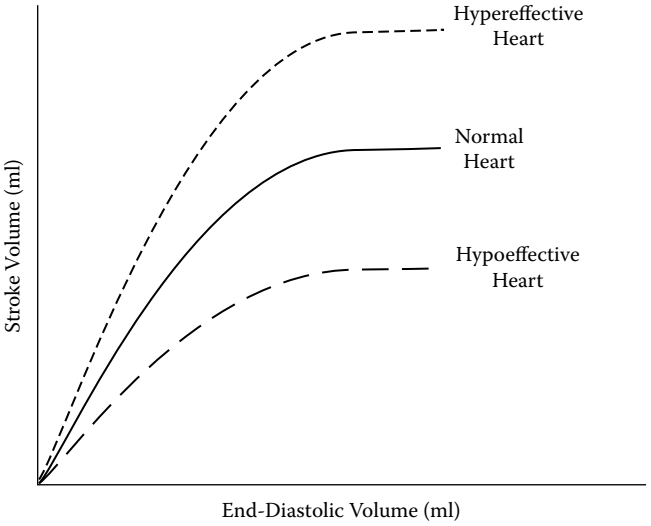


Figure 15.2 Cardiac function curve. This curve illustrates the stroke volume pumped by the heart for a given end-diastolic volume within the ventricle. As ventricular filling (end-diastolic volume) increases, stroke volume increases. Factors causing a positive inotropic effect (increased contractility) result in a greater stroke volume for a given amount of filling compared to the normal heart. This "hypereffective" heart is illustrated by the short dashed line that is shifted to the left of that of the normal heart. Sympathetic stimulation, the catecholamines, and some medications such as the cardiac glycosides increase contractility. Factors causing a negative inotropic effect (decreased contractility) result in a reduced stroke volume for a given amount of filling. This "hypoeffective" heart is illustrated by the long dashed line that is shifted to the right of that of the normal heart. Damage to the myocardium and some medications, such as β-adrenergic receptor antagonists and calcium channel blockers, decrease contractility.

muscle fibers and their component sarcomeres are stretched and move closer to the optimal length for cross-bridge cycling and tension development. This results in a stronger ventricular contraction and an increase in stroke volume. In other words, a healthy heart operates on the ascending portion of the cardiac function curve, such that an increase in preload results in an increase in stroke volume.

Stimulation of the ventricular myocardium by the sympathetic system will also increase stroke volume by increasing the *contractility* of the muscle. At any given end-diastolic volume, norepinephrine released from the sympathetic nerves to the heart will cause a more forceful contraction resulting in the ejection of more blood from the ventricles or an increase in stroke volume and an increase in cardiac output (see Figure 15.2). In other words, the cardiac function curve shifts to the left. Epinephrine released from the

PHARMACY APPLICATION:
DIURETICS AND CARDIAC OUTPUT

Diuretics are a group of therapeutic agents designed to reduce the volume of body fluids. Their mechanism of action is at the level of the kidney and involves an increase in the excretion of Na^+ ions and Cl^- ions and, consequently, an increase in urine production. As discussed in Chapter 2, sodium is the predominant extracellular cation, and due to its osmotic effects, it is a primary determinant of extracellular fluid volume. Therefore, if more sodium is excreted in the urine, then more water is also lost, which reduces the volume of the extracellular fluids including the plasma. As plasma volume decreases, then less blood is available for ventricular filling.

This reduction in preload is, in some cases, beneficial to patients experiencing heart failure. Unlike a healthy heart, a failing heart is unable to pump all of the blood returned to it. Instead, the blood dams up and overfills the chambers of the heart. This results in congestion and increased pressures in the heart and the venous system and the formation of peripheral edema. Because the failing heart is operating on the flat portion of a depressed cardiac function curve (see Figure 16.2), treatment with diuretics will relieve the congestion and edema but have little effect on stroke volume and cardiac output.

adrenal medulla has the same effect on contractility as direct sympathetic stimulation. The mechanism involves the stimulation of β-adrenergic receptors and the subsequent increase in the permeability to calcium. An increase in intracellular calcium results in increased cross-bridge cycling and greater tension development. Sympathetic stimulation, the circulating catecholamines, or any other factor that increases contractility has a *positive inotropic effect* on the heart. Therapeutic agents, such as β-adrenergic receptor blockers and calcium channel blockers, which inhibit calcium influx and, therefore, reduce contractility, have a *negative inotropic effect*.

The ventricles are very sparsely innervated by the vagus nerves. Therefore, parasympathetic stimulation to the heart has little or no significant effect on contractility and stroke volume.

The contractility of the myocardium determines the *ejection fraction* of the heart which is the ratio of the volume of blood ejected from the left ventricle per beat (stroke volume [SV]) to the volume of blood in the left ventricle at the end of diastole (end-diastolic volume [EDV]):

$$\text{Ejection fraction} = \frac{\text{SV}}{\text{EDV}}$$

Under normal resting conditions where the end-diastolic volume is 120 to 130 ml and the stroke volume is 70 ml/beat, the ejection fraction is 55% to 60%:

$$\text{Ejection fraction} = \frac{70 \text{ ml/beat}}{120 \text{ ml}} = 58\%$$

During exercise when sympathetic stimulation to the heart is increased, the ejection fraction may increase to more than 80% contributing to a greater stroke volume and a greater cardiac output.

Another factor determining cardiac performance is *afterload* or the diastolic blood pressure in the artery leading from the ventricle. Physiologically, afterload is determined primarily by vascular resistance. Vasoconstriction due to sympathetic nervous stimulation or compression of blood vessels due to skeletal muscle contraction will increase vascular resistance and, therefore, increase afterload.

When the ventricle contracts, it must develop a pressure greater than that in the associated artery in order to push open the semilunar valve and eject the blood (see Chapter 14). Typically, diastolic blood pressure in the aorta is 80 mmHg. Therefore, the left ventricle must develop a pressure slightly greater than 80 mmHg to open the aortic valve and eject the stroke volume. A dynamic exercise, such as running, may cause only a small increase in diastolic pressure (up to 90 mmHg). However, a resistance exercise, such as weight lifting, has a much greater impact on blood pressure, and diastolic pressure may be as high as 150 to 160 mmHg. A healthy heart can easily contract vigorously enough to overcome any increases in afterload associated with exercise or other types of physical exertion. In contrast, a diseased heart or one weakened by advanced age may not be able to generate enough force to effectively overcome a significantly elevated afterload. In this case, the stroke volume and the cardiac output would be reduced. In addition, a sustained or chronic increase in afterload, as observed in patients with hypertension, will also have a detrimental effect on cardiac workload. Initially, the left ventricle will hypertrophy, and the chamber walls will become thicker and stronger to compensate for this excess workload. This form of ventricular enlargement is referred to as *concentric* hypertrophy. However, eventually the balance between the oxygen supply and the oxygen demand of the heart is disrupted, leading to a decreased stroke volume, a decreased cardiac output, and heart failure.

In summary, afterload is determined largely by vascular resistance. An increase in vascular resistance leads to an increase in afterload. Stroke volume is inversely proportional to afterload. An increase in afterload may lead to a decrease in stroke volume.

Changes in *heart rate* also affect the contractility of the heart. As heart rate increases, so does ventricular contractility. The mechanism of this effect

involves the gradual increase of intracellular calcium. When the electrical impulse stimulates the myocardial cell, the permeability to calcium is increased, and calcium enters the cytoplasm from the extracellular fluid as well as from the sarcoplasmic reticulum, allowing it to contract. Between beats, the calcium is removed from the intracellular fluid and the muscle relaxes. When the heart rate is increased, the periods of calcium influx occur more frequently and the time for calcium removal is reduced. The net effect is an increase in intracellular calcium, an increase in the number of cross-bridges cycling, and an increase in tension development.

PHARMACY APPLICATION:
CARDIAC GLYCOSIDES AND CARDIAC OUTPUT

A patient is considered to be in heart failure when the cardiac output is insufficient to meet the metabolic demands of his or her body. One way to improve cardiac output is to enhance myocardial contractility which will increase stroke volume. The cardiac glycosides, including digoxin, have been used for many years to treat heart failure due to their positive inotropic effect. Digoxin binds to and inhibits the Na^+-K^+ ATPase in the myocardial cell membrane which ultimately leads to an increase in the intracellular concentration of calcium. As described previously, this increase in calcium may help to enhance myocardial contractility and, therefore, cardiac output.

15.4 Effect of exercise on cardiac output

Endurance training, such as running, alters the baseline or tonic activity of the autonomic nervous system. In the trained athlete, the dominance of the parasympathetic system is even greater than it is in the sedentary individual, resulting in a training-induced bradycardia. The resting heart rate is 70 to 75 beats/minute in a sedentary person, and the resting heart rate in the trained athlete may be 50 beats/min or lower. Due to the decrease in heart rate in these individuals, the length of diastole is increased. Assuming a constant rate of venous return, this longer filling period results in a greater end-diastolic volume and an increased stroke volume (40% to 50% greater in the elite athlete compared to the untrained individual). Therefore, at rest when their body's metabolic demands are similar, the cardiac output in the sedentary person and the athlete are also similar (see Table 15.4).

During exercise, the cardiac output increases substantially to meet the increased metabolic demand of the working muscles. However, endurance training results in significantly greater increases in cardiac output which improves oxygen and nutrient delivery to the working muscles (18 to 20 L/min in sedentary individuals; 30 to 35 L/min in trained athletes). As a result, exercise performance is enhanced, and fatigue is delayed. In order

Table 15.4 Effect of Exercise on Cardiac Output

	Sedentary Individual	Endurance-Trained Athlete
Rest		
Heart Rate	71 beats/minute	50 beats/minute
Stroke Volume	70 ml/beat	100 ml/beat
Cardiac Output	5,000 ml/min	5,000 ml/beat
Exercise		
Heart Rate	195 beats/minute	195 beats/minute
Stroke Volume	100 ml/beat	160 ml/beat
Cardiac Output	19,500 ml/min	31,200 ml/min

to increase cardiac output, both heart rate and stroke volume are increased. The maximum heart rate in all individuals is about 185 to 195 beats/minute. Therefore, the difference in cardiac output in athletically trained versus untrained people during exercise involves stroke volume. Stroke volume may increase as much as 50% to 60% during exercise. Because the athlete has a much larger stroke volume at rest, the increase in stroke volume during exercise is that much greater (see Table 15.4). In this way, even with a similar maximal heart rate, the endurance-trained athlete pumps a significantly greater volume of blood per minute. In order to accommodate these larger stroke volumes, the ventricles of these athletes hypertrophy such that the chambers become larger and increase their diameters. This form of ventricular enlargement is referred to as *eccentric hypertrophy*.

Medical terminology

Afterload (ăf′tĕr-lōd): Force against which a ventricle of the heart contracts; diastolic pressure in the corresponding artery of a ventricle.

Cardiac index (kăr′dē-ăk ĭn′dĕks): Cardiac output divided by body surface area.

Cardiac output (kăr′dē-ăk out′poot): Volume of blood pumped by the heart per minute.

Cardiac reserve (kăr′dē-ăk rē-zĕrv′): The difference between the resting cardiac output and the maximal cardiac output.

Chronotropic (krŏn″ō-trŏp′ĭk): Refers to an agent or factor that influences the heart rate.

Contractility (kŏn-trăk-tĭl′ĭ-tē): The force with which ventricular ejection occurs.

Ejection fraction (ē-jĕk′shŭn frăk′shŭn): The percentage of blood forced out of the ventricle during systole.

Inotropic (ĭn″ō-trŏp′ĭk): Refers to an agent or factor that influences the force of muscular contraction.

Preload (prē-lōd): Degree of stretch of the ventricles of the heart just prior to contraction; end-diastolic volume (EDV).

Venous return (vē′nŭs rē-tĕrn′): Volume of blood returning to the right atrium of the heart per minute.

Bibliography

American Heart Association, 2002, www.justmove.org/fitnessnews/healthf.cfm?Target=hartrates.html.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hoffman, B. B., Therapy of Hypertension, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 32.

Jackson, E. K., Diuretics, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 28.

Levy, M. N., and Pappano, A. J., *Cardiovascular Physiology,* 9th ed., Mosby, St. Louis, MO, 2007.

McArdle, W. D., Katch, F. I., and Katch, V. L., *Essentials of Exercise Physiology,* 2nd ed., Lippincott Williams & Wilkins, Baltimore, MD, 2000.

Porth, C. M., *Pathophysiology, Concepts of Altered Health States,* 7th ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 28.

Rhoades, R., and Pflanzer, R., *Human Physiology,* Thomson Learning, Pacific Grove, CA, 2003.

Rocco, T. P., and Fang, J. C., Pharmacotherapy Congestive Heart Failure, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 33.

Rowell, L. B., *Human Cardiovascular Control*, Oxford University Press, New York, 1993.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology: The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter sixteen

The circulatory system

Study objectives

- List the functions of the circulatory system
- Describe how the rate of blood flow to individual tissues differs according to specific tissue function
- Explain the function of each component of the blood vessel wall
- Distinguish between arteries, arterioles, capillaries, and veins in terms of their anatomical characteristics and their functions
- Distinguish between diastolic pressure, systolic pressure, and pulse pressure
- Understand the method by which mean arterial pressure is calculated
- Describe how blood pressure changes as the blood flows through the circulatory system
- Understand Ohm's Law and describe the relationship between blood flow, blood pressure, and vascular resistance
- List the factors that affect vascular resistance and explain their physiological significance
- Explain why mean arterial pressure must be closely regulated
- Explain how the autonomic nervous system alters cardiac output, total peripheral resistance, and therefore, mean arterial pressure
- List the sources of input to the vasomotor center
- Describe the mechanism of action and the physiological significance of the baroreceptor reflex, the chemoreceptor reflex, and the low-pressure receptor reflex
- Indicate the source, the factors regulating the release, and the physiological significance of the following vasoconstrictors: catecholamines, angiotensin II, vasopressin, endothelin, thromboxane A_2, and serotonin
- Indicate the source, the factors regulating the release or activation, and the physiological significance of the following vasodilators: prostacyclin, nitric oxide, atrial natriuretic peptide, histamine, bradykinin, adenosine, and epinephrine
- Compare and contrast the compliance of systemic arteries and systemic veins
- List the specific blood reservoirs and their common characteristics
- Explain how blood volume, sympathetic stimulation of the veins, skeletal muscle activity, and respiratory activity influence venous return
- Describe the effects of gravity on the circulatory system

- Describe the mechanism of active hyperemia
- Define *autoregulation of blood flow*
- Explain how the myogenic mechanism causes autoregulation of blood flow
- Describe the effects of acute exercise on the circulatory system
- Explain how blood flow through capillaries is regulated by vasomotion
- Describe the physiological significance of the Starling Principle
- Explain how hydrostatic forces and osmotic forces regulate the bulk flow of fluid across the capillary wall
- Describe the four general conditions that can lead to edema formation

16.1 Introduction

The *circulatory system* carries out many important functions that contribute to homeostasis. It obtains oxygen from the lungs, nutrients from the gastrointestinal tract, and hormones from the endocrine glands, and it delivers these substances to the tissues that need them. Furthermore, it removes metabolic waste products, such as carbon dioxide, lactic acid, and urea, from the tissues. Finally, it contributes to the actions of the immune system by transporting antibodies and leukocytes to areas of infection. Overall, the circulatory system plays a vital role in the maintenance of optimal conditions for cell and tissue function.

All tissues are *perfused*. In other words, all tissues receive blood flow. The amount of blood that flows through each tissue, however, depends upon that tissue's function. For example, many tissues, such as the heart, the brain, and the skeletal muscles, receive a blood flow that is sufficient to supply their metabolic needs. When metabolic activity increases, as it does during exercise, blood flow to these tissues increases accordingly. Other tissues, however, receive a blood flow that is in significant excess of their metabolic needs. These tissues, including the kidneys, the organs of the digestive system, and the skin, have important homeostatic functions. Among other vital activities, the kidneys filter the blood and remove waste products, the organs of the digestive system absorb nutrients into the blood, and thermoregulation involves the control of blood flow to the body surface where heat is eliminated. These functions are carried out most effectively and efficiently when the involved tissues receive an abundant blood flow. Under normal resting conditions, the kidneys, which account for only 1% of the body's weight, receive 20% of the cardiac output (CO), the gastrointestinal tract receives approximately 27% of the CO, and the skin receives 6% to 15% of the blood pumped by the heart per minute. Because these tissues receive more blood than they need to support metabolic activity, they can easily tolerate a sustained decrease in blood flow. During exercise when the metabolic demand of the working skeletal muscles and the heart increases substantially, blood flow is directed away from the kidneys and the organs of the digestive system and toward the skeletal muscles and the heart.

16.2 The blood vessels

The walls of the blood vessels may contain varying amounts of fibrous tissue, elastic tissue, and smooth muscle. All blood vessels are lined with a single layer of endothelial cells forming the endothelium. The *fibrous connective tissue* provides structural support and stiffens the vessel. The *elastic connective tissue* allows vessels to expand and hold more blood. It also causes the vessels to recoil and exert pressure on the blood within the vessels which pushes this blood forward. Most blood vessels contain *smooth muscle*. This muscle is arranged in either circular or spiral layers. Therefore, contraction of vascular smooth muscle, or *vasoconstriction*, narrows the diameter of the vessel and decreases the flow of blood through it. Relaxation of vascular smooth muscle, or *vasodilation*, widens the diameter of the vessel and increases the flow of blood through it. The smooth muscle of the vessel is innervated by the autonomic nervous system and is, therefore, physiologically regulated. Furthermore, this is where endogenous vasoactive substances and pharmacological agents exert their effects. The *endothelium* has several important physiological functions including contributing to the regulation of blood pressure, blood vessel growth, and the exchange of materials between the blood and the interstitial fluid of the tissues (see Table 16.1).

The circulatory system is composed of several anatomically and functionally distinct blood vessels:

- Arteries
- Arterioles
- Capillaries
- Veins

Arteries carry blood away from the heart (see Figure 16.1). These vessels contain fibrous connective tissue that strengthens them and enables them to withstand the high blood pressures generated by the heart. In general, the arteries function as a system of *conduits*, or pipes, transporting the blood under high pressure toward the tissues. There is little smooth muscle and, therefore, little physiological regulation of vessel diameter in these vessels.

Table 16.1 Functions of Endothelial Cells

• Selectively permeable barrier between the vascular compartment and the interstitial fluid of the tissues.
• Lining of blood vessels that prevents adherence of blood cells and platelets to the vessel walls.
• Production of vasoconstrictor and vasodilator substances that act on underlying vascular smooth muscle.
• Role in angiogenesis, or new capillary growth.

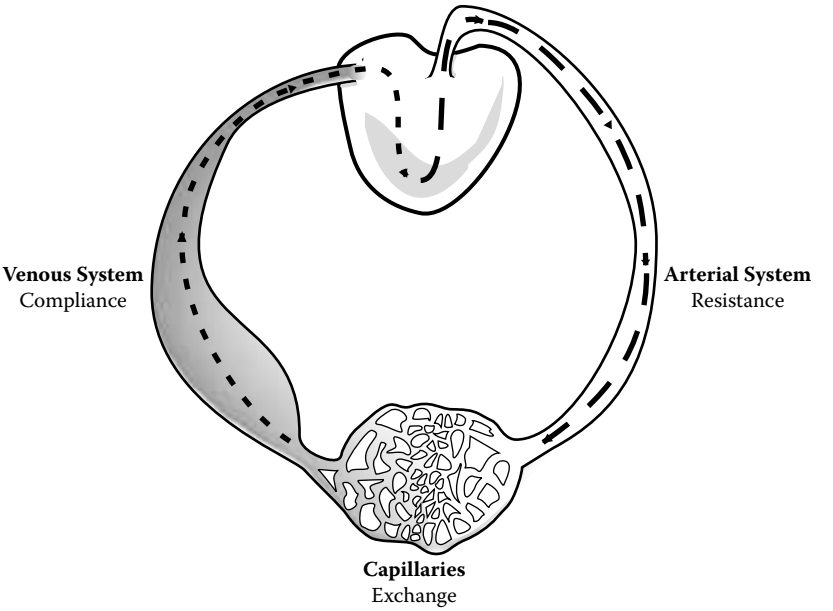


Figure 16.1 The circulatory system. Arteries carry blood away from the heart. The smallest arterial vessels, the arterioles, are composed mainly of smooth muscle and are the major resistance vessels in the circuit. The capillaries are the site of exchange between the blood and the tissues. Veins carry blood back toward the heart. The small veins are the major compliance vessels in the circuit and, under resting conditions, contain 64% of the blood volume.

Another noteworthy anatomical feature of the arteries is the presence of elastic connective tissue. When the heart contracts and ejects the blood, a portion of the stroke volume flows toward the capillaries. However, much of the stroke volume ejected during systole is retained in the distensible arteries. When the heart relaxes, the arteries recoil and exert pressure on the blood within them, forcing this "stored" blood to flow forward. In this way, a steady flow of blood toward the capillaries is maintained throughout the entire cardiac cycle.

As the arteries travel toward the peripheral organs and tissues, they branch and become smaller. Furthermore, the walls of the vessels become less elastic and more muscular. Finally, the smallest arterial vessels, the *arterioles*, are composed almost entirely of smooth muscle with a lining of endothelium. Therefore, depending upon the degree of constriction of the vascular smooth muscle, these vessels may alter their diameter and, consequently, their blood flow across a very wide range. For this reason, the arterioles are the major *resistance vessels* in the circulatory system. In fact, the primary function of arterioles is to regulate the distribution of the cardiac output and determine which tissues receive more blood and which tissues receive less blood

depending upon the tissue's and the body's needs. In addition, the resistance to blood flow in the arterioles contributes to the regulation of mean arterial pressure which is discussed in detail in Section 16.3.

From the arterioles, blood flows through the *capillaries*, the smallest vessels in the circulatory system. The capillaries are the *site of exchange* between the blood and the interstitial fluid surrounding the cells of the tissues. The primary mechanism of exchange is simple diffusion as substances move across the capillary walls "down" their concentration gradients, or from an area of high concentration to an area of low concentration. Several important factors influencing the process of diffusion include:

- Surface area of the barrier
- Thickness of the barrier
- Velocity of blood flow

As discussed in Chapter 2, Fick's Law of Diffusion states that as the surface area of a given barrier increases, so does the degree of diffusion. There are approximately 10 billion capillaries in the adult human body with a total exchange surface area of more than 6300 square meters — the equivalent of almost two football fields. Furthermore, most tissue cells are not more than 20 µm away from the nearest capillary.

The capillaries have the thinnest walls of all the blood vessels. They are composed of only a flat layer of endothelium that is one cell thick and is supported by a thin acellular matrix, referred to as the *basement membrane*. The total thickness of this barrier between the blood and the interstitial fluid is only 0.5 µm. As such, the anatomical characteristics of the capillaries that maximize the exchange surface area and minimize the thickness of the barrier render these vessels ideally suited for the exchange of materials by simple diffusion.

A third factor that influences diffusion across capillary walls is the velocity of blood flow. Although an individual capillary has a diameter of approximately 8 to 10 µm, the total cross-sectional area of all of the systemic capillaries taken together is 2500 to 3000 cm^2. As the cross-sectional area increases, then the velocity of blood flow decreases. In contrast to the aorta in which blood flows at a rate of 33 cm/sec, the capillaries have a velocity of blood flow of 0.33 mm/sec. This slow movement of blood facilitates the exchange of materials between the tissues and the vascular compartment.

Following the exchange of substances with the tissues, the blood begins its route back to the heart through the venous system. Blood flows from the capillaries into the *venules*. These small vessels consist mainly of a layer of endothelium and fibrous connective tissue. From the venules, the blood flows into *veins* that become larger as they travel toward the heart. As with the arteries, the walls of these vessels consist of a layer of endothelium, elastic connective tissue, smooth muscle, and fibrous connective tissue. However, veins have much thinner walls and wider diameters than the arteries they

accompany. These vessels are quite distensible and are capable of holding large volumes of blood at a very low pressure. For this reason, the veins are the major *compliance vessels* of the circulatory system (see Figure 16.1). In fact, approximately 64% of the blood volume is contained within the veins under resting conditions. During exercise, the pumping action of the contracting skeletal muscles and the constriction of smooth muscle in the walls of the veins force this blood toward the heart and increase venous return. Therefore, the veins are referred to as *blood reservoirs.* As such, they play an important role in the regulation of venous return and, consequently, cardiac output.

Another important anatomical characteristic of the veins is the presence of *valves*. These valves ensure the one-way flow of blood back toward the heart. They are most abundant in the lower limbs where the effects of gravity on the circulatory system are most prevalent and would tend to cause the pooling of blood in the feet and ankles.

Finally, the large veins and the *venae cavae* return the blood to the right atrium of the heart. As with the large arteries and the aorta, these vessels function primarily as conduits. There is little smooth muscle and, therefore, little physiological regulation of their diameter.

16.3 Blood pressure throughout the systemic circulation

The pressure generated by left ventricular contraction is the driving force for the flow of blood through the entire systemic circulation, from the aorta, through the tissues, and all the way back to the right atrium. The mean pressure in the aorta and the large arteries is typically very high (90 to 100 mmHg) due to the continual addition of blood to the system by the pumping action of the heart. However, this pressure is *pulsatile*. In other words, it fluctuates due to the alternating contraction and relaxation phases of the cardiac cycle. In a healthy resting adult, *systolic pressure* is approximately 120 mmHg and *diastolic pressure* is approximately 80 mmHg (see Figure 16.2). The *pulse pressure* (PP) is the difference between the systolic and the diastolic pressures:

$$\text{Pulse pressure} = \text{Systolic pressure} - \text{Diastolic pressure}$$

Therefore, using the average values of 120 mmHg (systolic) and 80 mmHg (diastolic), the pulse pressure may be determined:

$$\text{PP} = 120\ \text{mmHg} - 80\ \text{mmHg} = 40\ \text{mmHg}$$

The *mean arterial pressure* (MAP) is calculated as follows:

$$\text{MAP} = \text{Diastolic pressure} + 1/3\ (\text{Pulse pressure})$$

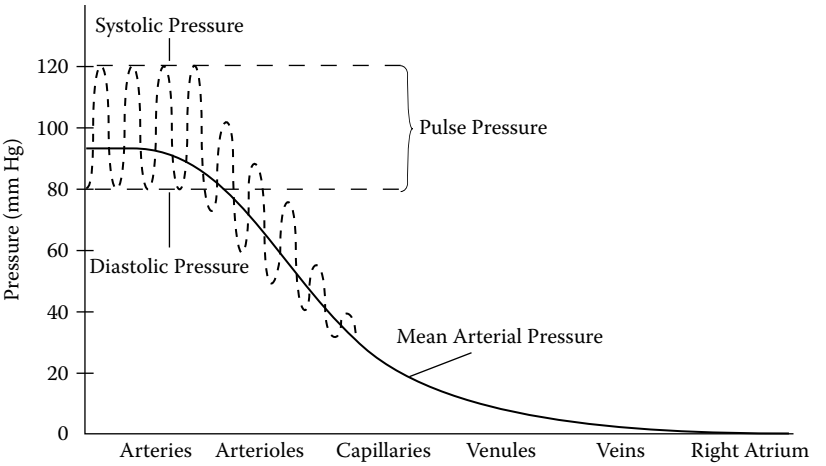


Figure 16.2 Pressures throughout the systemic circulation. At rest, blood pressure in the aorta and the other large arteries fluctuates between a low pressure of 80 mmHg during ventricular diastole and a high pressure of 120 mmHg during ventricular systole. The difference between the diastolic pressure and the systolic pressure is the pulse pressure. Mean arterial pressure in these arteries is approximately 93 mmHg. As the blood continues forward and flows through the arterioles, the pulse pressure is dampened. Because of the high resistance to blood flow in these vessels, the overall pressure drops dramatically. Furthermore, the fluctuations between diastolic pressure and systolic pressure are eliminated so that the blood pressure becomes nonpulsatile. Blood pressure continues to decline, although at a slower rate, as the blood flows through the capillaries and veins back toward the heart.

Therefore, using these same values, the MAP may be determined:

$$\text{MAP} = 80 \text{ mmHg} + 1/3\ (40 \text{ mmHg}) = 93 \text{ mmHg}$$

At rest, the MAP is closer to the diastolic pressure because the diastolic phase of the cardiac cycle lasts almost twice as long as the systolic phase. During exercise when heart rate increases and the length of diastole decreases, the systolic pressure contributes more to the MAP.

As the blood flows through the rest of the system, the pressure continually falls (see Figure 16.2). Furthermore, the pulsatile nature of the blood pressure is lost as the blood flows through the arterioles. The pulse pressure is damped out by the considerable resistance offered to blood flow by the arterioles. At the arteriolar end of the capillaries, the blood pressure is 30 to 35 mmHg, and at the venular end, the capillary pressure is approximately 10 mmHg. It is important that capillary pressure remains low so as to avoid the leakage of fluid out of the capillaries into the tissues. Venous pressure is approximately 6 to 8 mmHg, and pressure in the right atrium is close to zero.

16.4 Blood flow through a vessel

The *flow of blood through a vessel* is determined by two factors: the *pressure gradient* and *vascular resistance*.

The relationship between blood flow (Q, ml/min), the pressure gradient (ΔP, mmHg), and vascular resistance (R, mmHg/ml/min) is described by *Ohm's Law*:

$$Q = \frac{\Delta P}{R}$$

The *pressure gradient* is the difference between the pressure at the beginning of a blood vessel and the pressure at the end of the blood vessel. The inflow pressure is always greater than the outflow pressure, as substances, including blood and air, must flow "down" their pressure gradients — in other words, from an area of higher pressure to an area of lower pressure. The inflow pressure is initially generated by the contraction of the heart. As discussed previously, the blood pressure falls continuously as the blood flows through the circulatory system. This loss of driving pressure is due to the friction generated as the components of the flowing blood come into contact with the vessel wall as well as each other. Blood flow through a vessel is directly proportional to the pressure gradient; in other words, the greater the difference between the inflow pressure and the outflow pressure, the greater the flow of blood is through the vessel.

The second factor that determines the flow of blood through a vessel is *resistance*. In contrast to the pressure gradient, blood flow through a vessel is indirectly proportional to the resistance. In other words, resistance impedes or opposes blood flow. There are three factors that affect vascular resistance:

1. Blood viscosity
2. Vessel length
3. Vessel radius

Viscosity describes the friction developed between the molecules of a fluid as they interact with each other during flow. More simply put, the "thicker" the fluid, the greater is the viscosity of the fluid. Interestingly, blood is approximately three times more viscous than water.

Viscosity and resistance are directly proportional so that as the viscosity of the fluid increases, the resistance to flow increases. In the case of blood flow through the circulatory system, erythrocytes, or red blood cells, suspended in the blood are the primary factor determining viscosity. Blood cells exert a frictional drag against each other and against the wall of the blood vessel. *Hematocrit*, the percentage of the blood that consists of red blood cells, is 40% to 54% (average = 47%) for an adult male and 37% to 47% (average = 43%) for an adult female. Under normal physiological conditions, hematocrit and,

therefore, blood viscosity do not vary considerably within an individual. Only pathological conditions, such as chronic hypoxia, sickle cell anemia, and excess blood fibrinogen, may result in hyperviscosity and, consequently, impaired blood flow.

As mentioned, friction also develops as blood contacts the vessel wall while flowing through it. Therefore, the greater the vessel surface area in contact with the blood, the greater is the amount of friction developed and the greater the resistance to blood flow. Two factors determine the vessel surface area: the length of the vessel and the vessel radius.

The longer the vessel, the more the blood comes into contact with the vessel wall, and the greater the resistance. However, *vessel length* in the body remains constant. Therefore, as with blood viscosity, it is not a variable factor causing changes in resistance.

The most important physiological variable determining the resistance to blood flow is *vessel radius*. A given volume of blood comes into less contact with the wall of a vessel with a large radius compared to a vessel with a small radius. Therefore, as the radius of a vessel increases, the resistance to blood flow decreases. In other words, blood flows more readily through a larger vessel than it does through a smaller vessel.

Small changes in vessel radius result in significant changes in vascular resistance and in blood flow. This is because the resistance is inversely proportional to the fourth power of the radius:

$$R \propto \frac{1}{r^4}$$

If this equation is substituted into Ohm's Law, then blood flow may be calculated as follows:

$$Q = \frac{\Delta P}{1/r^4}$$

Assume two blood vessels of equal length each have a pressure gradient of 1 mmHg. However, blood vessel A has a radius of 1 mm and blood vessel B has a radius of 2 mm. The flow of blood through vessel A is 1 ml/min, and the flow of blood through vessel B is 16 ml/min. Simply doubling vessel radius causes a sixteenfold increase in blood flow.

As mentioned previously, the arterioles are the major resistance vessels in the circulatory system. Because the walls of these vessels contain primarily smooth muscle, they are capable of significant changes in their radius. Therefore, the regulation of blood flow to the tissues is carried out by the arterioles.

Ohm's Law may be rewritten to include the three factors that affect vascular resistance: blood viscosity (η), vessel length (L), and vessel radius (r). The following equation is known as *Poiseuille's Law*:

$$Q = \frac{\pi \Delta P r^4}{8 \eta L}$$

16.5 Regulation of arterial pressure

Mean arterial pressure (MAP) is the driving force for blood flow through the body's organs and tissues. The MAP must be closely monitored and regulated for several reasons. It must be high enough to provide a force sufficient to propel the blood through the entire systemic circuit: from the heart to the top of the head, to the tips of the toes, and back to the heart again. *Hypotension*, or a fall in blood pressure, may cause insufficient blood flow to the brain, causing dizziness and, perhaps, fainting. However, *hypertension*, or a blood pressure that is too high, may be detrimental to the cardiovascular system. An increase in diastolic pressure increases the afterload on the heart and increases cardiac workload. Furthermore, a chronic elevation in blood pressure increases the risk of various types of vascular damage, such as atherosclerosis and the rupture of small blood vessels. Atherosclerosis will often occur in arteries that supply the heart and the brain, and it impairs the flow of blood to these tissues. The rupture of small blood vessels allows fluid to move from the vascular compartment into the tissues resulting in edema formation. Several factors that may alter MAP are listed in Table 16.2.

Ohm's Law, which correlates the influence of blood pressure and vascular resistance on blood flow through a vessel ($Q = \Delta P/R$), may also be applied to blood flow through the entire systemic circulation, or cardiac output:

$$\text{Cardiac Output} = \frac{\text{Mean Arterial Pressure}}{\text{Total Peripheral Resistance}}$$

This equation can be reorganized to determine MAP:

$$\text{Mean Arterial Pressure} = \text{Cardiac Output} \times \text{Total Peripheral Resistance}$$

Total peripheral resistance (TPR) is the resistance to blood flow offered by all of the systemic vessels taken together, especially by the arterioles, which are the primary resistance vessels. Therefore, MAP is regulated by cardiac activity and vascular smooth muscle tone. Any change in CO or TPR causes a change in MAP. A summary of major cardiovascular principles may be found in Table 16.3.

The relationship between CO, TPR, and MAP may be considered further. Recall from the earlier discussion that MAP may be calculated from the systolic pressure (SP) and the diastolic pressure (DP):

Table 16.2 Factors That Alter Blood Pressure

Increase in Blood Pressure
• Decreased blood flow to the brain
• Pain originating in the skin
• Anger, anxiety
• Sexual activity
• Fight-or-flight response
• Exercise
• Collapse of the lungs
• Nicotine
• Caffeine
Decrease in Blood Pressure
• Pain originating in viscera or joints
• Sleeping
• Happiness
• Inflation of the lungs
• Dehydration
• Shock
• Cardiogenic
• Hypovolemic
• Septic
• Anaphylactic

$$MAP = DP + 1/3(SP - DP)$$

In a healthy individual under normal physiological conditions, systolic pressure is determined primarily by stroke volume. Therefore, any factor that alters stroke volume (and cardiac output) will alter systolic pressure and MAP. Diastolic pressure is determined primarily by TPR. Therefore, any factor that alters TPR will alter diastolic pressure and MAP.

The major factors that affect CO, TPR, and, therefore, MAP are summarized in Figure 16.3. These factors may be organized into several categories and will be discussed as such:

- Autonomic nervous system (ANS)
- Vasoactive substances
- Venous return
- Local metabolic activity

Table 16.3 Summary of Major Cardiovascular Principles

$$CO = VR$$

$$CO = HR \times SV$$

$$SV = EDV - ESV$$

$$Q = \frac{\Delta P}{R}$$

$$R \alpha \frac{1}{r^4}$$

$$\text{Pulse Pressure} = P_{systolic} - P_{diastolic}$$

$$MAP = P_{diastolic} + 1/3 \text{ (Pulse Pressure)}$$

$$MAP = CO \times TPR$$

$$VR = \frac{P_V - P_{RA}}{R_V}$$

Notes: CO, cardiac output; VR, venous return; HR, heart rate; SV, stroke volume; EDV, end-diastolic volume; ESV, end-systolic volume; Q, blood flow; ΔP, pressure gradient; R, resistance; r, vessel radius; $P_{systolic}$, systolic pressure; $P_{diastolic}$, diastolic pressure; MAP, mean arterial pressure; TPR, total peripheral resistance, P_v, venous pressure; P_{RA}, right atrial pressure; R_v venous resistance.

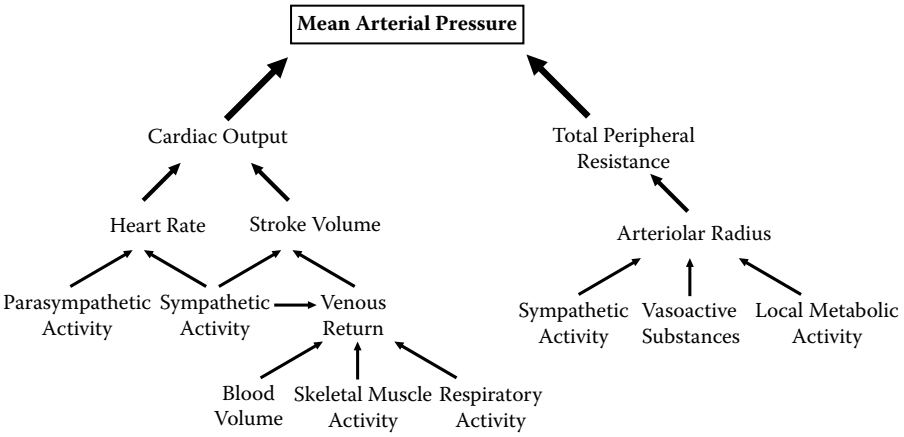


Figure 16.3 Factors that affect mean arterial pressure. Mean arterial pressure is determined by cardiac output and total peripheral resistance. Important factors that influence these two variables are summarized.

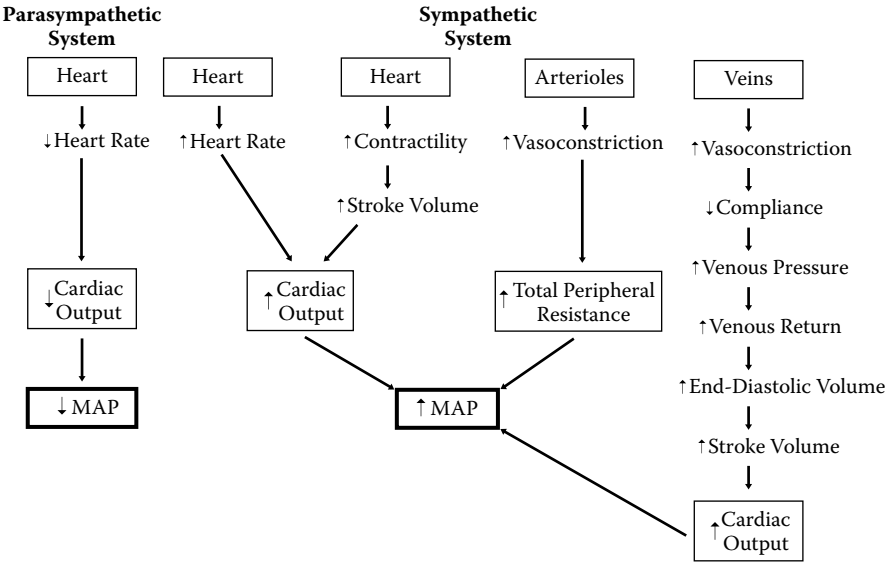


Figure 16.4 Effects of sympathetic and parasympathetic nervous activity on mean arterial pressure. The parasympathetic nervous system innervates the heart and, therefore, influences heart rate and cardiac output. The sympathetic nervous system innervates the heart and the veins and, therefore, influences cardiac output. This system also innervates the arterioles and, therefore, influences total peripheral resistance. The resulting changes in cardiac output and total peripheral resistance regulate mean arterial pressure.

16.6 Autonomic nervous system

The effects of the *autonomic nervous system* on MAP are summarized in Figure 16.4. The *parasympathetic system* innervates the SA node and the AV node of the heart. The major cardiovascular effect of parasympathetic stimulation, by way of the vagus nerves, is to decrease heart rate (HR) which decreases CO and MAP.

The *sympathetic system* innervates most tissues in the heart including the SA node, the AV node, and the ventricular muscle. Sympathetic stimulation causes an increase in HR as well as an increase in ventricular contractility which enhances stroke volume (SV). The increases in HR and SV cause an increase in CO and, therefore, MAP.

The sympathetic system also innervates vascular smooth muscle and regulates the radius of the blood vessels. All types of blood vessels except capillaries are innervated; however, the most densely innervated vessels include the arterioles and the veins. An increase in sympathetic stimulation of vascular smooth muscle causes vasoconstriction, and a decrease in stimulation causes vasodilation. Constriction of arterioles causes an increase in TPR and, therefore, MAP. Constriction of veins causes an increase in venous return

(VR) which increases end-diastolic volume (EDV), SV (Frank-Starling Law of the Heart), CO, and MAP.

Sympathetic nerves are distributed to most vascular beds. They are most abundant in the renal, gastrointestinal, splenic, and cutaneous circulations. Recall that these tissues receive an abundant blood flow, more than is necessary to simply maintain metabolism. Therefore, when blood is needed by other parts of the body, such as working skeletal muscles, sympathetic vasoconstrictor activity reduces flow to the tissues receiving excess blood so that it may be redirected to the muscles. Interestingly, there is no sympathetic innervation to cerebral blood vessels. In fact, these vessels do not have α_1-adrenergic receptors, so they cannot be affected by circulating catecholamines. There is no physiological circumstance where blood should be directed away from the brain.

16.6.1 *Vasomotor center*

Autonomic nervous activity to the cardiovascular system is regulated by the *vasomotor center* (see Figure 16.5). Located in the lower pons and the medulla of the brainstem, the vasomotor center is an integrating center for blood pressure regulation. It receives several sources of input, from the brain as well as the periphery of the body. It processes this information and then adjusts sympathetic and parasympathetic discharge to the heart and the blood vessels accordingly.

Sympathetic nerves going to the arterioles are tonically active. In other words, there is continuous discharge of these nerves causing *vasomotor tone*. As a result, under resting conditions, arterioles are partially constricted. This vasomotor tone is important because it helps to maintain MAP in the range of 90 to 100 mmHg. Without this partial vasoconstriction of the arterioles, MAP would fall precipitously, and blood flow to the vital organs would be compromised. Another physiological advantage of vasomotor tone is that the degree of vasoconstriction can be either increased or decreased. In this

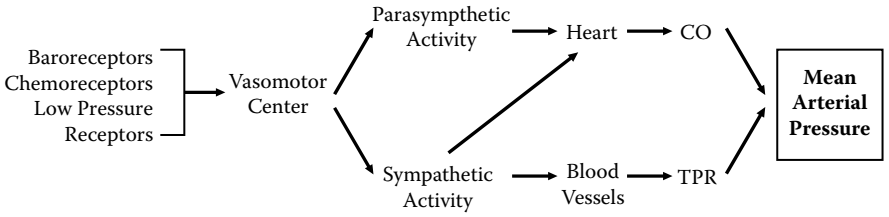


Figure 16.5 The vasomotor center. The baroreceptors, the chemoreceptors, and the low-pressure receptors provide nervous input to the vasomotor center in the brainstem. The vasomotor center integrates this input and determines the degree of discharge by the sympathetic nervous system and the parasympathetic nervous system to the cardiovascular system. Cardiac output and total peripheral resistance are adjusted so as to maintain mean arterial pressure within the normal range.

Table 16.4 Cardiovascular Receptors and Their Stimuli

Receptor	Stimulus
Baroreceptors	Blood pressure
Chemoreceptors	Blood gases ($\downarrow O_2$, $\uparrow CO_2$, $\downarrow$pH)
Low-pressure receptors	Blood volume

way, blood flow to the tissue can be either increased or decreased. Without tone, the vessels could only constrict, and blood flow to the tissue could only decrease.

Other regions of the vasomotor center transmit impulses to the heart via sympathetic nerves or the vagus nerves. An increase in sympathetic activity to the heart typically occurs concurrently with an increase in sympathetic activity to the blood vessels and a decrease in vagal stimulation of the heart. Therefore, the resulting increases in CO and TPR work together to elevate MAP more effectively. Conversely, an increase in vagal stimulation of the heart typically occurs concurrently with a decrease in sympathetic activity to both the heart and the blood vessels. Therefore, decreases in both CO and TPR work together to decrease MAP more effectively.

The vasomotor center receives input from multiple sources (summarized in Table 16.4 and Figure 16.5), including:

- Baroreceptors
- Chemoreceptors
- Low-pressure receptors

16.6.2 Baroreceptors

The *baroreceptors* provide the most important source of input to the vasomotor center. These receptors monitor *blood pressure* in the systemic circulatory system. They are found in two locations: the arch of the aorta and the carotid sinuses. As the aorta exits the left ventricle, it curves over the top of the heart, forming an arch, and then descends through the thoracic and abdominal cavities. The coronary arteries, which supply the cardiac muscle, branch off of the aorta in this most proximal portion of the aorta. The left and right common carotid arteries also branch off of the aortic arch and ascend through the neck toward the head. Each common carotid artery bifurcates, or divides, forming an external carotid artery, which supplies the scalp, and an internal carotid artery, which supplies the brain. The carotid sinus is located at the bifurcation of each common carotid artery. Because blood flow to a tissue is dependent, in large part, upon blood pressure, the baroreceptors are ideally located to monitor blood pressure in regions of the circulatory system responsible for delivering blood to the heart and brain, the two most vital organs in the body.

The baroreceptors respond to stretch or distension of the blood vessel walls. Therefore, they are also referred to as *stretch receptors*. A change in blood pressure will elicit the *baroreceptor reflex* that involves *negative feedback responses* that return the blood pressure to normal (see Figure 16.6). For example, an increase in blood pressure causes distension of the aorta and the carotid arteries which stimulates the baroreceptors. As a result, there is an increase in the number of afferent nerve impulses transmitted from these receptors to the vasomotor center. The vasomotor center processes this information and adjusts the activity of the autonomic nervous system accordingly. Sympathetic stimulation of vascular smooth muscle and cardiac muscle is decreased. Parasympathetic stimulation of cardiac muscle is increased. As a result, venous return, CO, and TPR all decrease so that MAP is decreased back toward its normal value.

On the other hand, a decrease in blood pressure causes less than normal distension or stretch of the aorta and the carotid arteries and a decrease in baroreceptor stimulation. Therefore, fewer afferent nerve impulses are transmitted from these receptors to the vasomotor center. The vasomotor center then alters autonomic nervous system activity such that sympathetic stimulation of vascular smooth muscle and cardiac muscle is increased, and parasympathetic stimulation of cardiac muscle is decreased. As a result, venous return, CO, and TPR all increase so that MAP is increased back toward its normal value. The effects of the autonomic nervous system on the cardiovascular system are summarized in Figure 16.5.

It is important to note that the baroreceptor reflex is elicited whether blood pressure increases or decreases. Furthermore, these receptors are most sensitive in the normal range of blood pressures so that even a small change in MAP will alter baroreceptor, vasomotor center, and autonomic nervous system activity. As such, the baroreceptor reflex plays an important role in the short-term regulation of blood pressure. Without this reflex, changes in blood pressure in response to changes in posture, hydration (blood volume), cardiac output, regional vascular resistance, and emotional state would be far more pronounced. The baroreceptor reflex helps to minimize unintentional changes in MAP and maintain adequate blood flow to the tissues.

16.6.3 Chemoreceptors

The *peripheral chemoreceptors* include the *carotid bodies*, located at the bifurcation of the common carotid arteries, and the *aortic bodies*, located in the arch of the aorta. These receptors are stimulated by a decrease in arterial oxygen (*hypoxemia*), an increase in arterial carbon dioxide (*hypercapnia*), and a decrease in arterial pH (*acidosis*). Therefore, as may be expected, the chemoreceptors are primarily concerned with the regulation of ventilation. A secondary function of these receptors is to influence MAP by providing input to the vasomotor center. A decrease in blood pressure causes a decrease in blood flow to the carotid and aortic bodies. Assuming a constant rate of

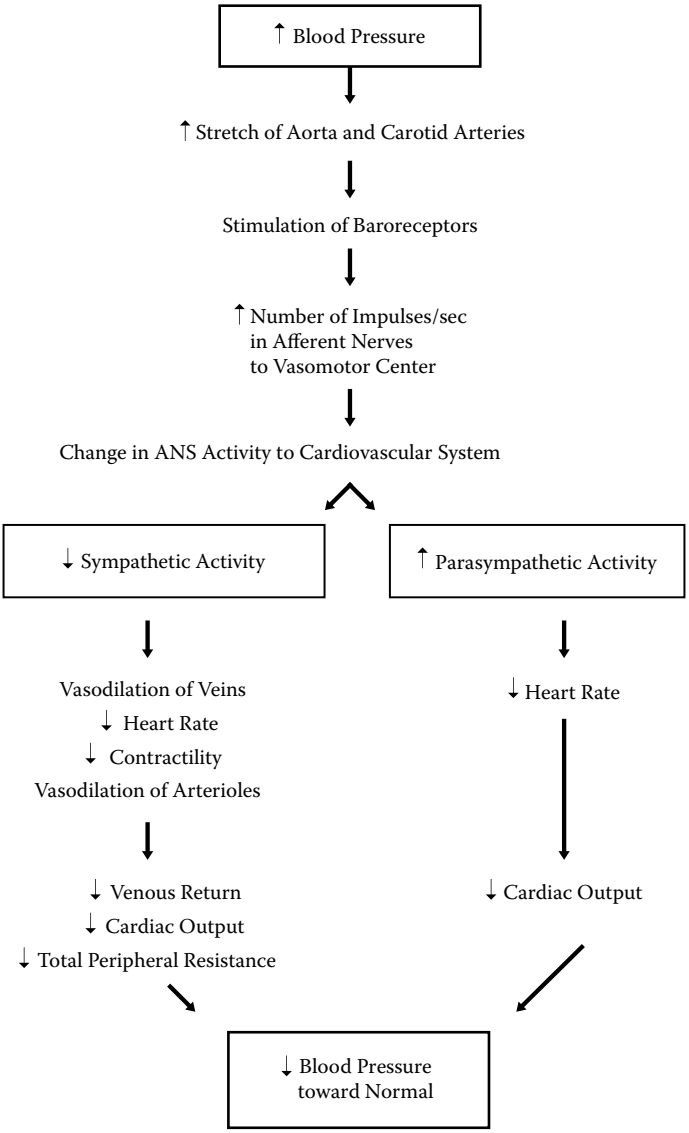


Figure 16.6 The baroreceptor reflex. The baroreceptors are the most important source of input to the vasomotor center. The reflex elicited by these receptors is essential in the maintenance of normal blood pressure.

metabolism in these tissues (constant oxygen consumption as well as carbon dioxide and hydrogen ion production), then a decrease in blood flow results in hypoxemia, hypercapnia, and a decrease in local pH. These conditions stimulate the chemoreceptors and cause an increase in the number of nerve impulses transmitted from these receptors to the vasomotor center.

The vasomotor center processes this input and adjusts the activity of the autonomic nervous system accordingly. Sympathetic discharge to the cardiovascular system is increased. The predominant effect is an increase in TPR. As a result of this negative feedback mechanism, MAP is increased, and blood flow to the chemoreceptors is increased back toward its normal value. Interestingly, the chemoreceptor reflex does not affect the cardiovascular system until MAP decreases below 80 mmHg. Therefore, unlike the baroreceptor reflex, this reflex does not help to minimize the daily variations in MAP. Instead, it supplements the activity of the baroreceptor reflex at lower pressures only.

16.6.4 Low-pressure receptors

The *low-pressure receptors* are located in the walls of the atria and the pulmonary arteries. Similar to the baroreceptors, low-pressure receptors are also stretch receptors. However, stimulation of these receptors is caused by changes in blood volume in these low-pressure areas. An overall increase in blood volume results in an increase in venous return, an increase in the blood volume in the atria and the pulmonary arteries, and stimulation of the low-pressure receptors. Transmission of nerve impulses to the vasomotor center occurs by way of the vagus nerves. The vasomotor center then elicits reflexes that parallel those of the baroreceptors. An increase in blood volume will initially increase MAP. Therefore, sympathetic discharge decreases and parasympathetic discharge increases so that MAP decreases back toward its normal value. The simultaneous activity of both the baroreceptors and the low-pressure receptors makes the total reflex system more effective in the control of MAP.

The low-pressure receptors may also be referred to as *atrial stretch receptors*. Increased atrial filling and stimulation of these receptors elicit additional compensatory responses, including:

- Reflex tachycardia (Bainbridge reflex)
- Reflex vasodilation of renal afferent arterioles
- Decreased secretion of antidiuretic hormone (ADH)
- Increased secretion of atrial natriuretic peptide (ANP)
- Increased urine output (volume reflex)

The net result is a decrease in plasma volume and a decrease in MAP back toward normal.

16.7 Vasoactive substances

Substances released from many cells and tissues in the body, including the endothelium lining the blood vessels, endocrine glands, and myocytes in the heart, may all affect vascular smooth muscle tone (see Table 16.5). These

Table 16.5 Vasoactive Substances

Vasoconstrictors	**Source**
Catecholamines	Adrenal medullae
Angiotensin II	Plasma
Vasopressin	Neurohypophysis
Endothelin	Endothelium
Thromboxane A_2	Platelets
Serotonin	Platelets
Vasodilators	**Source**
Prostacyclin	Endothelium
Nitric oxide	Endothelium
Atrial natriuretic peptide	Atrial myocardial cells
Histamine	Mast cells, basophils
Bradykinin	Plasma, interstitial fluid
Adenosine	Hypoxic cells
Epinephrine	Adrenal medullae

substances may either stimulate this muscle to cause vasoconstriction or inhibit this muscle to cause vasodilation. As expected, vasoconstriction will increase TPR and, therefore, MAP. Vasodilation will decrease TPR and, therefore, MAP.

16.7.1 Vasoconstrictors

There are many substances produced in the human body that cause vasoconstriction under physiological and pathophysiological conditions. *Vasoconstrictors* of particular importance include:

- Catecholamines
- Angiotensin II
- Vasopressin
- Endothelin
- Thromboxane A_2
- Serotonin

The major circulating hormones that influence vascular smooth muscle tone are the *catecholamines*, epinephrine and norepinephrine. These hormones are released from the adrenal medulla in response to sympathetic nervous stimulation. In humans, 80% of catecholamine secretion is epinephrine and 20% is norepinephrine. Stimulation of α_1-adrenergic receptors causes vasoconstriction. The selective α_1-adrenergic receptor antagonist, prazosin, is

effective in the management of hypertension as it causes the relaxation of both arterial and venous smooth muscle.

Angiotensin II is a circulating peptide with powerful vasoconstrictor properties. The formation of angiotensin II is initiated by the enzyme renin, which converts the plasma-borne precursor, angiotensinogen, into angiotensin I. Angiotensin-converting enzyme (ACE) then converts angiotensin I into the active molecule angiotensin II. The location of ACE is ideal for this function, as it is found on the surface of endothelial cells in the lungs which are exposed to the entire cardiac output. The release of renin from specialized cells in the kidneys occurs in response to sympathetic stimulation and when there is a decrease in renal blood flow.

Angiotensin II causes vasoconstriction by the direct stimulation of AT_1 receptors on the vascular smooth muscle. It also enhances the release of the neurotransmitter norepinephrine from the sympathetic nerve fibers present in the blood vessels. The vasopressor effects of angiotensin II may be inhibited pharmacologically in order to decrease TPR and treat hypertension. An important class of orally active drugs is the ACE inhibitors, including captopril and enalopril, which prevent the formation of angiotensin II. More recently, angiotensin receptor antagonists have been developed which act at the vascular smooth muscle. These drugs, including losartin and valsartan, are also orally active.

Vasopressin (antidiuretic hormone) is a peptide synthesized in the hypothalamus and secreted from the neurohypophysis of the pituitary gland. This substance plays an important role in the long-term regulation of blood pressure through its action on the kidney to increase the reabsorption of water. The major stimulus for the release of vasopressin is an increase in plasma osmolarity. The resulting reabsorption of water dilutes the plasma back toward its normal value of 290 mOsM. This activity is discussed in more detail in Chapter 10 and Chapter 21.

Vasopressin also plays an important role in the short-term regulation of blood pressure through its action on vascular smooth muscle. This hormone is one of the most potent known endogenous vasoconstrictors. Two types of vasopressin receptors have been identified: V_1 and V_2 receptors. V_{1A} receptors mediate vasoconstriction, and V_2 receptors mediate the antidiuretic effects of this hormone. Specific V_{1A} receptor antagonists of the vasoconstrictor activity of vasopressin are under development.

The vascular endothelium produces a number of substances that are released basally into the blood vessel wall to alter vascular smooth muscle tone. One such substance is *endothelin*. Endothelin exerts its effects throughout the body causing vasoconstriction as well as positive inotropic and chronotropic effects on the heart. The resulting increases in TPR and CO both contribute to an increase in MAP.

Synthesis of endothelin appears to be enhanced by many stimuli, including angiotensin II, vasopressin, and the mechanical stress of blood flow on the endothelium. Synthesis is inhibited by vasodilator substances, such as

prostacyclin, nitric oxide, and atrial natriuretic peptide. There is evidence that endothelin is involved with the pathophysiology of many cardiovascular diseases, including hypertension, heart failure, and myocardial infarction. Endothelin receptor antagonists are currently available for research use only.

Another vasoactive substance produced by the endothelium is *thromboxane A_2* (*TxA_2*). Typically, small amounts of TxA_2 are released continuously; however, increased synthesis appears to be associated with some cardiac diseases. Synthesized from arachidonic acid, a plasma membrane phospholipid, TxA_2, is a potent vasoconstrictor. Furthermore, this substance stimulates platelet aggregation, suggesting that it plays a role in thrombotic events such as myocardial infarction (heart attack). Nonsteroidal antiinflammatory drugs, such as aspirin and ibuprofen, block the formation of TxA_2 and reduce the formation of blood clots.

Serotonin is released from platelets. Along with TxA_2, serotonin enhances the vasoconstriction that occurs when a blood vessel is damaged. The resulting decrease in blood flow contributes to hemostasis, which is discussed in Chapter 17.

PHARMACY APPLICATION: ANTIHYPERTENSIVE DRUGS

Hypertension is the most common cardiovascular disease. In fact, nearly 25% of the adults in the United States are considered hypertensive. Hypertension is defined as a consistent elevation in blood pressure, such that systolic/diastolic pressures are >140/90 mmHg. Over time, chronic hypertension can cause pathological changes in the vasculature and in the heart. As a result, hypertensive patients are at increased risk for atherosclerosis, aneurysm, stroke, myocardial infarction, heart failure, and kidney failure.

There are several categories of antihypertensive agents:

Diuretics: Initially, the primary mechanism by which diuretics reduce blood pressure is to decrease plasma volume. Acting at the kidney, diuretics increase sodium loss and, due to the osmotic effects of sodium, increase water loss. The decrease in plasma volume results in a decrease in VR, CO, and MAP. However, the long-term hypotensive effect of the diuretics appears to be due to a decrease in TPR.

Sympatholytics: Sympathetic stimulation of the cardiovascular system may be altered by several mechanisms:

Centrally acting agents exert their effects at the vasomotor center in the brainstem and inhibit sympathetic discharge. Reduced sympathetic stimulation of the heart and, especially the vascular smooth muscle, results in some decrease in CO, especially in older patients, and a marked decrease in TPR.

Beta-adrenergic receptor antagonists reduce myocardial contractility and CO. These agents also reduce the secretion of renin. Therefore, their effects on reducing blood pressure can be explained, in part, by a reduction in the formation of angiotensin II. However, this mechanism does not fully account for the antihypertensive effects of beta blockers. These drugs also reduce peripheral vascular resistance; however, the mechanism of this effect is not known.

Alpha-adrenergic receptor antagonists reduce peripheral vascular resistance and, therefore, reduce blood pressure.

Vasodilators: Hydralazine causes direct relaxation of arteriolar smooth muscle. An important consequence of this vasodilation, however, is reflex tachycardia (↑CO). It may also cause sodium retention (↑ plasma volume). The resulting increase in CO tends to offset the effects of the vasodilator. Therefore, these drugs are most effective when administered along with sympathetic agents, such as β-adrenergic receptor antagonists, which prevent unwanted compensatory responses by the heart.

Ca^{++}-channel blockers: The dihydropyridine agents, such as amlodipine (Norvasc®) and felodipine (Plendil®), are the preferred types of calcium channel blockers used in the treatment of hypertension because they lack inotropic and chronotropic effects on the heart. Therefore, the mechanism of action of these drugs involves a marked decrease in peripheral vascular resistance. Furthermore, they cause less reflex tachycardia than nifedipine.

ACE inhibitors: ACE inhibitors not only cause vasodilation (↓TPR), they inhibit the aldosterone response to net sodium loss. Typically, aldosterone, which enhances the reabsorption of sodium in the kidney, would oppose diuretic-induced sodium loss. Therefore, the coadministration of ACE inhibitors would enhance the efficacy of diuretic drugs.

Angiotensin II receptor antagonists: These agents promote vasodilation (↓TPR), increase sodium and water excretion and, therefore, decrease plasma volume. ACE inhibitors and angiotensin II receptor antagonists are the drugs of first choice in patients with heart failure.

Drug Classification	Generic Agents	CO	TPR	PV
Diuretics		↔	↓	↓/↔
Thiazides	Hydrochlorothiazide			
Loop Diuretics	Furosemide			
K^+-Sparing Diuretics	Amiloride Spironolactone			
Sympatholytics				
Centrally Acting	Methyldopa Clonidine	↓/↔	↓	↔
β-Antagonist	Propranolol	↓	↓	↔
α-Antagonist	Prazosin	↔	↓	↔
Vasodilators				
Arterial	Hydralazine	↑	↓	↑
Arterial and Venous	Nitroprusside	↔	↓	↔
Ca^{++}-Channel Blockers	Verapamil	↓	↓	↔
	Nifedipine	↔	↓	↔
	Diltiazem	↓	↓	↔
ACE Inhibitors	Captopril Enalapril	↑	↓	↓
AgII-Antagonists	Losartin Valsartan	↑	↓	↓

16.7.2 Vasodilators

Many substances produced in the human body cause vasodilation under physiological and pathophysiological conditions. *Vasodilators* of particular importance include:

- Prostacyclin
- Nitric oxide
- Atrial natriuretic peptide
- Histamine
- Bradykinin
- Adenosine
- Epinephrine

Another metabolite of arachidonic acid is *prostacyclin* (*PGI_2*). As with TxA_2, PGI_2 is produced continuously. Synthesized by vascular smooth muscle and endothelial cells, with the endothelium as the predominant source, PGI_2 mediates effects that are opposite to those of TxA_2. Prostacyclin causes vasodilation and inhibits platelet aggregation. As a result, PGI_2 contributes importantly to the antithrombogenic nature of the vascular wall.

First described in the 1980s as "endothelium-derived relaxing factor," *nitric oxide* (*NO*) is a vasodilator believed to play a role in the regulation of blood pressure under both physiologic and pathophysiological conditions. For example, the pharmacologic inhibition of NO synthesis under normal conditions and during septic shock results in a significant elevation of blood pressure.

PHARMACY APPLICATION: NITROGLYCERIN AND ANGINA

Angina pectoris (chest pain) is the most common symptom of chronic ischemic heart disease. Angina is caused by an imbalance between the oxygen supply and the oxygen demand of the cardiac muscle. Myocardial oxygen demand increases during exertion, exercise, and emotional stress. If coronary blood flow does not increase proportionately to meet this demand, then the affected tissue becomes ischemic, and pain develops. This ischemia and pain may be treated pharmacologically with nitroglycerin. This drug causes vasodilation and an increase in blood flow. However, this effect occurs not only in the coronary arteries, but also in blood vessels throughout the body. Therefore, in addition to improving coronary blood flow, administration of nitroglycerin may decrease systemic blood pressure. The mechanism of action of nitroglycerin involves the release of NO in the vascular smooth muscle. Most frequently, this drug is administered in the sublingual form and its effects are apparent within 1 to 3 minutes.

Atrial natriuretic peptide (ANP) is produced by specialized myocytes in the atria of the heart. Secretion is stimulated by increased filling and stretch of the atria in response to plasma volume expansion. The effects of ANP include vasodilation, diuresis (increased urine production), and increased sodium excretion. Taken together, these effects decrease blood volume and blood pressure back toward normal.

Histamine and *bradykinin* are important mediators of inflammation. The actions of these substances are discussed in detail in Chapter 18.

Adenosine is an important mediator of active hyperemia. This phenomenon is discussed in Section 16.10.

Epinephrine may cause vasodilation at higher concentrations due to the stimulation of β_2-adrenergic receptors on vascular smooth muscle. These receptors are found primarily on blood vessels in skeletal muscles and in the hepatic and coronary circulations.

16.8 Venous return

The vessels of the circulatory system have varying degrees of *distensibility*. This feature allows them to accommodate changes in blood volume. For

example, the abrupt addition of the stroke volume to the aorta and large arteries during ventricular systole causes these vessels to expand and actually "store" a portion of the blood pumped by the heart. The subsequent elastic recoil of the arteries forces the stored blood forward during ventricular diastole. Therefore, the slight distensibility of arteries results in the maintenance of blood flow to the tissues throughout the cardiac cycle.

The most distensible vessels in the circulatory system are the veins. As with the arteries, this feature of the veins also has important physiological implications, as it allows them to serve as *blood reservoirs*. The veins are so distensible that they are capable of holding large volumes of blood at very low pressures. In fact, under resting conditions, 64% of the blood volume is contained within these vessels.

Compliance (*C*) in the circulatory system describes the relationship between vascular blood volume (V) and intravascular pressure (P):

$$C = \frac{V}{P}$$

In other words, it is a measure of the inherent distensibility of the blood vessels. The more compliant the vessel, the greater is the volume of blood that it is capable of accommodating.

As mentioned, all blood vessels are compliant. However, the marked difference in distensibility between the arteries and the veins is illustrated by the following:

Compliance of the systemic arteries at rest:

$$C = \frac{13\% \text{ of the blood volume}}{100 \text{ mmHg}}$$

Compliance of the systemic veins at rest:

$$C = \frac{64\% \text{ of the blood volume}}{8 \text{ mmHg}}$$

Due to the significant amounts of elastic connective tissue and smooth muscle in their walls, arteries tend to recoil rather powerfully which keeps the pressure within them high. Of course, this elevated pressure is necessary to drive the blood through the circulatory system. In contrast, veins contain less elastic connective tissue and smooth muscle, so the tendency to recoil is significantly less, and the pressure remains low.

The venous system, in general, serves as a reservoir for the circulatory system. However, there are some tissues that are particularly important in this

respect. These *specific blood reservoirs* include the spleen, the liver, the large abdominal veins, and the venous plexus beneath the skin. Common features of the vascular beds within these tissues are that they are very extensive and very compliant. In this way, under normal, resting conditions, these vascular beds can accommodate large volumes of blood. In fact, taken together, these tissues may hold up to 1 L of blood, or 20% of the blood volume. Under pathological conditions, such as hemorrhage or dehydration, blood may be mobilized from these tissues, allowing the circulatory system to function relatively normally until the blood volume is restored to normal.

In addition to serving as blood reservoirs, veins help to *regulate cardiac output* (CO) by way of changes in *venous return* (VR). Venous return is defined as the volume of blood that flows from the systemic veins into the right atrium per minute. As discussed in Chapter 14, a healthy heart pumps all of the blood returned to it. Therefore, CO is equal to VR:

$$CO = VR$$

On the other hand, the heart can only pump whatever blood it receives. Therefore, in order to increase CO, then VR must also increase.

As with blood flow through a vessel, blood flow through the venous system is determined by Ohm's Law ($Q = \Delta P/R$). In other words, it depends on the pressure gradient in the venous system and venous resistance. Ohm's Law may be rewritten to calculate VR:

$$VR = \frac{P_V - P_{RA}}{R_V}$$

The pressure gradient, or the inflow pressure minus the outflow pressure, is determined by the pressure at the beginning of the venous system (P_V) and right atrial pressure (P_{RA}) at the end of the system. The smaller compliant veins offer very little resistance to blood flow. The slightly stiffer large veins offer a small degree of resistance (R_V).

There are several factors that influence VR:

- Blood volume
- Sympathetic stimulation of the veins
- Skeletal muscle activity
- Respiratory activity

16.8.1 Blood volume

Blood volume has a direct effect on blood pressure. It also has an important effect on VR. A decrease in blood volume resulting from a hemorrhage or

dehydration causes a decrease in venous pressure and a decrease in VR. An increase in blood volume following oral or venous rehydration or a transfusion causes an increase in venous pressure and an increase in VR.

16.8.2 Sympathetic stimulation of the veins

The smaller, more compliant veins that serve generally as blood reservoirs as well as the specific blood reservoirs are densely innervated by the *sympathetic system*. Stimulation of the vascular smooth muscle in the walls of these vessels causes vasoconstriction and a decrease in venous compliance. The vasoconstriction increases venous pressure in the veins. The blood is squeezed out of the veins and, due to the presence of the one-way valves, moves toward the heart so that VR increases. A decrease in sympathetic stimulation allows the veins to relax and distend. The vessels become more compliant and capable of holding large volumes of blood at low pressures. In this case, VR decreases.

The effect of sympathetic stimulation on venous resistance is minimal. As previously stated, the larger, less flexible veins provide resistance to blood flow. However, these blood vessels are sparsely innervated. Therefore, there is little change in vessel radius, and the physiological effect on blood flow is relatively insignificant.

16.8.3 Skeletal muscle activity

In the extremities (arms and legs), many veins lie between the skeletal muscles. Contraction of these muscles causes compression of the veins and an increase in venous pressure. This external compression squeezes the blood out and forces it back toward the heart causing an increase in VR. This action is referred to as the *skeletal muscle pump*.

The effect of the skeletal muscle pump is essential during exercise. Although there is a mass sympathetic discharge and venous vasoconstriction that enhances VR, this mechanism alone is insufficient to increase VR and, therefore, CO to meet the metabolic demands of strenuous exercise. The skeletal muscle pump mobilizes the blood stored in these tissues and keeps it flowing toward the heart. As the number of muscles involved in the exercise increases, so does the magnitude of the increase in VR and CO.

16.8.4 Respiratory activity

Pressures in the venous system are altered during respiration. *Inspiration* causes a decrease in thoracic pressure and, therefore, a decrease in pressure within the venae cavae and the right atrium. Furthermore, the downward movement of the diaphragm causes an increase in abdominal pressure. Many large veins and specific blood reservoirs are located in the abdomen.

Compression of these tissues by the diaphragm causes an increase in venous pressure in this region. Therefore, the overall effect of inspiration is to increase the pressure gradient between extrathoracic and intrathoracic veins. This results in an increase in VR.

16.9 Effects of gravity on the circulatory system

Gravitational forces may have a profound influence on blood flow through the circulatory system. As a result, both VR and CO may be affected. Imagine that the circulatory system is a column of blood that extends from the heart to the feet. As in any column of fluid, the pressure at the surface is equal to zero. Due to the weight of the fluid, the pressure increases incrementally below the surface. This pressure is referred to as the *hydrostatic pressure*.

In an upright adult, the hydrostatic pressure of the blood in the feet may be as high as 90 mmHg. When this pressure is added to the pressure in the veins generated by the pumping activity of the heart, the total pressure in the veins in the feet may be as high as 100 mmHg. The valves in the veins effectively prevent the backward flow of blood toward the feet. However, the valves have no effect on the buildup of pressure in the veins in the lower extremities. The capillaries in the feet are also subjected to the effects of gravity. Pressure in these vessels may be as high as 135 mmHg.

The increased hydrostatic pressures in the veins and capillaries have two very detrimental effects on the circulatory system: the pooling of blood and edema formation. *Blood tends to pool* in the highly distensible veins. Furthermore, excessive filtration of fluid out of the capillaries and into the tissues occurs, causing *edema* or swelling of the ankles and feet. As a result, VR and, therefore, CO are decreased. This leads to a decrease in MAP. The fall in MAP can cause a decrease in cerebral blood flow and, possibly, syncope (fainting).

Compensatory mechanisms in the circulatory system are needed to counteract the effects of gravity. Two important mechanisms include the *baroreceptor reflex* and *skeletal muscle activity*.

Baroreceptors are sensitive to changes in MAP. As VR, CO, and MAP decrease, there is diminished excitation of the baroreceptors. Consequently, the frequency of nerve impulses transmitted from these receptors to the vasomotor center in the brainstem is reduced. This elicits a reflex that will increase HR, increase contractility of the heart, and cause vasoconstriction of both arterioles and veins. The increases in CO and TPR effectively increase MAP and, therefore, cerebral blood flow. Constriction of the veins assists in forcing blood back toward the heart and enhances venous return.

The *skeletal muscle activity* associated with simply walking, decreases venous pressure in the lower extremities significantly. Contraction of the skeletal muscles in the legs compresses the veins and blood is forced toward the heart.

16.10 Regulation of blood flow through tissues

The blood flow to most tissues in the body is determined by the metabolic needs of those tissues. Metabolically active tissues require enhanced delivery of oxygen and nutrients as well as enhanced removal of carbon dioxide and waste products. In general, as the metabolic activity of a tissue increases, its blood flow increases. An important feature of the circulatory system is that each tissue has the intrinsic ability to control its own local blood flow in proportion to its metabolic needs.

16.10.1 Active hyperemia

The increase in blood flow caused by enhanced tissue activity is referred to as *active hyperemia*. Assuming a constant blood pressure, then according to Ohm's Law ($Q = \Delta P/R$), the increase in blood flow is the result of a decrease in local vascular resistance. Tissue metabolism causes several local chemical changes that can mediate this metabolic vasodilation:

- Decreased oxygen
- Increased carbon dioxide
- Increased hydrogen ions
- Increased potassium ions
- Increased adenosine

As metabolism increases, *oxygen consumption* and *carbon dioxide production* are enhanced. The concentration of *hydrogen ions* is also enhanced as more carbonic acid (formed from carbon dioxide) and lactic acid are produced by the working tissue. Furthermore, the concentration of *potassium ions* in the interstitial fluid is increased. The rate of potassium release from the cells due to repeated action potentials exceeds the rate of potassium return to the cells by way of the Na^+-K^+ pump. Finally, the release of *adenosine* is also believed to play an important role in the regulation of resistance vessels, particularly in the heart and the skeletal muscle.

Each of these chemical changes promotes *vasodilation of arterioles*. In addition, the increase in *tissue temperature* associated with increased metabolism further contributes to metabolic vasodilation. The resulting increase in local blood flow restores these substances to their resting values. More oxygen is delivered, and the excess carbon dioxide, hydrogen ions, potassium ions, and adenosine are removed.

It is important to note that local regulatory mechanisms override the effects of extrinsic sympathetic stimulation. For example, during exercise, the mass sympathetic discharge would tend to cause widespread vasoconstriction, even in skeletal muscles. However, local vasodilation in working muscles due to active hyperemia supersedes the effects of the sympathetic nerves. Consequently, these tissues are supplied with the blood that they need.

16.10.2 Autoregulation

A different situation arises when the metabolic rate of a tissue remains constant, but the blood pressure changes. According to Ohm's Law ($Q = \Delta P/R$), an increase in blood pressure would tend to increase blood flow to a tissue. However, if there is no change in the metabolic activity of the tissue, then an increase in blood flow is unnecessary. In fact, blood flow to the tissue returns most of the way back to normal rather rapidly. The maintenance of a relatively constant blood flow to a tissue, in spite of changes in blood pressure, is referred to as *autoregulation*. Once again, resistance changes in the arterioles are involved.

Arteriolar resistance changes that take place in order to maintain a constant blood flow may be explained by the *myogenic mechanism*. According to this mechanism, vascular smooth muscle contracts in response to stretch. For example, consider the situation where blood pressure is increased. The increase in pressure causes an initial increase in blood flow to the tissue. However, the increased blood flow is associated with increased stretch of the vessel wall. This leads to the opening of stretch-activated calcium channels in the vascular smooth muscle. The ensuing increase in intracellular calcium results in vasoconstriction and a decrease in the blood flow to the tissue back toward normal.

Another mechanism that may explain autoregulation is the *metabolic theory* described previously as causing active hyperemia. Consider the situation where there is a decrease in blood pressure. The decrease in pressure would cause an initial decrease in blood flow. Assuming a constant rate of metabolism, the decrease in blood flow would lead to a decrease in tissue oxygen and an increase in tissue carbon dioxide, hydrogen ions, potassium ions, and adenosine. Once again, each of these chemical changes promotes vasodilation of the arterioles and an increase in blood flow to the tissue back toward normal.

16.11 Effects of acute exercise on the circulatory system

The primary goal of the circulatory system during exercise is to *increase blood flow to the working muscles*. This is accomplished by *increasing MAP* and *decreasing local vascular resistance*:

$$\uparrow\uparrow Q = \frac{\uparrow MAP}{\downarrow R}$$

At the onset of exercise, signals from the cerebral cortex are transmitted to the vasomotor center in the medulla of the brainstem. This *central command* not only inhibits parasympathetic activity, it also initiates the *mass sympathetic*

discharge associated with exercise. Sympathetic activity (including the release of catecholamines from the adrenal medulla) increases proportionally with the intensity of exercise.

Sympathetic stimulation of the heart results in the following:

Increased HR and Increased myocardial contractility
→ Increased SV and, therefore, Increased CO

Sympathetic stimulation of the veins and other blood reservoirs results in

Increased P_v → Increased VR → Increased EDV
→ Increased SV, and therefore, Increased CO

In other words, the increase in CO occurs by both extrinsic (sympathetic stimulation) and intrinsic (increased VR and the Frank-Starling Law of the Heart) mechanisms. Venous return is also markedly increased by the compression of blood vessels in the working muscles. The increase in CO causes an *increase in MAP*. The increase in MAP contributes to the increase in muscle blood flow.

Sympathetic stimulation of the arterioles results in increased TPR. Most arterioles of the peripheral circulation are strongly constricted by direct sympathetic stimulation. This widespread vasoconstriction serves two purposes. First, it contributes to the increase in MAP. Second, it is an important factor in the *redirection of blood flow* away from the inactive tissues and toward the working muscles. For example, at rest, the kidneys and the abdominal organs receive 20% and 27% of the cardiac output, respectively. During exercise, profound vasoconstriction in these vascular beds may reduce the blood flow to each of these circulations to as little as 3% of the cardiac output. On the other hand, the skeletal muscles that normally receive 20% of the cardiac output at rest may receive as much as 70% to 80% of the cardiac output during exercise.

Resistance in the arterioles of the working muscles is regulated locally. As discussed previously, active hyperemia results in the production of several factors that cause *metabolic vasodilation*. Exercising muscles generate CO_2, H^+ ions, K^+ ions, heat, and adenosine. The vasodilator effect of these locally produced substances overrides the vasoconstrictor effect of the sympathetic system in the muscle. As a result, *local vascular resistance is decreased*. The combination of increased driving pressure and decreased local vascular resistance causes an increase in blood flow to the working muscles. Changes in the cardiovascular system during exercise are summarized in Table 16.6.

Table 16.6 Changes in the Cardiovascular System during Moderate Exercise and Their Mechanisms

- Increased cardiac output: increased heart rate and increased stroke volume.
- Increased heart rate: increased sympathoadrenal input and decreased parasympathetic input to the sinoatrial (SA) node.
- Increased stroke volume: increased sympathoadrenal input to the ventricles increases ejection fraction from 60% to as much as 90%; increased venous return and increased end-diastolic volume (Frank-Starling Law of the Heart).
- Increased venous return: increased sympathetic input to the veins; skeletal muscle pumping activity.
- Increased total peripheral resistance: increased sympathetic input to vascular smooth muscle in metabolically inactive tissues (may be partially or entirely offset by vasodilation in working skeletal muscles).
- Increased mean arterial pressure: increased cardiac output and increased total peripheral resistance (due primarily to increased stroke volume and increased cardiac output).
- Decreased blood flow to kidneys and abdominal organs: increased sympathetic input to vascular smooth muscle (facilitates redirection of blood flow).
- Increased blood flow to skin: decreased sympathetic input to vascular smooth muscle (facilitates thermoregulation).
- Increased blood flow to skeletal muscles and heart: active hyperemia, increased mean arterial pressure.

16.12 Capillary exchange

The *capillaries* are the *site of exchange* between the blood and the interstitial fluid surrounding the cells of the tissues. Tissues with a higher metabolic rate have a more extensive capillary network. In other words, there are a greater number of capillaries per unit area. Because of the extensive branching of these vessels, the cells of the body are typically within 20 μm of the nearest capillary. Consequently, the distance that substances must travel between the blood and the cells is minimized. Capillaries are permeable to water and small water-soluble substances, such as glucose, amino acids, lactic acid, and urea. Capillaries are impermeable to proteins.

The *velocity of blood flow* through capillaries is slow compared to the rest of the circulatory system. This is because of the very large *total cross-sectional surface area* of the capillaries. Although each individual capillary has a diameter of only about 8 to 10 μ, when the cross-sectional areas of all of the billions of capillaries are combined, the total is well over 1000 times larger than that of the aorta. As the total cross-sectional area increases, the velocity of blood flow decreases. The physiological significance of this low velocity of blood flow is that it allows for adequate time for the exchange of materials between the blood and the cells of the tissues.

Blood flow through individual capillaries is *intermittent*, or sporadic. At the beginning of each capillary, where it branches off of the arteriole, there is a ring of smooth muscle referred to as the *precapillary sphincter.* This sphincter alternately contracts and relaxes. When contracted, blood flow through the capillary is interrupted. When relaxed, blood flow through the capillary resumes. This process of contraction and relaxation of the precapillary sphincter is referred to as *vasomotion*. It is regulated by the rate of metabolism in the tissue, or the oxygen demand of the tissue. As metabolism in the tissue and its need for oxygen increase, the rate of vasomotion increases, and the periods of relaxation are longer. In this way, blood flow through the capillary is markedly increased, and the metabolic needs of the tissue are met. At rest, approximately 10% to 20% of the body's capillaries are perfused at any given moment. During exercise, these changes in vasomotion allow blood to flow through all of the capillaries in the working muscles, contributing significantly to the enhanced perfusion.

There are three primary mechanisms by which substances are exchanged across the capillary wall:

1. Diffusion
2. Transcytosis
3. Bulk flow

The most important mechanism is *diffusion*. If a substance is permeable, it moves in or out of the capillary down its concentration gradient. *Lipid-soluble substances* can diffuse through the endothelial cells at any point along the capillary. These molecules, especially oxygen and carbon dioxide, can pass directly through the lipid bilayer. However, *water-soluble substances* move across the membrane only through water-filled pores in the endothelial cells. Small, water-soluble substances, such as glucose, amino acids, and ions, pass readily through the pores. Water may also cross capillary walls through water-selective channels in the endothelial cells referred to as *aquaporins.*

Large, nonlipid-soluble molecules may cross the capillary wall by way of *transcytosis*. This mechanism involves the transport of *vesicles* from one side of the capillary wall to the other. Many hormones, including the catecholamines and those derived from proteins, exit the capillaries and enter their target tissues by way of transcytosis.

The third mechanism of capillary exchange is *bulk flow.* In this case, water and dissolved solutes move across the capillaries according to *hydrostatic pressures* and *osmotic pressures*. When the balance of these two forces causes fluid to move out of the capillary, it is referred to as *filtration***.** When these forces cause fluid to move into the capillary, it is referred to as *reabsorption*.

An interesting phenomenon in the circulatory system is that, even though capillaries have numerous pores in their walls, all of the fluid does not leak out of them into the interstitial space. If a balloon filled with water had multiple pinpricks in it, all of the water would clearly leak out of it. What prevents

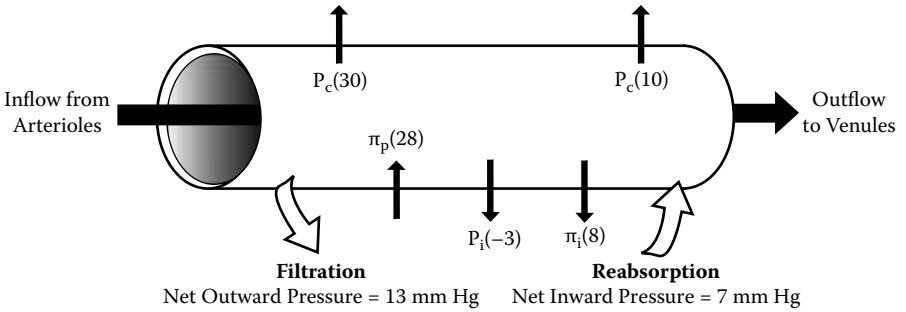


Figure 16.7 The Starling Principle. Illustrated is a summary of the forces determining the bulk flow of fluid across the wall of a capillary. Hydrostatic forces include capillary pressure (P_c) and interstitial fluid pressure (P_i). Capillary pressure pushes fluid out of the capillary. Interstitial fluid pressure is negative and acts as a suction pulling fluid out of the capillary. Osmotic forces include plasma colloid osmotic pressure (π_p) and interstitial fluid colloid osmotic pressure (π_i). These forces are caused by proteins that pull fluid toward them. The sum of these four forces results in the net filtration of fluid at the arteriolar end of the capillary (where P_c is high) and net reabsorption of fluid at the venular end of the capillary (where P_c is low).

this from happening in the capillaries? The *Starling Principle* describes the process by which plasma is held within the vascular compartment.

There are four forces that determine the movement of fluid into or out of the capillary (see Figure 16.7):

1. Capillary hydrostatic pressure (P_c)
2. Interstitial fluid hydrostatic pressure (P_i)
3. Plasma colloid osmotic pressure (π_p)
4. Interstitial fluid colloid osmotic pressure (π_i)

Capillary hydrostatic pressure forces fluid out of the capillary. This pressure is higher at the inflow end of the capillary (30 mmHg) than it is at the outflow end (10 mmHg). The *interstitial fluid hydrostatic pressure* would tend to force fluid into the capillary if it were positive. However, this pressure is usually negative and, instead, acts to suction and pulls fluid out of the capillary. Although it varies depending upon the specific tissue, the average interstitial fluid hydrostatic pressure is about –3 mmHg.

Plasma colloid osmotic pressure is generated by the proteins in the plasma that cannot cross the capillary wall. These proteins exert an osmotic force, pulling fluid into the capillary. In fact, the plasma colloid osmotic pressure, which is about 28 mmHg, is the only force holding the fluid within the capillaries. The *interstitial fluid colloid osmotic pressure* is generated by the small amounts of plasma proteins that leak into the interstitial space. Because these proteins are normally returned to the blood by way of the lymphatic system,

the protein concentration in the interstitial fluid is very low. The average interstitial fluid colloid osmotic pressure is 8 mmHg.

Note that, except for the capillary hydrostatic pressure, the magnitude of these forces remains constant throughout the length of the capillary. The capillary hydrostatic pressure decreases steadily as the blood flows from the arteriolar end to the venular end of the capillary. The steady decline in this pressure results in the filtration of fluid at one end and the reabsorption of fluid at the other end of the capillary.

At the *arteriolar end of the capillary*, the pressures forcing fluid out of the capillary include the following:

$$\text{Outward forces} = P_c + P_i + \pi_i$$

$$= 30 \text{ mmHg} + 3 \text{ mmHg} + 8 \text{ mmHg}$$

$$= 41 \text{ mmHg}$$

Although the interstitial fluid hydrostatic pressure is "negative," it causes fluid to be pulled out of the capillary, so this pressure is "added" to the other outward forces. The only force pulling fluid into the capillary is the plasma colloid osmotic pressure:

$$\text{Inward force} = \pi_p$$

$$= 28 \text{ mmHg}$$

The sum of the outward forces (41 mmHg) exceeds that of the inward force (28 mmHg), resulting in a *net filtration pressure* of 13 mmHg. In other words, there is a net movement of fluid out of the capillary at the arteriolar end.

At the *venular end of the capillary*, the sum of the pressures forcing fluid out of the capillary is decreased due to the fall in capillary hydrostatic pressure:

$$\text{Outward forces} = P_c + P_i + \pi_i$$

$$= 10 \text{ mmHg} + 3 \text{ mmHg} + 8 \text{ mmHg}$$

$$= 21 \text{ mmHg}$$

The plasma colloid osmotic pressure remains constant:

$$\text{Inward force} = \pi_p$$

$$= 28 \text{ mmHg}$$

Therefore, at the venular end of the capillary, the inward force (28 mmHg) exceeds the sum of the outward forces (21 mmHg), resulting in a *net reabsorption pressure* of 7 mmHg. In other words, there is a net movement of fluid into the capillary at the venular end.

Bulk flow plays only a minor role in the exchange of specific solutes between the blood and the cells of the tissues. A far more important function of bulk flow is to *regulate the distribution of the extracellular fluid* between the vascular compartment (plasma) and the interstitial space. The maintenance of an appropriate circulating volume of blood is an important factor in the maintenance of blood pressure. For example, dehydration and hemorrhage will cause a decrease in blood pressure. This leads to a decrease in the capillary hydrostatic pressure. As a result, net filtration decreases and net reabsorption increases, causing the movement, or bulk flow, of extracellular fluid from the interstitial space into the vascular compartment. This fluid shift expands the plasma volume and compensates for the fall in blood pressure.

Over the course of the day, approximately 20 L of fluid are filtered from the capillaries and about 17 L of fluid are reabsorbed back into the capillaries. The remaining 3 L of fluid is returned to the vascular compartment by way of the *lymphatic system*.

The *lymphatic capillaries* are close-ended vessels in close proximity to the blood capillaries. As with blood capillaries, the lymphatic capillaries are composed of a single layer of endothelial cells. However, large gaps in between these cells allow not only fluid, but also proteins and particulate matter to enter the lymphatic capillaries quite readily. Once the fluid has entered these capillaries, it is referred to as *lymph*. Not surprisingly, the composition of this fluid is similar to that of the interstitial fluid.

Lymphatic capillaries join together to form larger *lymphatic vessels*. These vessels have *valves* within them to ensure the one-way flow of lymph. The lymph is moved along by two mechanisms. Automatic, rhythmic waves of contraction of the smooth muscle in the walls of these vessels are the primary mechanism by which lymph is propelled through the system. Second, the contraction of skeletal muscles causes compression of lymphatic vessels. As in the veins, this pumping action of the surrounding skeletal muscles contributes to the movement of the lymph. Ultimately, the lymph is returned to the blood when it empties into the subclavian and jugular veins near the heart.

Four general conditions can lead to *edema* formation, or excess fluid accumulation in the tissue:

1. Increased capillary hydrostatic pressure
2. Blockage of lymph vessels
3. Increased capillary permeability
4. Decreased concentration of plasma proteins

Increased capillary hydrostatic pressure promotes filtration and inhibits reabsorption. As a result, excess fluid is forced out of the capillary into the interstitial space. An increase in capillary pressure is generally caused by an increase in venous pressure. For example, under conditions of right-sided congestive heart failure, the heart cannot pump all of the blood returned to it. Consequently, the blood becomes backed up in the venous system. This increases the hydrostatic pressure of both the veins and the capillaries, particularly in the lower extremities. Left-sided congestive heart failure may cause pulmonary edema.

Another condition that can impair venous return is pregnancy. As the uterus enlarges during gestation, it may cause compression of the veins draining the lower extremities. Once again, venous pressure and capillary pressure are increased. Filtration is enhanced, reabsorption is inhibited, and edema develops in the lower extremities.

Blockage of lymph vessels prevents the return of the excess filtered fluid to the vascular compartment. Instead, this fluid remains within the tissue. Impaired lymph drainage may be caused by local inflammation, cancer, and parasites.

Increased capillary permeability may allow plasma proteins to leak into the interstitial spaces of a tissue. The presence of excess protein in these spaces causes an increase in the interstitial fluid colloid osmotic pressure and pulls more fluid out of the capillaries. Mediators of inflammation, such as histamine and bradykinin, which are active following tissue injury and during allergic reactions, increase capillary permeability, and cause swelling.

A *decrease in the concentration of plasma proteins* causes a decrease in the plasma colloid osmotic pressure. As a result, filtration is increased, reabsorption is decreased, and fluid accumulates in the tissue. Most plasma proteins are made in the liver. Therefore, a decrease in protein synthesis due to liver failure is an important cause of this condition. Malnutrition may also impair protein synthesis. Finally, kidney disease leading to proteinuria (protein loss in the urine) decreases the concentration of plasma proteins.

Medical terminology

Acidosis (ăs″ĭ-dō′sĭs): Increased acidity or increased hydrogen ion concentration in arterial blood.

Baroreceptors (băr′ō-rē-sĕp-tors): In the circulatory system, sensory nerve endings located in the aorta and the carotid arteries which are sensitive to changes in pressure.

Chemoreceptors (kē′mō-rē-sĕp-tors): In the circulatory system, sensory nerve endings located in the aorta and the carotid arteries which are sensitive to changes in arterial oxygen, carbon dioxide, and pH.

Hematocrit (hē-măt′ō-crĭt): Percentage of total blood volume that consists of red blood cells (erythrocytes).

Hypercapnia (hī″pĕr-kăp′nē-ă): Abnormally high level of carbon dioxide in arterial blood.

Hypertension (hī″pĕr-tĕn′shŭn): An increase in blood pressure above normal (higher than 140 mmHg systolic and 90 mmHg diastolic).

Hypotension (hī″pō-tĕn′shŭn): A decrease in blood pressure below normal.

Hypoxemia (hī″pŏks-ē′mē-ă): Abnormally low level of oxygen in arterial blood.

Perfusion (pĕr-fū′zhŭn): Blood flow through a tissue.

Proteinuria (prō″tē-ĭn-ū′rē-ă): Loss of proteins in the urine.

Pulsatile (pŭl′să-tīl): Fluctuating; characterized by a rhythmic beat or throbbing.

Resistance (rĭ-zĭs′tăns): In the circulatory system, the opposition to blood flow through a blood vessel.

Syncope (sĭn′kō-pē): Loss of consciousness.

Vasoactive (vās″ō-ăk′tĭv): Affecting vascular smooth muscle.

Vasomotion (vās″ō-mō′shŭn): Change in diameter of a blood vessel.

Bibliography

Armstrong, A. W., Myers, C. W., Yeh, D. C., and Rocco, T. P., Integrative Cardiovascular Pharmacology: Hypertension, Ischemic Heart Disease, and Congestive Heart Failure, in *Principles of Pharmacology, The Pathophysiologic Basis of Drug Therapy,* Golan, D. E., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 23.

Benowitz, N. L., Antihypertensive Agents, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 11.

Foegh, M. L., and Ramwell, P. W., The Eicosanoids: Prostaglandins, Thromboxanes, Leukotrienes and Related Compounds, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 18.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hoffman, B. B., Adrenoceptor Antagonist Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 10.

Hoffman, B. B., Therapy of Hypertension, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 32.

Honig, C. R., *Modern Cardiovascular Physiology*, 2nd ed., Little, Brown, Boston, MA, 1988.

Jackson, E. K., Vasopressin and Other Agents Affecting the Renal Conservation of Water, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 29.

Katzung, B. G., and Chatterjee, K., Vasodilators and the Treatment of Angina Pectoris, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 12.

Levy, M. N., and Pappano, A. J., *Cardiovascular Physiology,* 9th ed., Mosby, St. Louis, MO, 2007.

McCulloch, K. M., and McGrath, J. C. Neurohumoral Regulation of Vascular Tone, in *An Introduction to Vascular Biology, From Physiology to Pathophysiology*, Halliday, A., Hunt, B. J., Poston, L., and Schachter, M., Eds., Cambridge University Press, Cambridge, United Kingdom, 1998, chap. 5.

Michael, T., Treatment of Myocardial Ischemia, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 31.

Reid, I. A., Vasoactive Peptides, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 17.

Rhoades, R., and Pflanzer, R., *Human Physiology*, Thomson Learning, Pacific Grove, CA, 2003.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Smyth, E. M., Burke, A., and FitzGerald, G. A., Lipid-Derived Autacoids: Eicosanoids and Platelet-Activating Factor, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 26.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Zdanowicz, M. M., *Essentials of Pathophysiology for Pharmacy*, CRC Press, Boca Raton, FL, 2003.

chapter seventeen

Blood and hemostasis

Study objectives

- Discuss the major functions of the plasma proteins
- Describe the morphological characteristics and the function of erythrocytes
- Explain how various blood types are determined and what blood types are compatible for transfusion
- Describe how the Rh factor may lead to hemolytic disease of the newborn
- Discuss the major functions of each of the five types of leukocytes: neutrophils, eosinophils, basophils, monocytes, and lymphocytes
- Describe the origin of the thrombocytes
- Discuss the role of platelets in the various aspects of hemostasis
- Describe the role of vascular constriction in hemostasis
- Explain how a platelet plug is formed
- Distinguish between the extrinsic pathway and the intrinsic pathway of blood coagulation
- Explain how blood clot growth is limited
- Explain how blood clots are dissolved

17.1 Introduction

Blood consists of cellular elements (red blood cells, white blood cells, and platelets) as well as plasma, the fluid in which the blood cells are suspended. Normally, the total circulating blood volume is about 8% of the body weight (about 5 L in women and 5.5 L in men). Adipose tissue is relatively avascular and, therefore, contains little blood compared to other tissues.

The cellular elements of the blood have a short life span and must be continuously replaced. The formation of red blood cells, white blood cells, and platelets, collectively, is referred to as *hematopoiesis*. This process takes place in the red bone marrow. In adults, red bone marrow is found in the pelvis, the ribs, and the sternum.

17.2 Plasma

The fluid portion of the blood, the *plasma*, accounts for 55% to 60% of the total blood volume. Plasma is about 90% water. The remaining 10% contains

Table 17.1 Composition of Blood

Plasma

- Water (90%)
- Organic Molecules
- Proteins (8%)
  - Albumins (60% of plasma proteins)
  - Globulins (38% of plasma proteins)
  - Fibrinogen
- Amino Acids
- Glucose
- Lipids
- Nitrogenous Wastes
- Ions (Na^+, K^+, Cl^-, H^+, Ca^{++}, HCO_3^-)
- Trace Elements and Vitamins
- Gases (O_2, CO_2)

Cellular Elements

- Red Blood Cells ($3.9–5.6 \times 10^6$ cells/μL; females) ($4.5–6.5 \times 10^6$ cells/μL; males)
- White Blood Cells ($4–11 \times 10^3$ cells/μL)
  - Neutrophils (50% to 70%)
  - Eosinophils (1% to 4%)
  - Basophils (<1%)
  - Monocytes (2% to 8%)
  - Lymphocytes (20% to 40%)
- Platelets ($200–500 \times 10^3$/μL)

proteins (8%) and other substances (2%), including hormones, enzymes, nutrient molecules, gases, electrolytes, and excretory products (see Table 17.1). All of these substances are either dissolved in the plasma (e.g., oxygen) or are colloidal materials (dispersed solute materials that do not precipitate out; e.g., proteins). The three major plasma proteins include:

1. Albumin
2. Globulins
3. Fibrinogen

Albumin is the most abundant (about 60%) of the plasma proteins. An important function of albumin is to bind with various molecules in the blood and serve as a *carrier protein*, transporting these substances throughout the circulation. Substances that bind with albumin include hormones, amino acids, fatty acids, bile salts, and vitamins. Albumin also serves as

an *osmotic regulator*. Because capillary walls are impermeable to plasma proteins, these molecules exert a powerful osmotic force on the water in the blood. In fact, the plasma colloid osmotic pressure exerted by plasma proteins is the only force that retains water within the vascular compartment and, therefore, maintains blood volume (see Chapter 16). Albumin is synthesized in the liver.

The *globulins* account for about 38% of the plasma proteins. There are three types of globulins: alpha (α), beta (β), and gamma (γ). The alpha and beta globulins are involved with several activities. They *transport substances* in the blood (hormones, cholesterol, iron), function as *clotting factors,* and serve as *precursor molecules* (angiotensinogen). The gamma globulins function as *antibodies*, which play an important role in the immune response. The alpha and beta globulins are synthesized in the liver, and the gamma globulins are made by the lymphocytes (a type of white blood cell).

Fibrinogen also plays a role in the blood-clotting process. It serves as a precursor for *fibrin*, which forms the *meshwork of a blood clot*. Fibrinogen is synthesized in the liver.

17.3 Erythrocytes

The most numerous of the cellular elements in the blood are the *erythrocytes* (*red blood cells*). On average, there are 5 million red blood cells per microliter (μL) of blood, or a total of about 25 to 30 trillion red blood cells in the adult human body. The percentage of the blood that is made up of red blood cells is referred to as *hematocrit*. An average hematocrit is about 45% (42% females, 47% males). As such, the *viscosity* of the blood is determined primarily by these elements.

Red blood cells are small, flat biconcave discs. Each cell is approximately 7.5 μ in diameter and 2 μ thick. This shape maximizes the surface area of the cell and facilitates the diffusion of oxygen across the cell membrane. Furthermore, red blood cells are very flexible and easily change their shape. This feature allows them to squeeze through capillaries as narrow as 3 μ in diameter. However, as the red blood cells age, their membranes become quite fragile, and the cells are prone to rupture. Aged cells are removed by phagocytic cells in the *spleen*, *liver*, and *bone marrow.* The average life span of a red blood cell is about 120 days. As such, red blood cells must be replaced at a rate of 2 to 3 million cells per second. Erythrocyte production is regulated by the hormone *erythropoietin*. Low levels of oxygen stimulate the release of erythropoietin from the kidneys into the blood.

The primary function of red blood cells is to transport oxygen to the tissues. The red, oxygen-carrying molecule within the erythrocyte is *hemoglobin*. This molecule has two components: the *globin portion* and the *heme portion*. There are four globin proteins, each of which is bound to a heme group. Each heme group contains an *iron* atom that binds reversibly with oxygen. The average hemoglobin content in the blood is about 15 g/100 ml of blood, all of it within

PHARMACY APPLICATION:
RECOMBINANT HUMAN ERYTHROPOIETIN

Anemia may result from a number of pathophysiologic conditions or their treatments, including:

- Hemodialysis
- Chronic renal failure
- AIDS
- Chemotherapy
- Surgery (e.g., cardiac procedures)

Epoetin alfa is a recombinant human erythropoietin. This compound, which was developed in the late 1980s, is administered by injection. Available preparations include Epogen® and Procrit®. Along with adequate iron intake, this therapy is highly effective in treating these anemias by enhancing red blood cell production. More recently, a longer-acting derivative of Epoetin alfa was developed. Darbepoietin (ARANESP®) remains in the circulation for 24 to 26 hours.

These drugs may also be misused by athletes. In an effort to increase hemoglobin levels and improve their performance, some athletes take these drugs illegally, a practice referred to as "blood doping."

the red blood cells. In fact, because of their high hemoglobin content, each red blood cell has the capacity to transport more than one billion oxygen molecules. This hemoglobin/oxygen-carrying capacity of the red blood cell is facilitated by the lack of a nucleus and any other membranous organelles within these cells.

17.4 Blood types

Erythrocytes are labeled with *cell surface antigens* that determine *blood type*. There are two types of inherited antigens found on red blood cells: A antigens and B antigens. An individual obtains two genes, one from each parent, which determine the production of antigens. Accordingly, there are four possible blood types (see Table 17.2):

1. Type A (A antigen)
2. Type B (B antigen)
3. Type AB (both A and B antigens)
4. Type O (neither A nor B antigens)

Antibodies are specialized molecules produced by the immune system to attack foreign antigens. Therefore, an individual with type A blood produces

Table 17.2 Summary of ABO Blood Type System

Blood Type	Possible Genotypes	Antibodies Produced	Possible Transfusions	Frequency in United States
A	AA, AO	Anti-B	A, O	41%
B	BB, BO	Anti-A	B, O	10%
AB	AB	None	A, B, AB, O	4%
O	OO	Anti-A, Anti-B	O	45%

anti-B antibodies, which attack type B antigens. An individual with type B blood produces anti-A antibodies, which attack type A antigens. Consequently, the mixing of incompatible blood can cause red blood cell destruction. The antibodies produced against a foreign blood type may cause *agglutination* (clumping) or *hemolysis* (rupture) of the donated erythrocytes.

Type AB blood contains both A and B antigens on the red blood cells. Therefore, individuals with this blood type produce neither anti-A nor anti-B antibodies and can receive a transfusion of any blood type. Individuals with type AB blood are referred to as *universal recipients.*

Type O blood contains no antigens on the cell surface. In this case, any antibodies that the transfusion recipient may produce (anti-A or anti-B antibodies) have no antigens to attack. Therefore, there is no immune response against this blood. Individuals with type O blood are referred to as "*universal donors,*" as this blood is suitable for transfusion in all individuals.

Another type of cell surface antigen found on red blood cells is the *Rh factor.* This factor was named for the rhesus monkey in which the factor was first identified. Red blood cells that contain the Rh factor are referred to as *Rh-positive,* and erythrocytes without this factor are referred to as *Rh-negative*.

This antigen also stimulates antibody production. Therefore, Rh-negative individuals who produce anti-Rh antibodies should receive only Rh-negative blood. Rh-positive individuals who do not produce anti-Rh antibodies can receive either Rh-negative or Rh-positive blood. Approximately 85% of Caucasians are Rh-positive and 15% are Rh-negative. Over 99% of Asians, 95% of American blacks, and 100% of African blacks are Rh-positive.

Rh incompatibility may occur when an Rh-negative mother carries an Rh-positive fetus. At the time of delivery, a small amount of the baby's Rh-positive blood may gain access to the maternal circulation. In response, the immune system of the mother produces *anti-Rh antibodies.* During a subsequent pregnancy, the fetus is exposed to these antibodies as they cross the placenta. If this fetus is also Rh-positive, then the anti-Rh antibodies attack the fetal erythrocytes and cause *hemolytic disease of the newborn* (*erythroblastosis fetalis*). This may occur in about 3% of second Rh-positive babies and about 10% of third Rh-positive babies. The incidence continues to increase with subsequent pregnancies.

PHARMACY APPLICATION: ERYTHROBLASTOSIS FETALIS

When an Rh-negative mother has been exposed to Rh-positive blood during birth, miscarriage, or ectopic pregnancy, she may produce anti-Rh antibodies. As discussed previously, if the fetus in a subsequent pregnancy is also Rh-positive, these maternal antibodies may attack the fetal erythrocytes causing erythroblastosis fetalis. This hemolytic disease is characterized by anemia, jaundice, enlargement of the liver and spleen, and generalized edema.

An important advance in immunopharmacology has been the development of a treatment for preventing erythroblastosis fetalis. This treatment is based on the observation that a primary antibody response to the Rh antigen can be blocked if specific anti-Rh antibodies are administered passively at the time of the exposure to the Rh antigen. This technique is a *passive* form of immunization where the injected antibodies inactivate the Rh antigens. As a result, the mother is prevented from becoming *actively* immunized to these antigens.

Antibody production by the mother may be blocked by treating her with Rh_O (D) immune globulin at the time of exposure to the Rh-positive blood. Commercial forms of Rh_O (D) immune globulin, such as RHOGAM®, contain a high titer of anti-Rh antibodies. When women have received this prophylactic treatment, erythroblastosis fetalis has not been observed in subsequent pregnancies. Rh_O (D) immune globulin is administered intramuscularly. The antibodies have a half-life of approximately 21 to 29 days.

17.5 Leukocytes

There are normally 4,000 to 11,000 *leukocytes* (*white blood cells*) per microliter of human blood. However, leukocytes act primarily within the tissues. Those found in the blood are actually in transit. Because of their amoeboid movement, leukocytes can squeeze through pores in the capillary walls and move toward sites of infection. Leukocytes are also found in lymphoid tissues, such as the thymus, spleen, and lymph nodes. These cells are referred to as "white" blood cells because they lack hemoglobin and are essentially colorless.

Leukocytes are an important component of the immune system. General *inflammatory and immune functions* of these cells include the following:

- Destruction of invading microorganisms (bacteria and viruses).
- Identification and destruction of cancer cells.
- Phagocytosis of tissue debris, including dead and injured cells.

The immune system is discussed in detail in Chapter 18. However, many of the functions of the various types of leukocytes are summarized here.

There are five types of leukocytes that are classified as either granulocytes or agranulocytes:

- Granulocytes
  - Neutrophils
  - Eosinophils
  - Basophils
- Agranulocytes
  - Monocytes
  - Lymphocytes

The *granulocytes* are *phagocytic cells*. Their nuclei tend to be segmented into multiple lobes, and the cytoplasm of the cells contains numerous granules. These cells are identified by the staining properties of their granules.

Neutrophils are the most abundant of the leukocytes and account for about 60% of the total number of white blood cells. Mature neutrophils have lobulated nuclei, with two to five lobes connected by thin strands. Because of this unique appearance, neutrophils are also referred to as polymorphonuclear leukocytes (PMNs).

Neutrophils are usually the first to arrive at a site of injury or inflammation. Their primary function is to attack and destroy invading bacteria. In fact, bacterial infection is typically associated with pronounced *neutrophilia* (an increase in the number of circulating neutrophils). These leukocytes are also involved in the removal of tissue debris and, therefore, play a role in the healing process.

Neutrophils eliminate bacteria and tissue debris by way of *phagocytosis*. Small projections of the cell membrane extend outward and engulf the harmful organisms and particles. As a result, these materials are internalized within a cell-membrane-bound vesicle. A lysosome, an organelle filled with hydrolytic enzymes, then fuses with the vesicle. In this way, the phagocytized material is degraded by these enzymes without any damage to the rest of the cell. Neutrophils have the capacity to phagocytize 5 to 25 bacteria before they also die.

Eosinophils, which constitute only 1% to 4% of the total number of white blood cells, are only weak phagocytes. These leukocytes are produced in large numbers in individuals with *internal parasitic infections*. The eosinophils attach to the parasites and secrete substances that kill them. These substances include the following:

- Hydrolytic enzymes released from eosinophil granules (which are actually modified lysosomes).
- Highly reactive forms of oxygen that are particularly lethal.
- Major basic protein, a larvicidal polypeptide also released from granules.

Eosinophils also tend to accumulate at the sites of allergic reactions, particularly in the lungs and the skin. The functions of the eosinophils in these areas include the neutralization of inflammatory mediators released from mast cells as well as the phagocytosis of allergen–antibody complexes. In this way, the spread of the inflammatory reaction is limited.

Basophils are the least abundant of the leukocytes and account for less than 1% of the total number of white blood cells. They are similar structurally and functionally to the *mast cells* found in connective tissues, especially in the lungs, skin, and gastrointestinal tract.

Basophils and mast cells play an important role in allergic reactions. The granules of these cells contain many substances:

- Heparin, which prevents blood coagulation.
- Histamine, which promotes bronchoconstriction as well as the vasodilation and increased capillary permeability that lead to inflammation.

The leukocytes classified as *agranulocytes* contain very few granules in their cytoplasm. In further contrast to the granulocytes, these cells have a single, large, nonsegmented nucleus.

Monocytes are the largest of the leukocytes and account for about 5% of the total number of white blood cells in the blood. These leukocytes, which are immature in the blood, leave the vascular compartment and enter the tissues. Within the tissues, they enlarge, mature, and develop into *macrophages*. Macrophages are large phagocytic cells that can ingest bacteria, necrotic tissue, and even dead neutrophils. These cells survive much longer than neutrophils and may ingest up to 100 bacteria. The life span of the macrophage may range from months to years until it is ultimately destroyed as the result of phagocytic activity.

Lymphocytes are typically the second most numerous type of leukocyte and constitute about 30% of the total number of white blood cells. There are two types of lymphocytes: B lymphocytes and T lymphocytes: The primary function of the *B lymphocytes* is to produce *antibodies*, which are molecules that identify and lead to the destruction of foreign substances, such as bacteria. The B lymphocytes and the antibodies they produce are responsible for *humoral immunity*.

The *T lymphocytes* provide immunity against viruses and cancer cells. These lymphocytes directly attack and destroy their targets by forming holes in the target cell membrane, causing cell lysis. The T lymphocytes are responsible for *cell-mediated immunity*.

17.6 Platelets

The third of the cellular elements within the blood are the *platelets* (*thrombocytes*). Platelets are actually small, round, or oval *cell fragments*. They are about 2 to 4 μ in diameter and have no nuclei. Platelets are formed in the

red bone marrow as pinched-off portions of the very large *megakaryocytes*. Each megakaryocyte, which is confined to the bone marrow, can produce up to 1000 platelets. There are approximately 300,000 platelets per microliter of blood. They are replaced about once every 10 days.

Platelets are essential for many aspects of *hemostasis*, or the cessation of blood loss. They play an important role in blood clotting. In fact, they constitute much of the mass of the clot. Several substances are found within the cytoplasm of platelets that contribute to the arrest of bleeding as well as vessel repair:

- *Actin and myosin molecules, thrombosthenin*: Contractile proteins that enable platelets to contract.
- *Fragments of the endoplasmic reticulum and the Golgi apparatus*: Produce enzymes and store calcium.
- *Mitochondria and enzyme systems*: Form adenosine triphosphate (ATP) and adenosine diphosphate (ADP).
- *Enzyme systems that produce prostaglandins*: Substances involved with both the formation of platelet plugs as well as the limitation of clot growth.
- *Fibrin-stabilizing factor*: Protein involved with blood coagulation.
- *Growth factor*: Facilitates vascular endothelial cell, vascular smooth muscle cell, and fibroblast multiplication and growth, leading to the repair of damaged blood vessels.

These substances are discussed more fully in the following section.

17.7 Hemostasis

The prevention of blood loss from a damaged blood vessel is referred to as *hemostasis*. There are three inherent mechanisms that contribute to hemostasis:

1. Vascular constriction
2. Formation of a platelet plug
3. Blood coagulation

17.7.1 Vascular constriction

The first mechanism to occur is *vascular constriction*. Immediately after a blood vessel is cut or severed, the vascular smooth muscle automatically constricts. This results in a decrease in the flow of blood through the vessel which helps to limit blood loss. The vasoconstriction is caused by several factors:

- Sympathetic nerve reflexes in response to pain.
- Local myogenic vasospasm in response to the injury.
- Locally produced vasoconstrictors released from the damaged tissue and from platelets.

PHARMACY APPLICATION: ANTIPLATELET DRUGS

Platelets play a role in each of the mechanisms of normal hemostasis: vasoconstriction, formation of the platelet plug, and blood coagulation. However, they are also involved in pathological processes that lead to atherosclerosis and thrombosis (formation of a blood clot within the vascular system). Antiplatelet drugs interfere with platelet function and are used to prevent the development of atherosclerosis and the formation of arterial thrombi.

The prototype of antiplatelet drugs is aspirin. Aspirin inhibits cyclooxygenase, an enzyme involved in arachidonic acid metabolism. Inhibition of cyclooxygenase blocks the synthesis of thromboxane A_2, the platelet product that promotes both vasoconstriction and platelet aggregation. Because platelets are simply cell fragments, they are incapable of synthesizing new proteins, including enzymes. Therefore, the aspirin-induced inhibition of cyclooxygenase is permanent and lasts for the life of the platelet (7 to 10 days).

Aspirin is maximally effective as an antithrombotic agent at the comparatively low dose of 160 to 320 mg per day. (The antipyretic dose of aspirin in adults is 325 to 650 mg every 4 hours.) Low doses of aspirin cause a steady-state decrease in *platelet* cyclooxygenase activity. Higher doses of aspirin are actually contraindicated in patients prone to thromboembolism. At higher doses, aspirin also reduces the synthesis of prostacyclin, another arachidonic acid metabolite. Prostacyclin, produced by the *endothelium*, normally inhibits platelet aggregation.

The prophylactic administration of low-dose aspirin has been shown to increase survival following myocardial infarction, decrease the incidence of stroke, and assist in the maintenance of patency of coronary bypass grafts.

When the extent of the trauma to the vessel is increased, then the degree of vascular constriction is increased. Accordingly, a sharply cut blood vessel bleeds far more profusely than a blood vessel damaged by a more crushing injury. The vasoconstriction may last for many minutes or hours. This provides time for the two subsequent mechanisms to develop and get under way.

17.7.2 Formation of a platelet plug

The *formation of a platelet plug* physically blocks small holes in blood vessels. Platelets are typically unable to adhere to the endothelial lining of the blood vessels. The surface of the platelets contains a coat of glycoproteins that repels the normal endothelium. Interestingly, these same glycoproteins enable the platelets to adhere to damaged vessels.

When platelets come into contact with a damaged vascular surface, in particular, collagen fibers in the vessel wall or damaged endothelial cells, the platelets become activated. These platelets become "sticky" and adhere to the damaged tissue. Binding of the platelets to the collagen is facilitated by *von Willebrand's factor* (*vWF*). This protein, which is secreted by endothelial cells and platelets, binds to both the exposed collagen and the platelets. In this way, vWF acts as a bridge between the damaged vessel wall and the platelets. The adhered platelets also release ADP and thromboxane A_2, a prostaglandin metabolite, which enhance the stickiness of other platelets. Consequently, more and more platelets adhere to the damaged vessel, ultimately forming a plug. This process is also referred to as *agglutination*. Furthermore, thromboxane A_2, as well as serotonin, also released from the platelets, contribute to the initial mechanism of vasoconstriction.

17.7.3 Blood coagulation

The third major step in hemostasis is *coagulation*, or the formation of a blood clot. This complex process involves a series of reactions that results in the formation of a protein fiber meshwork that stabilizes the platelet plug.

There are three essential steps that lead to clotting (see Figure 17.1):

1. Activation of factor X.
2. Conversion of prothrombin into thrombin.
3. Conversion of fibrinogen into fibrin.

The resulting blood clot contains platelets and the physically stabilizing fibrin. In addition, this meshwork usually traps erythrocytes that give the clot its red color.

Altogether, there are 12 *clotting factors* in the plasma. These factors, which are proteins synthesized in the liver, are normally found circulating in the plasma in their inactive forms. The activation of one of these factors leads to the activation of another factor, and so on, resulting in a cascade of reactions culminating in the formation of fibrin. Interestingly, these factors are not numbered for their place in the clotting cascade. Instead, they are numbered in order of their discovery.

Activated *factor X*, along with Ca^{++} ion, factor V, and platelet factor 3 (collectively referred to as the *prothrombin activator*), catalyzes the conversion of *prothrombin* into *thrombin*. Thrombin then catalyzes the conversion of *fibrinogen* into *fibrin*, an insoluble, thread-like polymer. The fibrin threads form a meshwork that traps blood cells, platelets, and plasma to form the blood clot.

There are two mechanisms by which the clotting cascade may be elicited (see Figure 17.1): the *extrinsic pathway* and the *intrinsic pathway*.

The *extrinsic pathway* of blood coagulation begins when a blood vessel is ruptured and the surrounding tissues are damaged. The traumatized tissue releases a complex of substances referred to as *tissue thromboplastin*. The

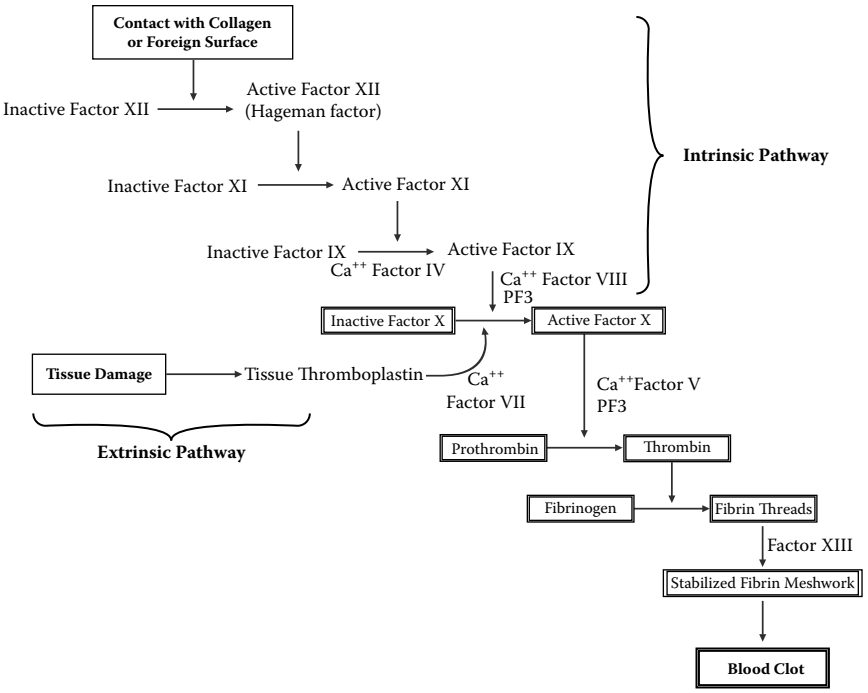


Figure 17.1 The coagulation pathways. Blood coagulation may be elicited by two pathways occurring independently or, more often, concurrently. The intrinsic mechanism begins when blood comes into contact with the collagen in a damaged vessel wall or with a foreign surface (e.g., test tube). This causes the activation of factor XII, or Hageman Factor, followed by the activation of other clotting factors and, finally, factor X. The extrinsic pathway occurs when damaged tissue releases tissue thromboplastin. This pathway activates factor X directly. The activation of factor X leads to the conversion of prothrombin into thrombin. Thrombin then leads to the conversion of fibrinogen into fibrin threads. The fibrin forms the stabilized meshwork that traps blood cells and forms the blood clot. (PF3, platelet factor 3.)

tissue thromboplastin further complexes with factor VII and Ca^{++} ions to activate factor X directly.

The *intrinsic pathway* of blood coagulation causes the blood to clot within the vessel. It is activated when the blood is traumatized or when it comes into contact with the exposed collagen of a damaged vessel wall. This contact activates *factor XII* (*Hageman Factor*) in the blood. Simultaneously, platelets are activated, such that they begin adhering to the collagen in the vessel wall to form the platelet plug. In addition to ADP and thromboxane A_2, these aggregated platelets also release *platelet factor 3*. This substance plays a role in subsequent clotting reactions. (It is important to note at this point that platelets are involved in all three mechanisms of hemostasis: vascular constriction, formation of the platelet plugs, and blood coagulation.)

Activated factor XII leads to the activation of *factor XI*. Activated factor XI, along with Ca^{++} ions and factor IV, leads to the activation of *factor IX*. Activated factor IX, along with Ca^{++} ions, factor VIII, and platelet factor 3, leads to the activation of factor X. From the point of factor X activation, the extrinsic and intrinsic mechanisms follow the same common pathway to fibrin formation.

The extrinsic pathway and the intrinsic pathway typically occur simultaneously. The extrinsic mechanism coagulates the blood that has escaped into the tissue prior to the sealing of the vessel. The intrinsic mechanism coagulates the blood within the damaged vessel. Another important difference involves the speed at which these two mechanisms cause coagulation. Because the extrinsic mechanism causes the activation of factor X directly, clotting begins within seconds. The intrinsic mechanism is much slower, usually requiring 1 to 6 minutes to form a clot. However, the cascade of reactions characteristic of this mechanism allows for amplification. Each molecule of a given activated clotting factor may activate many molecules of the clotting factor in the next step of the cascade. Therefore, a few molecules of activated Hageman Factor can lead to the activation of hundreds of molecules of factor X and a very powerful coagulation response.

Once the clot is formed, the platelets trapped within it contract, shrinking the fibrin meshwork. This *clot retraction* pulls the edges of the damaged vessel closer together.

Blood coagulation is limited to the site of damage. Once the blood-clotting factors have carried out their activities, they are rapidly inactivated by enzymes present in the plasma and in the surrounding tissue.

17.7.4 Positive feedback nature of clot formation

Thrombin promotes clot formation at several points in the coagulation cascade through *positive feedback*. Activities of thrombin include the following:

- Acting on prothrombin to make more thrombin, thus facilitating its own formation.
- Accelerating the actions of several blood-clotting factors (VIII, IX, X, XI, and XII).
- Enhancing platelet adhesion and activation.
- Activating factor XIII, which strengthens and stabilizes the fibrin meshwork of the clot.

17.7.5 Clot dissolution

Once the blood vessel has been repaired, the clots must be removed in order to prevent permanent obstruction. *Plasmin* is a proteolytic enzyme that digests fibrin (*fibrinolysis*). It is synthesized from its precursor, *plasminogen*. The conversion of plasminogen into plasmin involves several substances, including factor XII (Hageman Factor), which are also involved in the coagulation

cascade. Within a few days after the blood has clotted, enough plasmin has been formed to dissolve it. The residue of the clot dissolution is removed by the phagocytic white blood cells, neutrophils, and macrophages.

PHARMACY APPLICATION: ANTICOAGULANT DRUGS

Anticoagulant drugs include heparin and warfarin (Coumadin®). These agents are used to prevent deep vein thrombosis. They are also used to prevent the formation of emboli due to atrial fibrillation, valvular heart disease, and other cardiac disorders.

Heparin, which is not absorbed by the gastrointestinal tract, is available only by injection. Its mechanism of action involves the activation of *antithrombin III*. This plasma protein binds with and inactivates thrombin and several other clotting factors. The effect of heparin is immediate.

The most commonly used oral anticoagulant drug in the United States is warfarin. This agent acts by altering vitamin K such that it is unavailable to participate in the synthesis of vitamin-K-dependent coagulation factors in the liver (coagulation factors II, VII, IX, and X). Because of the presence of preformed clotting factors in the blood, the full antithrombotic effect of warfarin therapy may require 36 to 72 hours.

The major adverse effect of warfarin is bleeding. (Ironically, this compound was originally introduced as a very effective rodenticide. As the active ingredient in rodent poison, is causes death due to internal hemorrhaging.) Furthermore, it readily crosses the placenta and can cause a hemorrhagic disorder in the fetus. Therefore, it is contraindicated in pregnant women.

17.7.6 Prevention of blood clotting and platelet aggregation in the normal vascular system

Several factors contribute to the prevention of blood clotting in the normal vascular system:

- *Smoothness* of the endothelial lining: Prevents the contact activation of the intrinsic mechanism.
- Layer of *glycocalyx* on the endothelium: Repels clotting factors and platelets
- *Thrombomodulin*: Protein on the endothelium that (1) binds with thrombin, reducing its availability for the clotting process, and (2) activates protein C, which acts as an anticoagulant by inactivating factors V and VIII.

- *CD 39*: Enzyme on the endothelium that breaks down ADP in the blood to adenosine monophosphate (AMP) and phosphate (recall that ADP promotes platelet aggregation).
- *Prostacyclin* and *nitric oxide*: Secreted by the endothelium, these substances act as vasodilators and inhibitors of platelet aggregation.
- *Tissue plasminogen activator*: Activates plasmin to dissolve fibrin that is continuously made at low levels.

Medical terminology

Agglutination (ă-gloo″tĭ-nā′shŭn): Clumping of platelets.
Antibody (ăn′tĭ-bŏd″ē): Immunoglobulin molecule produced by B lymphocytes (plasma cells) that destroys or neutralizes antigens.
Anticoagulant (ăn″tĭ-kō-ăg′ū-lănt): Agent that prevents coagulation of the blood.
Antigen (ăn′tĭ-jĕn): Protein marker on the cell surface that allows the immune system to distinguish between "self" cells and "nonself" cells.
Antipyretic (ăn-tĭ-pī-rĕt′ĭk): Agent that reduces fever.
Coagulation (kō-ăg″ū-lā′shŭn): Formation of a blood clot.
Erythrocyte (ĕ-rĭth′rō-sīt): Mature red blood cell.
Erythropoietin (ĕ-rĭth″rō-poy′ĕ-tĭn): Hormone made by the kidney that stimulates red blood cell production in the bone marrow.
Fibrinolysis (fī″brĭn-ŏl′ĭ-sĭs): Breakdown of fibrin in a blood clot.
Hematopoiesis (hĕm″ă-tō-poy-ē′sĭs): Production of blood cells.
Hemolysis (hē-mŏl′ĭ-sĭs): Destruction of red blood cells.
Hemostasis (hē″mō-stā′sĭs): Prevention of blood loss.
Larvicidal (lăr-vĭ-sī′dăl): Destructive to larva.
Leukocyte (loo′kō-sīt): White blood cell.
Necrotic (nĕ-krŏt′ĭk): Relating to tissue death or destruction.
Neutrophilia (nū″trō-fĭl′ē-ă): Abnormal increase in neutrophils in the blood.
Phagocytosis (făg″ō-sī-tō′sĭs): Process by which phagocytes engulf and destroy microorganisms (e.g., bacteria) and cellular debris.
Plasma (plăz′mă): Liquid portion of the blood in which the blood cells are suspended.
Thrombocyte (thrŏm′bōsīt): Platelet.
Thrombosis (thrŏm-bō′sĭs): Formation of a blood clot within the vascular system.

Bibliography

Fox, S. I., *Human Physiology*, 9th ed., McGraw-Hill, New York, 2006.
Ganong, W. F., *Review of Medical Physiology*, 19th ed., Appleton & Lange, Stamford, CT, 1999.

Gaspard, K. J., Composition of Blood and Formation of Blood Cells, in *Pathophysiology, Concepts of Altered Health States*, 7th ed., Porth, C. P., Ed., Lippincott, Philadelphia, PA, 2005, chap. 14.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hambleton, J., and O'Reilly, R. A., Drugs Used in Disorders of Coagulation, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 34.

Kaushansky, K., and Kipps, T. J., Hematopoietic Agents: Growth Factors, Minerals, and Vitamins, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 53.

Krensky, A. M., Vincenti, F., and Bennett, W. M., Immunosuppressants, Tolerogens, and Immunostimulants, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 52.

Lake, D. F., Akporiaye, E. T., and Hersh, E. M., Immunopharmacology, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 56.

Majerus, P. W., Broze, G. J., Miletich, J. P., and Tollefsen, D. M., Anticoagulant, Thrombolytic, and Antiplatelet Drugs, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 9th ed., Hardman, J. G., and Limbird, L. E., Eds., McGraw-Hill, New York, 1996, chap. 54.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2001.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, 2005.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter eighteen

The immune system

Study objectives

- Describe the functions and activities of the immune system
- List the symptoms of inflammation
- Explain the steps of the inflammatory response and how they protect against bacterial invasion
- Explain how interferon defends against invading viruses and tumor cells
- Describe the function of natural killer (NK) cells
- Describe how the complement system defends against invading bacteria
- Compare and contrast the general characteristics of the innate immune system and the adaptive immune system
- List the different types of antibodies and their functions
- Describe the two functional regions of the antibody molecule
- Explain how antibodies defend against invading bacteria
- Explain the clonal selection theory
- Compare and contrast primary responses and secondary responses
- Distinguish between active immunity and passive immunity
- List the three types of T cells and their functions
- Explain how cytotoxic T cells defend against invading viruses
- Distinguish between class I major histocompatibility complex (MHC) glycoproteins and class II MHC glycoproteins

18.1 Introduction

Immunity is defined as the body's ability to eliminate foreign organisms or substances as well as abnormal cells. The immune system consists of tissues, cells, and molecules that work together to form an internal defense system. As such, it is capable of recognizing and then destroying or neutralizing organisms and substances that are foreign to the "normal self." The immune system performs several functions and activities that are important in maintaining health or causing disease:

- Defense against infection
- Removal of cells and tissue debris
- Immune surveillance
- Allergies and autoimmune diseases
- Rejection of tissue grafts and newly introduced proteins

The immune system provides *defense against infection* by pathogenic microorganisms such as bacteria, viruses, fungi, and parasites. Immune deficiency results in an increased susceptibility to infections. This is exemplified by patients with AIDS or those receiving chemotherapy or taking immunosuppressant drugs.

Another function of this system involves the *removal* of "worn out" cells, such as aged red blood cells, as well as tissue debris. Injury and disease cause tissue damage and cell death. The removal of the tissue debris is an important step in wound healing and the tissue repair process.

Immune surveillance involves the identification and destruction of abnormal or mutant cells that have originated within the body. This function serves as the primary internal defense mechanism against tumor growth and cancer.

Allergies and autoimmune diseases are inappropriate immune responses. An allergy, or type I hypersensitivity reaction, is a response to a seemingly harmless environmental agent (ragweed, dust, pet dander), food (nuts, shellfish), or drug (penicillin). Symptoms may include nasal congestion, sneezing, runny nose, asthma, hives, diarrhea, and anaphylaxis. Autoimmune diseases occur when the immune system erroneously produces antibodies against a particular type of the body's own cells, which damages tissues and disrupts normal organ function. Examples of autoimmune diseases include rheumatoid arthritis, systemic lupus erythematosus, type 1 diabetes mellitus, and Crohn's disease. Symptoms of each of these diseases are related to the organ systems affected.

Tissues and organs transplanted from a donor are recognized by the immune system as being foreign to the body. The resulting immune responses, or type II hypersensitivity reactions, may damage and destroy these tissues. Therefore, the recipient of a transplant must inhibit these responses with the administration of immunosuppressive drugs such as cyclosporine, rapamycin, mycophenolic acid, and leflunomide.

18.2 Agents of infectious disease

Many microorganisms, which are not visible to the human eye, are capable of causing infectious disease. These include bacteria, viruses, fungi, and parasites. This chapter will focus on defense provided by the immune system against bacteria and viruses.

Bacteria are nonnucleated, single-celled organisms. They consist of a single chromosome of DNA and a cytoplasm containing the reproductive and metabolic machinery of the cell. As such, bacteria can synthesize DNA, RNA, and proteins and can reproduce independently. These microorganisms cause disease primarily by releasing enzymes or toxins that physically injure or functionally disrupt cells and tissues.

Viruses are pathogens consisting of nucleic acid (DNA or RNA) enclosed by a protein coat. These microorganisms lack the cellular machinery for protein

synthesis and energy production. As such, they are incapable of reproducing outside of a living cell. Therefore, viruses must invade a susceptible living "host" cell and use the biosynthetic machinery of that cell in order to replicate. Viruses may damage tissues and cause disease in many ways, including the following:

- Destruction of the host cell by the immune system in order to gain access to and destroy the virus.
- Depletion of essential cellular components.
- Production by the cell of substances that are toxic to itself.
- Transformation of the cell into a cancer cell.

18.3 Effector cells of the immune system

Leukocytes (white blood cells) are the effector cells of the immune system and are responsible for the destruction of invading microorganisms, identification and destruction of cancer cells, and phagocytosis of tissue debris including dead and injured cells. Leukocytes are present in the blood only transiently. These cells leave the vascular compartment and enter the tissues, moving toward areas of inflammation or infection where they carry out their effects. The structure and function of white blood cells are discussed in some detail in Chapter 17. Briefly, the primary characteristics and functions of the five types of leukocytes are as follows:

- *Neutrophils*: Highly mobile phagocytes that attack and destroy invading bacteria and remove tissue debris.
- *Monocytes*: Immature leukocytes that leave the blood, enter the tissue, and transform into large, tissue-bound phagocytes.
- *Eosinophils*: Granulocytes that destroy parasitic worms and are involved in allergic reactions.
- *Basophils*: Leukocytes that are structurally and functionally similar to connective tissue mast cells such that they release histamine and heparin and are involved in allergic reactions.
- *Lymphocytes*:
  - *B lymphocytes*: Transform into plasma cells that produce antibodies; antibodies lead to the indirect destruction or neutralization of bacteria and bacterial toxins.
  - *T lymphocytes*: Are involved in the direct destruction of virus-invaded cells and mutant cells.

Other effector cells include *natural killer* (NK) *cells* and *mast cells*. These effector cells are discussed in detail in subsequent sections of this chapter.

18.4 Immune responses

The immune system has two major components: the *innate immune system* and the *adaptive immune system*. Working together to protect the body from infection, these two types of immune systems elicit responses that differ in their timing and selectivity.

18.5 Innate immune system

Also referred to as the *nonspecific immune system*, the innate immune system consists of inherent, built-in defense mechanisms that *nonselectively* defend the body against foreign material and substances. Because these responses are nonspecific, they are elicited *immediately* and are broadly reacting. These mechanisms provide the initial defense against infectious agents, chemical irritants, and tissue injury.

Neutrophils and macrophages play a key role in the innate immune system. These phagocytes contain *toll-like receptors* (*TLRs*) on their plasma membranes. There are several different types of TLRs, each of which recognizes different molecular patterns expressed by pathogens (PAMPs). These PAMPs are shared by many infectious agents, which reduces the need for a large number of receptor types. Furthermore, PAMPs are clearly distinguishable from "self" molecular patterns on the surface of the body's own cells. Engagement of the PAMP with the TRL on the phagocyte triggers phagocytosis and the secretion of cytokines that stimulate inflammation.

In summary, the innate immune system elicits rapid and nonspecific defense mechanisms. However, the effectiveness of these mechanisms is limited in that the responses are not very powerful. The benefit of the innate immune system is that it contains and limits the spread of infection until the slower, but more powerful, adaptive immune system is activated.

Innate immune responses include the following:

- Inflammation
- Interferon
- Natural killer cells
- Complement system

18.6 Inflammation

The inflammatory response is a complex reaction involving the accumulation of phagocytes and plasma proteins at a site of infection, toxin exposure, or tissue injury. The purpose of this response is to:

- Isolate, destroy, or inactivate infectious microbes.
- Remove tissue debris.
- Promote tissue repair.

The cardinal signs or symptoms of foremost importance in regard to inflammation include the following:

- *Rubor*: Redness due to increased blood flow to the affected area.
- *Tumor*: Swelling due to increased capillary permeability and fluid accumulation in the affected area.
- *Calor*: Increased temperature (heat) due to the increased blood flow.
- *Dolor*: Pain.
- *Functio laesa*: Altered or impaired function of the affected area.

The inflammatory response involves a series of steps that is similar in most instances.

18.6.1 Defense by tissue macrophages

A small number of macrophages may be found in the body's tissues at any given time. These macrophages provide the first step in protection from infection by phagocytizing invading microbes.

18.6.2 Localized vasodilation

Vasodilation of arterioles increases blood flow to the affected area. As a result, the delivery of phagocytes and plasma proteins is also increased. This vasodilation is caused primarily by histamine released from activated mast cells as well as activated bradykinin. Mast cells are found in connective tissues, particularly in areas of potential microbial entry to the body such as the lungs, skin, and gastrointestinal tract. As mentioned previously, vasodilation in the inflamed area leads to redness and heat.

18.6.3 Increased capillary permeability

In addition to vasodilation, histamine and bradykinin cause the pores between the endothelial cells to become larger. As a result, capillary permeability increases, and plasma proteins escape into the tissue spaces.

18.6.4 Localized edema

Vasodilation and increased capillary permeability lead to localized edema. As discussed in Chapter 16, vasodilation and increased blood flow lead to an increase in capillary pressure (P_c). Furthermore, the presence of escaped plasma proteins within the tissue spaces leads to an increase in interstitial fluid colloid osmotic pressure (π_i). The increase in P_c and π_i increases the filtration and reduces the reabsorption of fluid. In other words, fluid moves out of the capillaries and accumulates in the tissue. In addition to redness

and heat (due to increased blood flow), symptoms of edema include swelling and pain (due to distension of the tissue).

18.6.5 Walling off of the inflamed area

Several types of plasma proteins escape from the vascular compartment and enter the tissue spaces. Of particular interest in the inflammatory response are the *clotting* and *anticlotting factors* that circulate in the blood. As discussed in Chapter 17, the inactive plasma protein *fibrinogen* is converted into the active *fibrin* that forms clots within the tissue spaces. This effectively walls off the injured area and prevents or delays the spread of infection. Subsequently, anticlotting factors, which are activated more slowly, dissolve these clots when they are no longer needed.

18.6.6 Infiltration of phagocytes

Increased numbers of phagocytes are needed in order to engulf and destroy infectious microbes, remove tissue debris, and prepare the injured area for healing and repair. This step is facilitated by the increase in blood flow and the increase in capillary permeability. Neutrophils arrive first, typically within 1 hour. Monocytes arrive within 8 to 12 hours and proceed to swell and mature into macrophages within the next 8 to 12 hours.

The emigration of phagocytes from the vascular compartment toward the injured area involves three steps:

- *Margination*: Phagocytes adhere to the endothelial cells of the capillary wall.
- *Diapedesis*: Phagocytes squeeze through the capillary pores and enter the tissue space.
- *Chemotaxis*: Phagocytes move through the tissue up a concentration gradient of *chemotaxins* that are released at the site of injury.

18.6.7 Opsonization

Opsonization is the process by which bacteria are marked for phagocytosis. The two most important *opsonins* or chemicals that bind to and label the bacteria are *antibodies* and *complement protein C3b.* Opsonization by way of antibodies will be discussed later in this chapter, as this is a form of adaptive immunity. Opsonization by way of C3b is a form of innate immunity. The concentration of C3b in the injured area is increased due to the increased capillary permeability and leakage of plasma proteins into the tissue space. This complement factor binds nonspecifically with the invading bacteria. In addition, there are receptors specific for C3b on the surface of the phagocytes. Binding of the C3b with its receptor creates a linkage between the

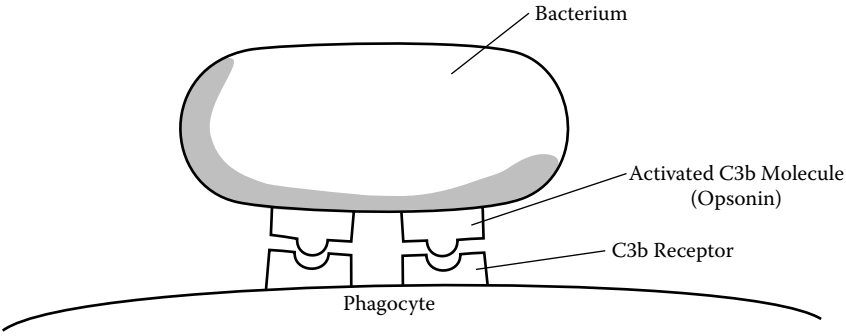


Figure 18.1 Opsonization. Activated complement factor C3b binds nonspecifically with an invading microbe such as a bacterium. C3b also binds with a receptor on the surface of a phagocyte. In this way, the bacterium and the phagocyte are linked, and phagocytosis is facilitated.

bacterium and the phagocyte that prevents the "escape" of the bacterium (Figure 18.1). In this way, bacteria are more readily and more efficiently engulfed by phagocytes.

18.6.8 Phagocytosis

Phagocytosis is the process of ingestion and digestion of bacteria, foreign particles, and tissue debris. These substances are internalized within a vesicle formed from the plasma membrane. The vesicle then fuses with a *lysosome*. Phagocytes have an abundance of lysosomes, which are organelles filled with hydrolytic enzymes. The enzymes then degrade the substances within the vesicle. Inevitably, some amount of these destructive enzymes escapes into the cytoplasm and kills the phagocyte. Neutrophils are capable of engulfing 5 to 25 bacteria before they are killed by the enzymes. Macrophages engulf and destroy as many as 100 bacteria before they die. *Pus* is a fluid found in an infected wound that consists of living and dead phagocytes, necrotic tissue, and bacteria.

18.7 Interferon

A second form of innate immune response involves the release of *interferon* from virus-infected cells. This protein helps to contain viral infection by interfering with the replication of viruses in other host cells. Interferon binds to specific receptors on uninfected cells. This receptor binding leads to the synthesis of over two dozen proteins that contribute to viral resistance via multiple mechanisms. These mechanisms include the inhibition of transcription, translation, protein processing, and virus maturation. As a result, when the virus enters the interferon-altered cell, it is unable to replicate and multiply.

PHARMACY APPLICATION: ANTIINFLAMMATORY DRUGS

In some instances, the inflammatory response may be exaggerated, prolonged, or inappropriate. Inflammation in these situations is without benefit and may, in fact, cause serious injury to the tissue or impair the tissue's function. Nonsteroidal antiinflammatory drugs (NSAIDs) include a chemically heterogeneous group of compounds with antiinflammatory, analgesic, and antipyretic effects. The prototype drug in this class is aspirin. The major mechanism of action of these drugs involves the inhibition of cyclooxygenase, the enzyme involved in the synthesis of the prostaglandins. This family of compounds contributes to and exacerbates the inflammatory response. Most over-the-counter medications, such as ibuprofen (Advil®, Motrin®) and naproxen (Aleve®), inhibit both cyclooxygenase 1 (COX-1) and cyclooxygenase 2 (COX-2). The inhibition of COX-2 mediates the therapeutic effects of these drugs. Although COX-1 inhibition also leads to therapeutic effects via prevention of platelet aggregation, this inhibition is also primarily responsible for the undesirable side effects of these drugs such as gastric ulcers.

More recently, selective COX-2 inhibitors have been developed. These substances, which include celecoxib (Celebrex®), rofecoxib (Vioxx®), and valdecoxib (Bextra®), were very effective therapeutically and caused fewer gastric side effects. However, it was not until these drugs were widely prescribed to patients that it was determined that they may be associated with an increased risk of serious cardiovascular events such as heart attack and stroke. Both rofecoxib and valdecoxib have been removed from the market.

Glucocorticoids can prevent or reduce inflammation in response to mechanical, chemical, infectious, and immunological stimuli. In fact, these drugs suppress almost every aspect of the inflammatory response. They are useful in treating inappropriate immune responses such as allergic reactions (poison ivy rash, asthma) and the inflammation associated with arthritis. However, it should be noted that the use of glucocorticoids does not address the underlying cause of the inflammation, only the symptoms.

The mechanism of action of the glucocorticoids involves decreased release of substances that contribute to the inflammatory response, including histamine, prostaglandins, cytokines, and endothelial leukocyte adhesion molecule-1 (ELAM-1). As a result, vasodilation, leukocyte extravasation, and chemotaxis are all reduced. In addition, the activation of T cells and antibody production is decreased. Therefore, an undesired effect of this immunosuppression is an accompanying increased risk of infection.

Other functions of interferon include the following:

- Increasing macrophage activity.
- Increasing production of antibodies.
- Increasing activity of natural killer cells and cytotoxic T cells.
- Decreasing cell division and decreasing tumor growth.

18.8 Natural killer cells

Natural killer (NK) cells are lymphocytes that are distinct from B cells and T cells. NK cells nonspecifically destroy virus-infected cells and tumor cells. These cells are referred to as *natural killer cells* because, unlike cytotoxic T cells, the recognition of a specific antigen is not necessary for their activation. However, the nature of the NK cell surface receptor that allows these cells to identify their targets remains largely unknown. The mechanism of NK cytotoxicity is similar to that of cytotoxic T cells. These cells cause the direct lysis of virus-infected cells and tumor cells by the production and release of pore-forming proteins.

18.9 Complement system

The fourth form of innate immune response involves the *complement system*. This family of proteins is typically found circulating in the blood in their inactive forms. As with fibrinogen and the clotting factors, complement proteins leak into the tissue spaces during an inflammatory response.

The complement system may be activated by three pathways:

1. Classical pathway
2. Alternative pathway
3. Lectin pathway

The *classical pathway* is elicited when complement protein C1 binds to an antibody. Activation of C1 also leads to the activation of the other complement proteins in the system. This pathway is more rapid and efficient than the alternative and lectin pathways. The *alternative pathway* is triggered when complement protein C3b binds with proteins or polysaccharides on the surface of a microbe. This binding and activation of C3b leads to the activation of other complement proteins. Finally, the *lectin pathway* is triggered when plasma mannose-binding lectin (MBL) attaches to a microbe. MBL is structurally similar to C1 and, therefore, serves to activate the complement system. It is important to note that the alternative and lectin pathways are components of the innate immune system, whereas the classical pathway is involved in the adaptive immune system. Regardless, the effects of complement activation are the same. The overall effects include *direct lysis of the invading microbe* and *enhancement of the inflammatory response*.

Complement-mediated cytolysis of microbes involves the formation of the *membrane attack complex* (MAC). Binding of C3b to the microbe leads to the activation of complement factors C5 through C9 in a stepwise fashion. These factors then aggregate and embed in the membrane of the microbe forming a pore. The ensuing osmotic flux of water into the cell results in cell lysis.

Complement augments several steps in the inflammatory response including the following:

- Enhanced release of histamine
- Vasodilation and increased permeability
- Chemotaxis
- Opsonization

Complement factors C3a and C5a stimulate the release of histamine from tissue mast cells. The histamine then promotes the vascular changes resulting in increased delivery of phagocytes to the tissue. Complement factors C3a and C5a also serve as chemotaxins drawing the phagocytes through the tissue to the site of injury. Finally, as described previously, C3b serves as an opsonin and enhances phagocytosis.

18.10 Adaptive immune system

Adaptive immunity is also referred to as *specific* or *acquired* immunity. Adaptive immune responses develop more slowly than the innate immune responses; however, they are much more powerful and effective at eliminating infection. In addition, these responses are highly specific in that they are directed toward a particular microbial invader. As such, the immune system requires prior exposure to the infectious agent in order to elicit these responses.

There are two types of lymphocytes: B cells and T cells. Each type is derived from stem cells in the bone marrow. However, B cells differentiate and mature in the bone marrow, and T cells differentiate and mature in the thymus, the lymphoid tissue located in the midline of the thoracic cavity, above the heart and between the lungs. Mature B cells and T cells are then released into the blood. Some of these cells remain in the blood or enter the lymph and body tissues. In this way, they conduct immune surveillance and are on the constant search for invading microbes and tumor cells. However, most of the lymphocytes (approximately 2 trillion) establish colonies in peripheral lymphoid tissues, such as the lymph nodes, spleen, adenoids, tonsils, and appendix. Regardless of location, exposure to a specific microbial protein, or antigen, stimulates cell division and the production of new generations of a particular type of lymphocyte.

Adaptive immune responses are triggered by *antigens*. These molecules are large (mol. wt. > 10,000), complex, and unique. Typically, antigens are foreign proteins and may include microbial cell surface receptors and bacterial toxins. Large polysaccharides may also be antigenic in nature. B cells

and T cells have surface receptors that recognize a specific antigen. Binding of an antigen to its receptor elicits a specific immune response. There are two types of immune responses: antibody-mediated immunity and cell-mediated immunity.

18.11 Antibody-mediated immunity

Also referred to as humoral immunity, antibody-mediated immunity involves the activation of B lymphocytes. These responses begin when an antigen binds to its specific B cell surface receptor. Under the influence of cytokines released from *helper T cells*, the B cell matures into a *plasma cell*. The plasma cell produces the antibodies. These cells have an abundant rough endoplasmic reticulum (the site of protein synthesis) and can produce as many as 2000 antibodies per second. These antibodies eventually gain access to the blood and are referred to as *gamma globulins* or *immunoglobulins*. Under normal conditions, antibodies account for approximately 20% of all plasma proteins.

18.11.1 Classification of antibodies

There are five classes of antibodies:

1. IgM: Serves as the B cell surface receptor for antigen binding; is associated with the primary immune response; acts as an efficient activator of the complement system.
2. IgG: Circulates as the predominant (75%) antibody in the serum; is produced in secondary immune responses; inactivates pathogens (e.g., opsonization) and activates the complement system; crosses the placenta and provides neonatal immunity.
3. IgA: Protects sites of potential microbial invasion from infection due to its presence in the secretions and the mucus membranes of the digestive, respiratory, reproductive, and urinary systems; is found in saliva and tears; provides neonatal immunity due to its presence in maternal milk.
4. IgE: Serves as the mediator for acute inflammatory responses and allergic responses due to its presence on the surface of mast cells; provides protection against parasitic worms.
5. IgD: Serves as the B cell surface receptor for antigen binding; circulates in trace amounts; physiological activities unclear.

18.11.2 Structure of antibodies

Antibodies are composed of four interlinked polypeptide chains: *two long, heavy chains* and *two short, light chains*. These chains are arranged so that the antibody molecule is in the shape of a "Y" (Figure 18.2). The arm regions contain the *antigen-binding fragments* (Fab). These regions are unique for each

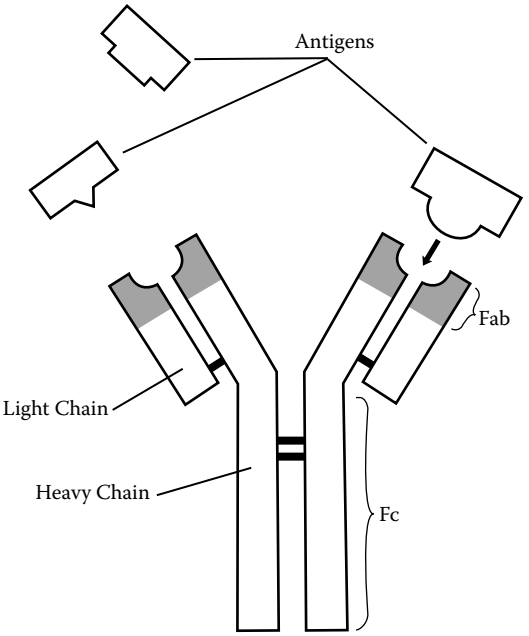


Figure 18.2 Structure of antibodies. The antibody molecule is shaped like a "Y." The antigen-binding fragments (Fab) bind specifically with a particular antigen. The constant region (Fc) determines the type of protective action to be carried out by that antibody.

different antigen. As a result, these fragments determine antibody specificity. The tail portion of the antibody molecule is the *constant region* (Fc). This region determines the type of protective action carried out by the antibody. For example, the Fc region of an IgG antibody binds to the surface of phagocytes. As a result, the IgG serves as an opsonin. The Fc region of an IgE antibody binds to the surface of mast cells. Following antigen binding to the IgE antibody, the mast cell degranulates and releases histamine and other mediators of allergic and inflammatory reactions. The Fc region of an IgM antibody binds with complement protein C1 and leads to the activation of the complement system.

18.11.3 Actions of antibodies

Antibodies exert their protective effects by several different mechanisms (see Figure 18.3):

- Neutralization
- Agglutination
- Activation of the complement system
- Opsonization

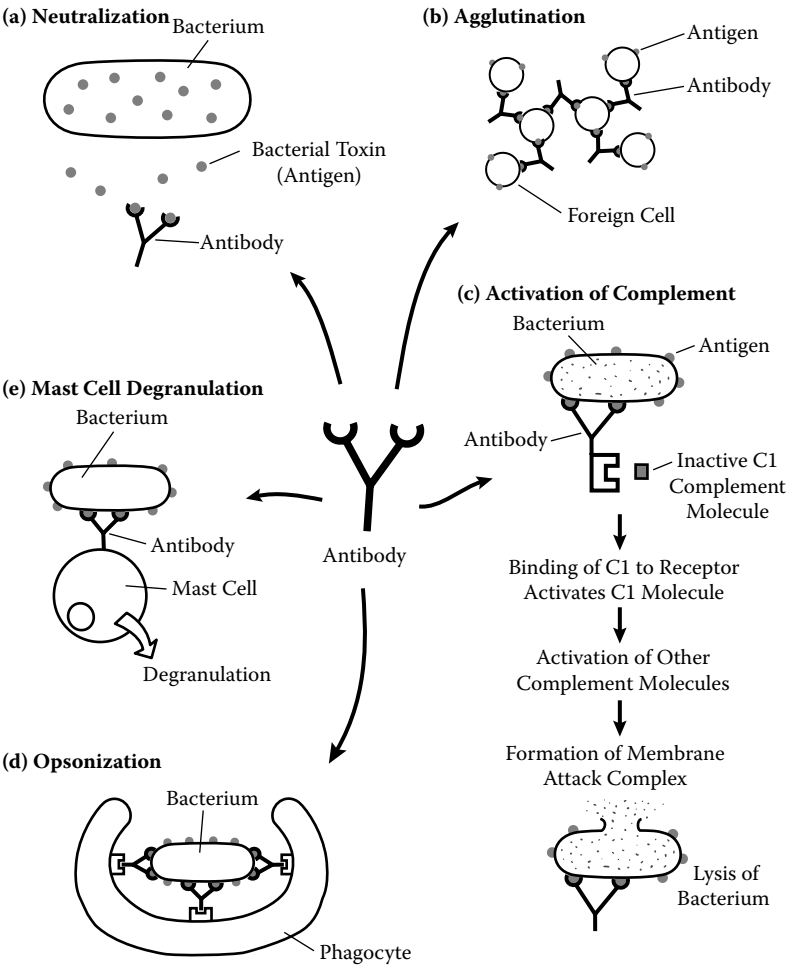


Figure 18.3 Actions of antibodies. Antibodies may physically hinder antigens by way of neutralization (a) or agglutination (b). Antibodies may also amplify innate immune responses including the activation of the complement system (c), opsonization (d), and mast cell degranulation (e).

- Mast cell degranulation
- Stimulation of natural killer cells

Neutralization and *agglutination* are forms of *physical hindrance* of antigens which, in fact, plays a relatively minor role in antibody protection against antigens. Specifically, neutralization involves the binding of bacterial toxins with their specific antibodies. As a result, the toxins are unable to interact with or damage susceptible cells. Agglutination involves the clumping of foreign cells such as bacteria or mismatched red blood cells. This process

enhances the likelihood of phagocytosis. Agglutination reactions are also used clinically to identify bacteria and to type red blood cells. IgM and, to a lesser extent, IgG are effective agglutinating agents.

Activation of the complement system, *opsonization*, *mast cell degranulation*, and *stimulation of killer cells* all involve the amplification of innate immune responses. These are the most powerful protective actions of antibodies.

Antibodies, specifically IgM molecules, are the most potent activators of the complement system. As discussed previously, the complement system causes the direct lysis of bacteria via the membrane attack complex. In addition, this system enhances every aspect of the inflammatory response including chemotaxis and opsonization.

Opsonization involves the binding of IgG antibodies to the surface of phagocytes as well as to invading bacteria. The linkage between the phagocyte, the IgG, and the bacterium prevents the escape of the microbe and enhances phagocytosis.

Mast cells have IgE antibodies on their cell surfaces. The binding of specific antigens to the IgE molecules results in mast cell degranulation and the release of chemicals such as histamine which mediate the inflammatory response.

Killer cells are similar to the natural killer cells discussed previously in that they cause the lysis of cells. However, in order to carry out their effects, killer cells require the target bacterium to be coated with antibodies. The antigen-bound antibody then binds to the surface of the killer cell by way of the Fc region. As a result, the killer cell is activated to release *perforin molecules.* These molecules aggregate and insert themselves into the bacterial cell membrane that forms a pore and causes cell lysis. A summary of the innate and adaptive immune defenses against bacterial invasion can be found in Table 18.1.

Table 18.1 Defenses against Bacterial Invasion

Inflammation
• Walling off the affected area with a fibrin clot
• Phagocytosis of bacteria by neutrophils and phagocytes
Complement
• Bacterial lysis by the membrane attack complex
• Opsonization
Antibody-Mediated Immunity
• Neutralization
• Agglutination
• Activation of the complement system
• Opsonization
• Stimulation of killer cells

18.11.4 Clonal selection theory

There are literally millions of different types of antigens that may elicit an antibody-mediated immune response. However, each type of B lymphocyte with its specific IgM antibody as its cell surface receptor may respond to only one particular antigen. Therefore, millions of different types of B cells are needed to recognize each of these antigens. As such, a remarkably diverse population of B cells is produced during fetal development. Each of these B cells is capable of synthesizing antibodies against a specific antigen. All of the offspring from a given B cell are identical to the original lymphocyte and form *clones*. Furthermore, all of these cells produce the same specific antibody. Because millions of different B lymphocytes are required to produce so many different antibodies, it is clear that the number of a particular type of B cell must be quite limited. When exposed to an antigen, expansion of the clone, or an increase in the number of antibody-producing cells, is necessary to achieve an effective immune response.

The *clonal selection theory* is illustrated in Figure 18.4. Exposure to an antigen causes the selective activation of a *naïve lymphocyte*, one that has not been previously exposed to the antigen. The B cell then proliferates, forming *plasma cells* and *memory cells*.

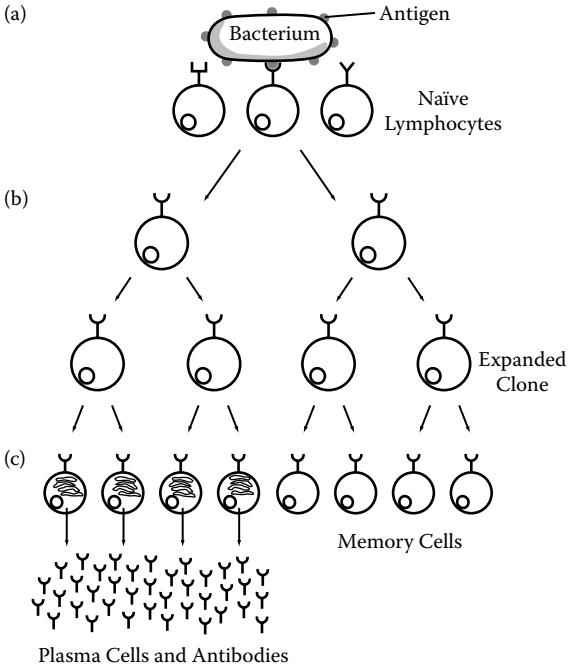


Figure 18.4 Clonal selection theory. (a) A naïve B cell is exposed to an antigen. (b) The now-activated B cell proliferates and expands the clone. (c) Some of the progeny become antibody-producing plasma cells and others become memory cells.

Most of the cellular offspring are plasma cells. These cells have abundant rough endoplasmic reticulum, the protein-producing organelle of the cell. As such, they are the actual antibody-producing cells during an immune response. Although the newly produced antibodies have the same antigen-binding sites as the IgM cell surface receptors, they are, in fact, IgG molecules. These IgG antibodies then exert their protective effects against the foreign antigen by way of opsonization or activation of the complement system.

18.11.5 Primary versus secondary responses

The initial exposure to an antigen elicits the *primary response*. Characteristics of this antibody response include the following:

- The response is delayed several days until sufficient numbers of plasma cells are formed and antibody production is prolific.
- The peak of this response is reached in 2 to 3 weeks.
- Symptoms of the particular microbial infection occur.

Subsequent exposure to the same antigen elicits the *secondary response*. This response involves the activation of the long-lived memory cells that were formed during the primary response. The presence of these memory cells expanded the clone and, therefore, increased the number of lymphocytes available to respond to the antigen. In addition, memory cells are more readily stimulated by the antigen because the cell surface receptors have a greater affinity for the antigen than did the original B lymphocyte that produced the clone. As a result, the characteristics of the secondary response are quite different from those of the primary response and include the following:

- The response is more rapid (days versus weeks), more potent (the magnitude of antibody production is 100 times greater), and longer lasting (several weeks).
- This faster, more powerful response prevents or minimizes overt infection and the development of symptoms.
- This response provides long-term immunity against a specific disease.

The primary response to antigen exposure may occur by way of actual exposure to the microbe or vaccination.

A microbe typically includes two components: the *virulent portion* and the *antigenic portion*. The virulent portion of the microbe elicits disease. The antigenic portion elicits the immune response. Vaccine development involves stripping a microbe of its disease-inducing capability while leaving its antigenic nature intact. As a result, the patient develops an immune response against this now-harmless antigen and forms the memory cells necessary for long-lasting immunity. Once again, the development of overt disease is averted upon exposure to the actual microbial antigen.

Interestingly, some microbial infections do not elicit the formation of memory cells. In these cases, exposure to the antigen does not result in long-lasting immunity. An important example of this phenomenon is streptococcal infection or "strep throat." The course of the disease and the intensity of the symptoms are the same following each exposure.

18.11.6 Active versus passive immunity

Active immunity involves the production of antibodies as a result of an exposure to an antigen. The discussion of antibody-mediated immunity thus far has described the active immune process.

Passive immunity involves the transfer of preformed antibodies from one individual (human or animal such as horse and sheep) to another. Examples of passive immunity include the following:

- Transplacental passage of IgG from mother to fetus.
- Acquisition of IgA from mother's colostrum and milk by a nursing infant.
- Injection of human polyclonal antibodies to patients with tetanus infection (antitoxin).
- Administration of horse polyclonal antibodies to patients with botulism (antitoxin).
- Injection of human polyclonal antibodies to patients following a bite from an animal possibly infected with rabies (antiviral).

The clinical administration of preformed antibodies is very effective in providing immediate protection from virulent infectious agents or lethal toxins to which a patient may have been exposed. The transfer of antibodies from mother to offspring is quite beneficial as the maturity of an infant's immune system is incomplete for several months. Passively acquired antibodies are usually broken down within 1 month.

18.12 Cell-mediated immunity

Cell-mediated immunity involves the activation of T cells. This form of immunity provides defense against microbial invaders located within the host's cells where antibodies and the complement system are unable to reach and destroy them. Specifically, T lymphocytes cause the destruction of virus-infected cells and tumor cells.

Similar to B cells, T cells are clonal and antigen specific. In addition, receptors capable of recognizing foreign antigens are found on the surface of the T cell. Unlike B cells, T cells are activated only when their receptors bind with the foreign antigen as well as the self-antigen on the host cell surface. The exception to activation by way of the foreign antigen and self-antigen complex involves the immune response to whole transplanted foreign cells.

PHARMACY APPLICATION: ANTIBIOTIC AGENTS

The body's natural defense mechanisms are not always adequate in preventing microbial infection. However, not until after World War II were effective pharmacotherapies developed. Antibacterial agents are generally referred to as antibiotics. Penicillin is one of the earliest discovered and widely used antibiotic agents. The use of penicillin began in the 1940s. Since that time, the development of new antibiotic agents has been explosive, with different classes of drugs acting at different target sites in bacteria. Mechanisms of antibiotic action include:

Inhibition of synthesis of the bacterial cell wall (e.g., penicillins, cephalosporins).
Inhibition of bacterial protein synthesis (e.g., erythromycin, tetracyclines)
Altered bacterial protein synthesis (e.g., aminoglycosides).
Interruption of nucleic acid synthesis (e.g., fluoroquinolones).
Interference with normal metabolism (e.g., sulfonamides).

The selection of an antibiotic agent may be made empirically, definitively, or prophylactically.

Of significant concern is the increasing prevalence of bacterial resistance to antibiotic agents. There are three general categories of bacterial resistance:

1. Drug does not reach its target (e.g., loss of effective porin channels decreases the rate of entry of drug into the bacterium).
2. Drug is not active (e.g., production of drug-modifying enzymes by the bacterium).
3. Target is altered (e.g., mutation of the natural target).

The responsible use of antibiotics is essential to avoid or minimize the development of resistance.

18.12.1 Types of T cells

There are three types of T lymphocytes produced by the body:

1. *Cytotoxic T cells*: Also referred to as killer T cells or CD8 cells, these lymphocytes destroy virus-infected cells, tumor cells, and transplanted cells.
2. *Helper T cells*: Also referred to as CD4 cells, these lymphocytes are more abundant and account for 60% to 80% of circulating T cells. These lymphocytes are selectively destroyed by the AIDS virus. The resulting loss of the numerous contributions of these cells to immune responses leaves the patient susceptible to infection.

3. *Suppressor T cells*: By inhibiting the activities of cytotoxic T cells and helper T cells, suppressor T cells prevent the development of an excessive immune reaction that may actually be harmful to the body.

18.12.2 Actions of T cells

Cytotoxic T cells *directly* destroy target cells. In contrast, the immune functions of the helper T cells are *indirect*. These cells modulate the activities of other effector cells of the immune system and increase the overall magnitude of the immune response:

Cytotoxic T cells: These cells secrete chemicals that destroy their target cells (Figure 18.5). The following mechanism describes how the cytotoxic T cell destroys a virus-infected cell that is its most frequent target.

As mentioned previously, viruses must invade a living host cell in order to replicate. As it enters the host cell, the virus leaves a portion of its antigenic protein coat on the cell surface in association with the host's self-antigen. (The viral DNA proceeds into the cell to carry on metabolism and reproduction.) The cytotoxic T cell with the appropriate receptor recognizes and binds to the foreign antigen and self-antigen complex. The now activated T cell releases granules containing *perforin molecules* into the extracellular fluid. These molecules aggregate and insert themselves into the membrane of the host cell forming a pore-like channel. This channel allows water and salt to enter the host cell and cause cell lysis.

The activated cytotoxic T cell may also release chemicals referred to as *granzymes*. These enzymes enter the infected cell through the perforin channels and elicit self-destruction of the cell by way of *apoptosis*.

Lysis of virus-infected cells causes the release of the viruses into the extracellular fluid. The now exposed viruses are engulfed and destroyed by macrophages.

Helper T cells: These cells secrete chemicals that amplify the activity of other effector cells of the immune system. These *cytokines* include all of the chemicals (other than antibodies) that are secreted by leukocytes. Most of the cytokines are secreted by helper T cells. These cytokines and their immune actions include the following:

B-cell growth factor: Enhances the development of antigen-stimulated B lymphocytes into antibody-secreting plasma cells (secreted by *T helper 2 cells*)

T-cell growth factor (interleukin 2, IL-2): Enhances the activity of the T lymphocytes in the activated clone; enhances the activity of appropriate B lymphocytes (secreted by *T helper 1 cells*)

Chemotaxins: Attract phagocytes to the infected area

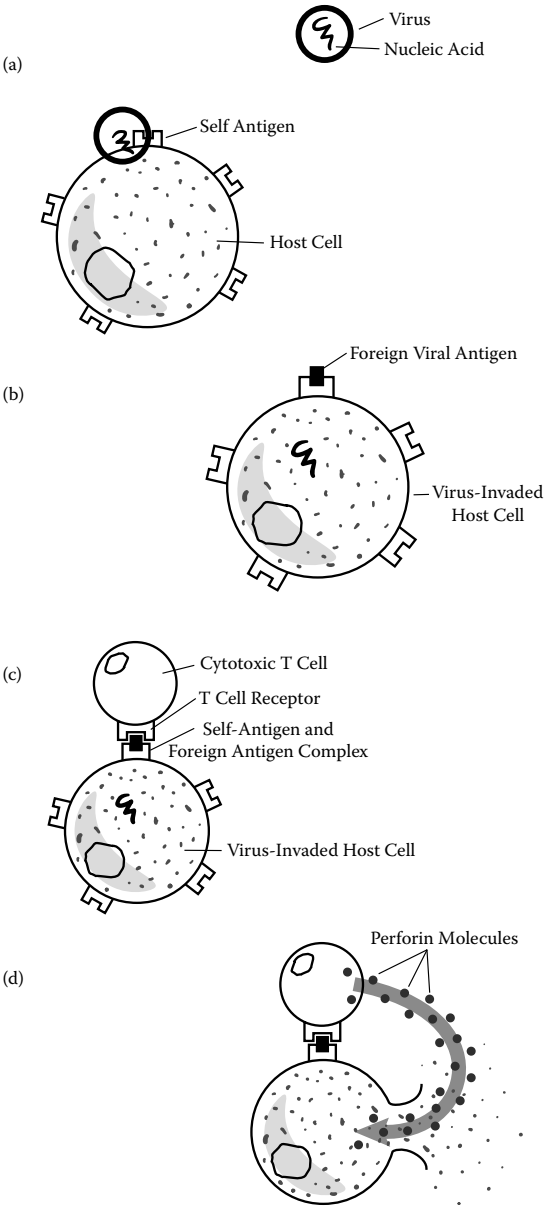


Figure 18.5 Mechanism of action of cytotoxic T cells. Lysis of virus-infected cells involves several steps including (a) viral invasion of the host cell; (b) incorporation of foreign antigen into the host cell membrane in association with the self-antigen; (c) binding of the specific cytotoxic T cell to the self-antigen and foreign antigen complex; and (d) release of perforin molecules that form channels in the host cell membrane, allow for the influx of water and salt into the cell, and cause cell lysis. Viruses released from the lysed cell are then removed and destroyed by macrophages.

Table 18.2 Chemicals Active during Immune Responses

Antibodies: Gamma globulins, immunoglobulins; secreted by plasma cells to defend against invading bacteria.
B-Cell Growth Factor: Enhances the development of antigen-stimulated B cells into plasma cells.
Chemotaxins: Molecules that attract phagocytes to an area of infection or inflammation.
Complement: Family of proteins that cause lysis of invading bacteria and enhance the inflammatory response (e.g., chemotaxis, opsonization).
Cytokines: Proteins released by one type of immune cell that influence the growth and activity of other immune cells.
Granzymes: Cytotoxic enzymes that initiate apoptosis in virus-infected cells.
Histamine: Released from mast cells and basophils; initiates the inflammatory response by causing vasodilation and increased capillary permeability.
Interferon: Protein released from virus-infected cells; inhibits replication of viruses in other host cells.
Macrophage-Migration Inhibition Factor: Limits the outward migration of macrophages; enhances the phagocytic activity of macrophages.
Membrane Attack Complex: Composed of complement proteins C5–C9; causes lysis of invading bacteria.
Perforin: released from cytotoxic T cells and killer cells; causes lysis of virus-infected cells and tumor cells.
T-Cell Growth Factor: Enhances the activity of T cells in an activated clone; enhances the activity of select B cells.

Macrophage-migration inhibition factor: Causes the accumulation of macrophages in the infected area by limiting their outward migration; enhances the phagocytic activity of macrophages

These chemicals and others that are active during immune responses are summarized in Table 18.2. A summary of the innate and adaptive immune defenses against viral invasion can be found in Table 18.3.

18.12.3 MHC molecules

Self-antigens are also referred to as *MHC molecules* because their synthesis occurs by way of a group of genes called the *major histocompatibility complex*. Each individual has their own unique pattern of MHC molecules or cell markers. The presence of these glycoproteins on the surface of the body's cells identifies the cells as "self," and accordingly, T lymphocytes do not respond to them.

There are two classes of MHC molecules: *class I MHC glycoproteins* and *class II MHC glycoproteins*.

Table 18.3 Defenses against Viral Invasion

Macrophages
• Phagocytosis of viruses released into extracellular fluid
Interferon
• Prevention of viral replication in host cells
• Increase in macrophage activity
• Increase in activity of natural killer cells
• Increase in activity of cytotoxic T cells
Natural Killer Cells
• Lysis of virus-infected cells
Cytotoxic T Cells
• Lysis of virus-infected cells
Helper T Cells
• Release of T-cell growth factor to enhance the activity of cytotoxic T cells
• Release of macrophage-migration inhibition factor to increase the numbers of macrophages in the affected area

Class I MHC glycoproteins are found on the surface of all nucleated cells in the body. This includes virus-infected cells and tumor cells. These molecules enable *cytotoxic T cells* to respond to foreign antigens remaining on the surface of the infected cell. It is quite logical that cytotoxic T cells do not respond to foreign antigens in the absence of MHC molecules. As described previously, these T cells defend against foreign antigen only when it has become incorporated into the infected cell's membrane and not when it is in the free form.

Class II MHC glycoproteins are found on the surface of select immune cells including macrophages, B cells, and cytotoxic T cells. These molecules are recognized by *helper T cells*. As such, helper T cells respond to foreign antigens only when they are present on the surface of the immune cells with which they interact.

In summary, class I MHC glycoproteins facilitate the interaction of cytotoxic T cells with virus-infected cells, and class II MHC glycoproteins facilitate the interaction of helper T cells with the immune cells that they activate.

Medical terminology

Agglutination (ă-gloo″tĭ-nā′shŭn): Antigen–antibody reaction in which cells or particles are removed from solution.
Antigenic (ăn-tĭ-jĕn′ĭk): Capable of inducing a specific immune response.
Apoptosis (ă-pŏp-tō′sĭs): Programmed cell death.
Calor (kā′lor): Heat.

Chemotaxis (kē″mō-tăk′sĭs): Movement of phagocytes toward an area of inflammation.

Clone (klōn): Group of lymphocytes all capable of recognizing and responding to a specific foreign antigen.

Cytokine (sī′tō-kīn): Protein produced by an immune cell that regulates the immune response.

Diapedesis (dī″ă-pĕd-ē′sĭs): Movement of phagocytes out of the capillary into the tissue during an inflammatory response.

Dolor (dō′lor): Pain.

Edema (ĕ-dē′mă): Accumulation of fluid in a tissue.

Endoplasmic reticulum (ĕn-dō-plăs′mĭk rĕ-tĭk′ū-lŭm): Network of tubules within the cytoplasm where the synthesis and modification of molecules takes place.

Functio laesa (fŭnk′shē-ō lē′să): Impaired function.

Hypersensitivity (hī″pĕr-sĕn″sĭ-tĭv′ĭ-tē): Increased responsiveness to a stimulus.

Immunity (ĭ-mū′nĭ-tē): Protection from disease, especially infectious disease.

Immunoglobulin (ĭm″ū-nō-glŏb′ū-lĭn): Antibody.

Innate (ĭn-nāt′): Inherent, intrinsic, existing at birth.

Interferon (ĭn-tĕr-fēr′-ŏn): Glycoprotein released from virus-infected cells with antiviral and antitumoral effects.

Lysosome (lī′sō-sōm): Cellular organelle containing digestive enzymes.

Margination (măr″jĭ-nā′shŭn): Adhesion of phagocytes to the wall of the capillary during an inflammatory response.

Naïve (nă-ēv′): Natural, unlearned, unaffected.

Neutralization (nū″trăl-ĭ-zā′shŭn): Counteracting the effects of an agent that produces morbidity or disease.

Opsonization (ŏp″sō-nī-zā′shŭn): Facilitation of phagocytosis.

Perforin (pĕr′fōr-ĭn): Protein released by natural killer cells and cytotoxic T cells that causes the lysis of virus-infected cells.

Phagocytosis (făg″ō-sī-tō′sĭs): Process where phagocytes engulf and destroy microorganisms, particles, or cellular debris.

Polyclonal (pŏl″ē-klōn′ăl): Referring to proteins arising from multiple cell lines.

Prolific (prō-lĭf-ĭk): Fruitful, productive.

Pus (pŭs): Fluid containing dead phagocytes and cellular debris.

Rubor (roo′bor): Redness.

Transcription (trăn-skrĭp′shŭn): Synthesis of messenger RNA.

Translation (trăns-lā′shŭn): Synthesis of proteins.

Transplacental (trăns″plă-sĕn′tă): Across the placenta.

Tumor (tū′mor): Swelling.

Virulent (vĭr′ū-lĕnt): Infectious, pathogenic.

Bibliography

Abbas, A. K., and Lichtman, A. H., *Basic Immunology, Functions and Disorders of the Immune System,* 2nd ed., updated edition 2006–2007, Saunders/Elsevier, Philadelphia, PA, 2006–2007.

Chambers, H. F., Antimicrobial Agents: General Considerations, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 9th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 43.

Delves, P. J., Martin, S. J., Burton, D. R., and Roitt, I. M., *Roitt's Essential Immunology,* 11th ed., Blackwell, Malden, MA, 2006.

Dunne, Jr., W. M., Mechanisms of Infectious Disease, in *Pathophysiology, Concepts of Altered Health States,* 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 18.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Porth, C. M., Alterations of the Immune Response, in *Pathophysiology, Concepts of Altered Health States,* 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 21.

Roberts, L. J., and Morrow, J. D., Analgesic-Antipyretic and Antiinflammatory Agents and Drugs Employed in the Treatment of Gout, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 9th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 27.

Sherwood, L., *Human Physiology, From Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA. 2004.

Silverthorn, D. U., *Human Physiology: An Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Sommer, C., The Immune Response, in *Pathophysiology, Concepts of Altered Health States,* 7th ed., Porth, C. M., Ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005, chap. 19.

Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

www.fda.gov/CDER/Drug/advisory/nsaids.htm.

chapter nineteen

The respiratory system

Study objectives

- Describe the blood–gas interface and explain why the lungs are ideally suited for gas exchange
- List the components and the functions of the conducting airways
- Distinguish between the various types of airways in terms of epithelium and cartilage
- Describe the forces and factors responsible for maintaining inflation of the lungs
- Explain how inspiration and expiration take place
- Distinguish between atmospheric pressure, alveolar pressure, intrapleural pressure, and transpulmonary pressure
- Define *pulmonary compliance*
- Describe the role of elastic connective tissue in the elastic recoil of the lungs as well as in lung compliance
- Explain how surface tension affects the elastic behavior of the lungs
- Describe the functions of pulmonary surfactant
- Explain how interdependence promotes alveolar stability
- Describe the factors that determine airway resistance
- Define *tidal volume*, *residual volume*, *expiratory reserve volume*, and *inspiratory reserve volume*
- Define *functional residual capacity*, *inspiratory capacity*, *total lung capacity*, and *vital capacity*
- Distinguish between total ventilation and alveolar ventilation
- Distinguish between anatomical dead space, alveolar dead space, and physiological dead space
- Explain how each factor in Fick's Law of Diffusion influences gas exchange
- List the partial pressures of oxygen and carbon dioxide in the various regions of the respiratory and cardiovascular systems
- Explain how the PO_2 and the PCO_2 of the alveolar gas are determined
- Explain the effects of airway obstruction and obstructed blood flow on ventilation–perfusion (V/Q) matching
- Describe the local control mechanisms that restore the V/Q ratio to one
- Explain how oxygen is transported in the blood
- Describe the physiological significance of the steep portion and the plateau portion of the oxyhemoglobin dissociation curve

- Describe the effects of carbon dioxide, pH, temperature, 2,3-bisphosphoglycerate, anemia, and carbon monoxide poisoning on the transport of oxygen
- Explain how carbon dioxide is transported in the blood
- Compare and contrast the functions of the dorsal and the ventral respiratory groups in the medullary respiratory center
- List and describe the sources of input to the medullary respiratory center
- Compare and contrast the function of the peripheral and the central chemoreceptors
- Describe the ventilation response to exercise

19.1 Introduction

The cells of the body require a continuous supply of oxygen to produce energy and carry out their metabolic functions. Furthermore, these aerobic metabolic processes produce carbon dioxide that must be continuously eliminated. Adequate removal of carbon dioxide is important in the regulation of acid–base balance, or concentration of H^+ ion in the blood. Therefore, the primary functions of the respiratory system include obtaining oxygen from the external environment and supplying it to the body's cells and eliminating from the body the carbon dioxide produced by cellular metabolism. The process by which oxygen is taken up by the lungs and carbon dioxide is eliminated from the lungs is referred to as *gas exchange*.

19.2 Blood–gas interface

Gas exchange takes place at the *blood–gas interface*. This interface exists where the alveoli and the pulmonary capillaries come together. The *alveoli* are the smallest airways in the lungs. The *pulmonary capillaries* are found in the walls of the alveoli. Inspired oxygen moves from the alveoli into the capillaries for eventual transport to the tissues. Carbon dioxide, entering the lungs by way of the pulmonary circulation, moves from the capillaries into the alveoli for elimination by expiration. Both oxygen and carbon dioxide move across the blood–gas interface by way of *simple diffusion* — from an area of high concentration to an area of low concentration.

According to *Fick's Law of Diffusion*, the amount of gas that moves across the blood–gas interface is proportional to the surface area of the interface and inversely proportional to the thickness of the interface. In other words, gas exchange in the lungs is promoted when the *surface area* for diffusion is maximized and the *thickness of the barrier* to diffusion is minimized. In fact, anatomically, the lungs are ideally suited for the function of gas exchange. There are 300 million alveoli in the lungs. Furthermore, the walls of each alveolus are completely lined with capillaries. There are as many as 280 billion pulmonary capillaries or almost 1000 capillaries per alveolus.

This results in a vast surface area for gas exchange of approximately 70 m^2, which is roughly the size of a tennis court.

More specifically, the blood–gas interface consists of the alveolar epithelium, the capillary endothelium, and the interstitium. The alveolar wall is made up of a single layer of flattened *alveolar type I cells*. The capillaries surrounding the alveoli also consist of a single layer of cells, the *endothelial cells*. In between the alveolar epithelium and the capillary endothelium is a very small amount of *interstitium*. In some regions, the interstitium may be essentially absent. In this case, the basement membranes of the alveolar epithelial cell and the capillary endothelial cell fuse together. Taken all together, only 0.2 to 0.5 µm separates the air in the alveoli from the blood in the capillaries. The extreme thinness of the blood–gas interface further facilitates gas exchange by way of diffusion.

19.3 Airways

The airways of the lungs consist of a series of branching tubes. Each level of branching results in another generation of airways. As they branch, the airways become narrower, shorter, and more numerous. There are a total of 23 *generations of airways* with the alveoli composing the 23rd generation.

Air is carried to and from the lungs by the *trachea*, which extends toward the lungs from the larynx. The trachea divides into the *right* and *left main bronchi*. These *primary bronchi* each supply a lung. The primary bronchi branch and form the *secondary*, or *lobar*, *bronchi* — one for each lobe of lung. The left lung consists of two lobes, and the right lung has three lobes. The lobar bronchi branch and form the *tertiary*, or *segmental*, *bronchi* — one for each of the functional segments within the lobes. These bronchi continue to branch and move outward toward the periphery of the lungs. The smallest airways without alveoli are the *terminal bronchioles*. Taken all together, the airways from the trachea through and including the terminal bronchioles are referred to as the *conducting airways*. This region, which consists of the first 16 generations of airways, contains no alveoli. Therefore, there is no gas exchange in this area. Consequently, it is also referred to as *anatomical dead space*. The volume of the anatomical dead space is approximately 150 ml (or about 1 ml per pound of ideal body weight).

The conducting airways carry out three major functions:

1. They lead inspired air to the more distal gas-exchanging regions of the lungs.
2. They warm and humidify the inspired air as it flows through them.
3. They defend against microbes, toxic chemicals, and foreign matter.

The large airways provide a low-resistance pathway for airflow toward the respiratory zone. In other words, they function largely as conduits for the bulk flow of air.

Second, the alveoli are delicate structures and may be damaged by excessive exposure to cold, dry air. The addition of warmth and moisture, particularly in the winter months, protects the alveoli. In fact, this water vapor is visible when one exhales on a cold winter day. The defense against inhaled microbes or foreign particles is discussed in Section 19.3.1. In addition, vagally induced bronchoconstriction in response to irritant receptor stimulation is discussed in Section 19.8.3.

Branching from the terminal bronchioles are the *respiratory bronchioles*. This is the first generation of airways to have alveoli in their walls. Finally, there are the *alveolar ducts* that are completely lined with *alveolar sacs*. This region, from the respiratory bronchioles through the alveoli, is referred to as the *respiratory zone*. This zone makes up most of the lungs and has a volume of about 3000 ml at the end of a normal expiration.

19.3.1 Epithelium

All of the conducting airways, the trachea through the terminal bronchioles, are lined with *pseudostratified ciliated columnar epithelium*. There are approximately 300 cilia per epithelial cell. Interspersed among these epithelial cells are mucus-secreting *goblet cells*. Furthermore, *mucus glands* are found in the larger airways. Consequently, the surface of the conducting airways consists of a mucus-covered ciliated epithelium. The cilia beat upward at frequencies between 600 and 900 beats per minute. As a result, the cilia continuously move the mucus away from the respiratory zone and up toward the pharynx at a rate of 1 to 2 cm/min. This *mucociliary escalator* provides an important protective mechanism that removes inhaled particles from the lungs. As a result of this process, particles larger than about 6 μm do not normally reach the respiratory zone. Mucus that reaches the pharynx is usually swallowed or expectorated. An additional mechanism by which airway mucus protects the lungs involves the presence of immunoglobulins. These substances, also referred to as antibodies, destroy or neutralize inhaled pathogens. The activity of the immunoglobulins was discussed in detail in Chapter 18. Interestingly, the nicotine found in cigarette smoke paralyzes the cilia and impairs their ability to remove the many toxic substances found in smoke.

The respiratory bronchioles are lined with *cuboidal epithelial cells* that gradually flatten and become squamous-type cells. As mentioned previously, the alveoli are composed of large, flat, *simple squamous epithelium*.

19.3.2 Cartilage

The trachea and the primary bronchi contain *C-shaped hyaline cartilage rings* in their walls. The lobar bronchi contain *plates of cartilage* that completely encircle the airways. The cartilage in these large airways provides structural support and prevents the collapse of the airways. As the bronchi continue

to branch and move out toward the lung periphery, the cartilage diminishes progressively until it disappears in airways about 1 mm in diameter. Airways with no cartilage are referred to as *bronchioles*. As the cartilage becomes sparser, it is replaced by *smooth muscle*. Therefore, the bronchioles, which have no cartilage to support them but do have smooth muscle capable of vigorous constriction, are susceptible to collapse under certain conditions, such as an asthmatic attack.

19.4 The pleura

Each lung is enclosed in a double-walled sac referred to as the *pleura*. The *visceral pleura* is the membrane adhered to the external surface of the lungs. The *parietal pleura* lines the walls of the thoracic cavity. The space in between the two layers, the *pleural space*, is very thin and completely closed.

The pleural space is filled with *pleural fluid*. This fluid lubricates the membranes and reduces friction between the layers as they slide past each other during breathing. The pleural fluid also plays a role in maintaining lung inflation. The surface tension between the molecules of the pleural fluid keeps the two layers of the pleura "adhered" to each other. This concept is similar to the effect of water between two glass microscope slides. The two pieces of glass can easily slide over each other; however, the strong attraction of the water molecules for each other opposes the separation of the slides. In this way, the lungs are in contact with the thoracic wall, fill the thoracic cavity, and remain inflated. In other words, the surface tension in the pleural space opposes the tendency of the lungs to collapse.

19.5 Mechanics of breathing

The mechanics of breathing involve the volume changes and pressure changes that take place during ventilation that allow air to move in and out of the lungs. Air will move from an area of high pressure to an area of low pressure. Therefore, a pressure gradient between the atmosphere and the alveoli must be developed. This section will explain how changes in thoracic volume, lung volume, and pulmonary pressures occur in order to cause the pressure gradients responsible for inspiration and expiration.

19.5.1 Thoracic volume

The volume of the thoracic cavity increases during inspiration and decreases during expiration.

19.5.2 Inspiration

The most important muscle of inspiration is the *diaphragm*. The diaphragm is a thin, dome-shaped muscle that inserts into the lower ribs. A skeletal

muscle, it is supplied by the *phrenic nerves*. When the diaphragm contracts, it flattens and pushes downward against the contents of the abdomen. Therefore, contraction of the diaphragm causes an increase in the vertical dimension (top to bottom) of the thoracic cavity and an increase in thoracic volume. In fact, the diaphragm is responsible for 75% of the enlargement of the thoracic cavity during normal, quiet breathing.

Assisting the diaphragm with inspiration are the *external intercostal muscles*. These muscles connect adjacent ribs. When the external intercostal muscles contract, the ribs are lifted upward and outward (much like a handle on a bucket). Therefore, contraction of these muscles causes an increase in the horizontal dimension (front to back) of the thoracic cavity and a further increase in thoracic volume. The external intercostal muscles are supplied by the *intercostal nerves*.

Deeper inspirations are achieved by more forceful contraction of the diaphragm and the external intercostal muscles. Furthermore, *accessory inspiratory muscles*, including the scalenes, the pectoralis minor, and the sternocleidomastoid muscles, contribute to this process. Located mainly in the neck, the scalenes and the sternocleidomastoid muscles raise the sternum and elevate the first two ribs. As a result, the upper portion of the thoracic cavity is enlarged. The pectoralis minor muscles, which are located in the upper chest, also pull the rib cage superiorly.

19.5.3 Expiration

Expiration during normal, quiet breathing is *passive*. In other words, no active muscle contraction is required. When the diaphragm is no longer stimulated by the phrenic nerves to contract, it passively returns to its original, preinspiration position under the ribs. Relaxation of the external intercostal muscles allows the rib cage to fall inward and downward, largely due to gravity. As a result, these movements cause a decrease in thoracic volume.

During exercise or voluntary hyperventilation, expiration becomes an *active* process. Under these conditions, a larger volume of air must be exhaled more rapidly. Therefore, two muscle groups are recruited to facilitate this process. The most important muscles of expiration are the *muscles of the abdominal wall*. Contraction of these muscles pushes inward on the abdominal contents and increases abdominal pressure. As a result, the diaphragm is pushed upward more rapidly and more forcefully toward its preinspiration position.

Assisting the muscles of the abdominal wall are the *internal intercostal muscles*. These muscles are also found between the ribs; however, they are oriented in a direction opposite to that of the external intercostal muscles. Contraction of these muscles pulls the ribs inward and downward.

19.5.4 Lung volume

There are no real physical attachments between the lungs and the thoracic wall. Instead, the lungs literally float in the thoracic cavity, surrounded by pleural fluid. Therefore, the question arises, how does the volume of the lungs increase when the volume of the thoracic cavity increases? The mechanism involves the pleural fluid and the surface tension between the molecules of this fluid. As mentioned previously, the surface tension of the pleural fluid keeps the parietal pleura lining the thoracic cavity and the visceral pleura on the external surface of the lungs "adhered" to each other. In other words, the pleural fluid keeps the lungs in contact with the chest wall. Therefore, as the muscles of inspiration cause the chest wall to expand, increasing the thoracic volume, the lungs are pulled open as well. As a result, the lung volume also increases.

19.5.5 Pulmonary pressures

The changes in thoracic volume and lung volume cause the pressures within the airways and the pleural cavity to change. These pressure changes create the pressure gradients responsible for airflow in and out of the lungs. There are four pressures that must be considered (see Figure 19.1):

1. Atmospheric Pressure
2. Intrapleural Pressure
3. Alveolar Pressure
4. Transpulmonary Pressure

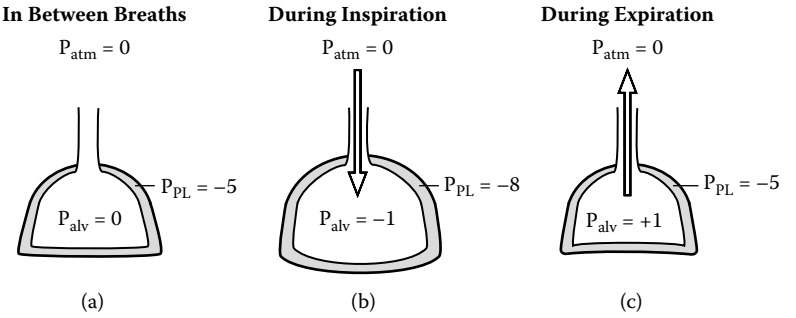


Figure 19.1 Pulmonary pressures. (a) In between breaths. Alveolar pressure (P_{alv}) is equal to atmospheric pressure (P_{atm}) which is 0 mmHg. No air flows in or out of the lungs. (b) During inspiration. As the lung volume increases, the alveolar pressure decreases and becomes subatmospheric (–1 mmHg). The pressure gradient between the atmosphere and the alveoli allows air to flow into the lungs. (c) During expiration. Following inspiration, the lungs recoil and lung volume decreases. Alveolar pressure increases and becomes greater than atmospheric pressure (1 mmHg). The pressure gradient between the atmosphere and the alveoli forces air to flow out of the lungs.

As it does with all objects on the surface of the earth, gravity exerts its effects on the molecules of the atmosphere. The weight generated by these molecules is referred to as *atmospheric*, or *barometric*, *pressure* (P_{atm}). At sea level, atmospheric pressure is 760 mmHg. In order to simplify this discussion, atmospheric pressure will be normalized to 0 mmHg and all other pressures are referenced to this.

Intrapleural pressure (P_{pl}) is the pressure within the pleural cavity. Under equilibrium conditions, the chest wall tends to pull outward and the elastic recoil of the lungs tends to pull them inward (like a collapsing balloon). These opposing forces create a subatmospheric or negative pressure within the pleural space. In between breaths, intrapleural pressure is –5 mmHg. During inspiration, the lungs follow the chest wall as it expands. However, the lung tissue resists being stretched so that the intrapleural pressure becomes even more negative and is –8 mmHg.

Alveolar pressure (P_{alv}) is the pressure within the alveoli. In between breaths, it is equal to 0 mmHg. Because there is no pressure gradient between the atmosphere and the alveoli, there is no airflow. However, in order for air to flow into the lungs, the alveolar pressure must fall below atmospheric pressure. In other words, alveolar pressure becomes slightly negative. According to Boyle's Law, at a constant temperature, the volume of a gas and its pressure are inversely related:

$$P \propto 1/V$$

Therefore, as lung volume increases during inspiration, the pressure within the alveoli decreases. Atmospheric pressure is now greater than alveolar pressure and air enters the lungs. Because the lungs are normally very compliant, or distensible, only a small pressure gradient is necessary for air to flow into the lungs. During a normal inspiration under resting conditions, alveolar pressure is –1 mmHg.

During expiration the opposite occurs. Lung volume decreases, and therefore, pressure within the alveoli increases. Alveolar pressure is now greater than atmospheric pressure and air flows out of the lungs. Alveolar pressure during a normal expiration under resting conditions is 1 mmHg.

Transpulmonary pressure (P_{tp}) is the pressure difference between the inside and the outside of the lungs. In other words, it is the pressure difference between the alveoli and the pleural space:

$$\begin{aligned} P_{tp} &= P_{alv} - P_{pl} \\ &= 0 \text{ mmHg} - (-5 \text{ mmHg}) \\ &= 5 \text{ mmHg} \end{aligned}$$

In between breaths, the transpulmonary pressure is 5 mmHg. The transpulmonary pressure is also referred to as the *expanding pressure* of the lungs. A force of 5 mmHg expands, or pushes outward on, the lungs so that they fill the thoracic cavity. As might be expected, during inspiration, the transpulmonary pressure increases, causing greater expansion of the lungs:

$$P_{tp} = (-1 \text{ mmHg}) - (-8 \text{ mmHg})$$

$$= 7 \text{ mmHg}$$

The entry of air into the pleural cavity is referred to as a *pneumothorax*. This may occur spontaneously when a "leak" develops on the surface of the lung allowing air to escape from the airways into the pleural space. It may also result from a physical trauma that causes penetration of the chest wall so that air enters the pleural space from the atmosphere. In either case, the pleural cavity is no longer a closed space, and the pressure within it equilibrates with the atmospheric pressure (0 mmHg). As a result, the transpulmonary pressure is also equal to 0 mmHg, and the lung collapses.

19.6 Elastic behavior of the lungs

In a healthy individual, the lungs are very distensible. In other words, the lungs can be inflated with minimal effort. Furthermore, during normal, quiet breathing, expiration is passive. The lungs inherently recoil to their preinspiratory position. These processes are attributed to the *elastic behavior* of the lungs. The elasticity of the lungs involves the following two interrelated properties: *elastic recoil* and *pulmonary compliance*.

The *elastic recoil* of the lungs refers to their ability to return to their original configuration following inspiration. It may also be used to describe the tendency of the lungs to oppose inflation. Conversely, *pulmonary compliance* describes how easily the lungs inflate. Compliance is defined as the change in lung volume divided by the change in transpulmonary pressure:

$$C = \frac{\Delta V}{\Delta P}$$

A highly compliant lung requires only a small change in pressure for a given degree of inflation. A less compliant lung requires a larger change in pressure for the same degree of inflation. For example, during normal, quiet breathing, human adults inhale a tidal volume of about 500 ml per breath. In an individual with healthy, compliant lungs, the transpulmonary pressure gradient needed to be generated by the inspiratory muscles is very small (approximately 2 to 3 mmHg). The patient with less compliant, or "stiff,"

lungs must generate a larger transpulmonary pressure to inflate the lungs with the same 500 ml of air. In other words, more vigorous contraction of the inspiratory muscles is required. Therefore, the *work of breathing* is increased.

The elastic behavior of the lungs is determined by two factors: the *elastic connective tissue in the lungs* and *alveolar surface tension*. The *elastic connective tissue* in the lungs consists of *elastin* and *collagen* fibers. These fibers are found in the alveolar walls and around the blood vessels and bronchi. When the lungs are inflated, the connective tissue fibers are stretched, or distorted. As a result, they have a tendency to return to their original shape and cause the elastic recoil of the lungs following inspiration. However, due to the interwoven, mesh-like arrangement of these fibers, the lungs remain very compliant and readily distensible.

The alveoli are lined with fluid. At an air–water interface, the water molecules are much more strongly attracted to each other than to the air at their surface. This attraction produces a force at the surface of the fluid referred to as *surface tension* (ST).

Alveolar surface tension exerts two effects on the elastic behavior of the lungs. First, it decreases the compliance of the lungs. For example, inflation of the lung would increase its surface area and pull the water molecules lining the alveolus apart from each other. However, the attraction between these water molecules, or the surface tension, resists this expansion of the alveolus. Opposition to expansion causes a decrease in compliance. In other words, the alveolus is more difficult to inflate and the work of breathing is increased. The greater the surface tension, the less compliant are the lungs.

The second effect of surface tension is that it causes the alveolus to become as small as possible. As the water molecules pull toward each other, the alveolus forms a sphere, which is the smallest surface area for a given volume. This generates a pressure directed inward on the alveolus, or a *collapsing pressure*. The magnitude of this pressure is determined by the *Law of LaPlace*:

$$P = \frac{2ST}{r}$$

The collapsing pressure (P) is proportional to the alveolar surface tension (ST) and inversely proportional to the radius (r) of the alveolus. In other words, the greater the surface tension and the smaller the radius, the greater is the collapsing pressure.

Due to this collapsing pressure, alveoli are inherently unstable. For example, if two alveoli of different sizes have the same surface tension, the smaller alveolus has a greater collapsing pressure and would tend to collapse and empty into the larger alveolus (see Figure 19.2a). Air flows from an area of higher pressure to an area of lower pressure. As a result, the air within the smaller alveolus flows into the larger one and an area of *atelectasis* (airway collapse) develops. Therefore, if alveolar surface tension were to remain

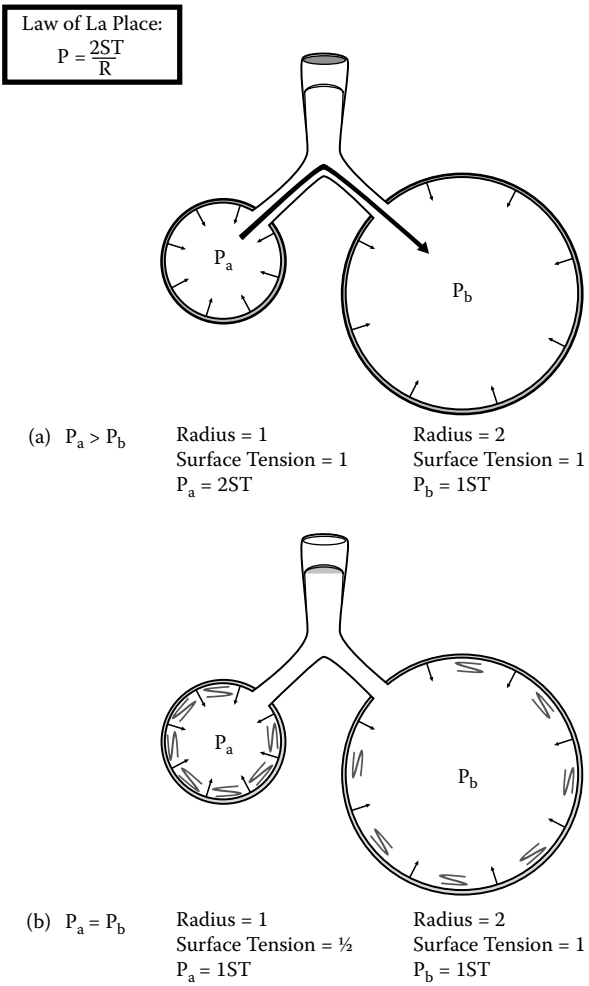


Figure 19.2 Effects of surface tension and surfactant on alveolar stability. (a) Effect of surface tension. According to the Law of LaPlace (P = 2ST/r), if two alveoli have the same surface tension (ST), the alveolus with the smaller radius (r) and, therefore, a greater collapsing pressure (P), would tend to empty into the alveolus with the larger radius. (b) Effect of surfactant. Surfactant decreases the surface tension and, therefore, the collapsing pressure, in smaller alveoli to a greater extent than it does in larger alveoli. As a result, the collapsing pressures in all alveoli are equal. This prevents alveolar collapse and promotes alveolar stability.

the same throughout the lungs, it has the potential to cause widespread alveolar collapse.

Healthy lungs, however, produce a chemical substance referred to as *pulmonary surfactant*. Made by *alveolar type II cells* within the alveoli, surfactant is a complex mixture of proteins (10% to 15%) and phospholipids (85% to 90%),

including dipalmitoyl phosphatidyl choline, the predominant constituent. By interspersing throughout the fluid lining the alveoli, it disrupts the cohesive forces between the water molecules. As a result, pulmonary surfactant has three major functions:

1. Decreases surface tension.
2. Increases alveolar stability.
3. Prevents transudation of fluid.

Pulmonary surfactant *decreases the surface tension* of the alveolar fluid. Reduced surface tension leads to a decrease in the collapsing pressure of the alveoli, an increase in pulmonary compliance (less elastic recoil), and a decrease in the work required to inflate the lungs with each breath. Second, pulmonary surfactant *promotes the stability of the alveoli*. Because the surface tension is reduced, there is a decreased tendency for small alveoli to empty into larger ones (see Figure 19.2b). Third, surfactant *inhibits the transudation of fluid* out of the pulmonary capillaries into the alveoli. Excessive surface tension would tend to reduce the hydrostatic pressure in the tissue outside of the capillaries. As a result, capillary filtration is promoted. The movement of water out of the capillaries may result in interstitial edema formation and excess fluid in the alveoli.

Surfactant is more concentrated in the smaller alveoli. Therefore, its effect of lowering surface tension is greater in those alveoli compared to the larger alveoli. In this way, the collapsing pressure is equalized among alveoli of different sizes.

The amount of surfactant produced by alveolar type II cells decreases when breaths are small and constant. On the other hand, the amount of surfactant produced is increased in response to deep breaths due to stretching of the alveolar type II cells. Patients who have had thoracic or abdominal surgery often take small, shallow breaths because of the pain. The patients must be encouraged to regularly take deep breaths. To facilitate this, patients are often provided with an *incentive spirometer.* With a mouthpiece and a piston that rises during inspiration, the patients can visualize how deeply they are breathing.

19.7 Interdependence

Another important factor in maintaining alveolar stability is *interdependence*. Each alveolus in the lungs is surrounded by other alveoli (see Figure 19.3a). Furthermore, all of these alveoli are interconnected with each other by connective tissue. Because of these interconnections, any tendency for an alveolus to collapse is opposed by the surrounding alveoli. As the central alveolus collapses, it pulls outward on the surrounding alveoli, stretching them and distorting their shape. In response, the distorted alveoli pull back in the opposite direction to regain their normal shape. In other words, they exert

PHARMACY APPLICATION: INFANT RESPIRATORY DISTRESS SYNDROME

Infant respiratory distress syndrome (IRDS), also known as hyaline membrane disease, is one of the most common causes of respiratory disease in premature infants. In fact, it occurs in 30,000 to 50,000 newborns per year in the United States; most commonly in neonates born before week 28 of gestation (60%). IRDS is characterized by areas of atelectasis, hemorrhagic edema, and the formation of hyaline membranes within the alveoli.

IRDS is caused by a deficiency of pulmonary surfactant. Alveolar type II cells, which produce surfactant, do not begin to mature until weeks 25 to 28 of gestation. Therefore, premature infants may have poorly functioning alveolar type II cells and insufficient surfactant production.

At birth, the first breath taken by the neonate requires high inspiratory pressures to cause the initial expansion of the lungs. Normally, the lungs will retain a portion of this first breath (40% of the residual volume) so that subsequent breaths require much lower inspiratory pressures. In infants lacking surfactant, the lungs collapse between breaths. The airless portions of the lungs become stiff and noncompliant. Therefore, every inspiration is as difficult as the first. In fact, a transpulmonary pressure of 25 to 30 mmHg is needed to maintain a patent airway (compared to the normal 5 mmHg). This results in a significant increase in the work of breathing and a decrease in ventilation. The inability of the neonate to ventilate adequately leads to progressive atelectasis, hypoxemia, hypercapnea (increased carbon dioxide), and respiratory acidosis. Furthermore, the formation of the hyaline membranes impairs gas exchange, which exacerbates these conditions.

The therapy for IRDS includes mechanical ventilation with continuous positive airway pressure. This maintains adequate ventilation and prevents airway collapse between breaths.

Therapy also includes the administration of exogenous pulmonary surfactant. There are two types of surfactant used to prevent and treat IRDS in the United States. These include surfactants prepared from animal sources as well as synthetic surfactants. Exogenous pulmonary surfactants are administered as a suspension (in saline) through the endotracheal tube used for mechanical ventilation.

Many exogenous pulmonary surfactants are derived from bovine extracts. For example, Infasurf® contains the active ingredient calfactant, which is an unmodified calf lung extract that includes mostly phospholipids and hydrophobic surfactant-specific proteins. Other exogenous pulmonary surfactants derived from bovine lung extracts include Survanta® (active ingredient, beractant) and Bovactant® (active ingredient, alveofact).

Exosurf® is a synthetic surfactant. It contains colfosceril palmitate, which is a phospholipid and an important constituent of natural and many synthetic pulmonary surfactant compounds.

radial traction on the central alveolus. As a result, the alveolus is pulled open and collapse is prevented.

19.8 Airway resistance

The factors determining the flow of air through the airways are analogous to those determining the flow of blood through the vessels and are described by Ohm's Law:

$$\text{Airflow} = \frac{\Delta P}{R}$$

Airflow through the airways is proportional to the gradient between atmospheric pressure and alveolar pressure (ΔP) and inversely proportional to the *airway resistance* (R).

The factors determining the resistance to airflow are also analogous to those determining the resistance to blood flow and include viscosity, length of the airway, and airway radius. Under normal conditions, the viscosity of the air is fairly constant, and the length of the airway is fixed. Therefore, airway radius is the critically important physiological factor determining airway resistance:

$$R \propto \frac{1}{r^4}$$

Airway resistance is inversely proportional to the *radius* (r) of the airway to the fourth power. In other words, when the radius is reduced by a factor of two (50%), the airway resistance increases sixteenfold.

In healthy lungs, the resistance to airflow is so small that small changes in pulmonary pressure result in large volumes of airflow. As discussed previously, a pressure gradient of 1 mmHg between the atmosphere and the alveolus allows 500 ml of air to enter the lungs. However, there are several factors that may alter airway resistance:

- Lung volume
- Airway obstruction
- Bronchial smooth muscle tone

19.8.1 Lung volume

As *lung volume* increases, the airway resistance decreases. In other words, as the lungs inflate, the airways expand and become larger. The increase in airway radius decreases airway resistance. Conversely, as lung volume decreases, the airway resistance increases. In fact, at very low lung volumes, the small airways may close completely. This is a problem especially at the base of the lungs where, due to the weight of the lungs, the airways are less well expanded.

19.8.2 Airway obstruction

Airway obstruction may be caused by several factors:

- Excess mucus production
- Inflammation and edema of the airway wall
- Airway collapse

Both asthma and chronic bronchitis are characterized by the *overproduction of thick, viscous mucus* (see Figure 19.3b). This mucus blocks the airways and, in effect, reduces the radius of the airways and increases airway resistance. A severe asthmatic attack may be accompanied by the formation of mucus plugs, which can completely obstruct airflow.

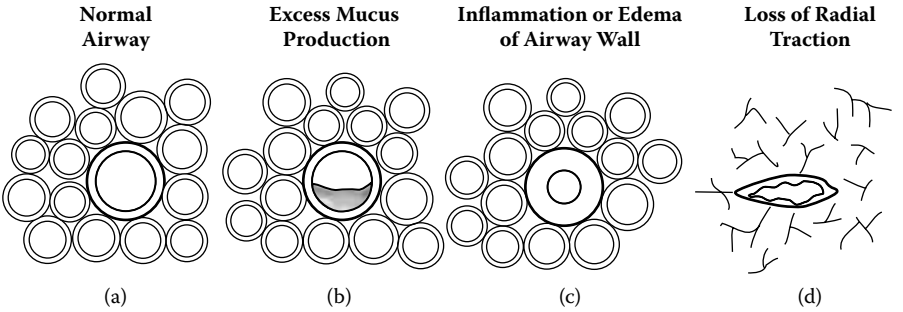


Figure 19.3 Airway obstruction. (a) Normal airway. Illustrated is a normal, patent airway with radial traction offered to it by the surrounding airways. Resistance in this airway is low and air flows through it freely. (b) Excess mucus production. The airway is obstructed by the presence of excess mucus and airway resistance is increased. Airflow is reduced. (c) Inflammation or edema of the airway wall. Thickening of the airway wall due to inflammation or edema narrows the lumen of the airway. The decrease in airway radius increases airway resistance and decreases airflow. (d) Loss of radial traction. Destruction of surrounding airways results in the loss of interdependence, or radial traction. Without the structural support offered by surrounding airways, the central airway collapses, and airflow through it is reduced.

Asthma and chronic bronchitis, which are considered chronic inflammatory conditions, are also characterized by *inflammation and edema of the airway walls* (see Figure 19.3c). This thickening of the airway wall narrows the lumen of the airway and increases airway resistance. The increase in airway resistance due to excess mucus production and inflammation is reversible pharmacologically.

The pathophysiology of emphysema involves the breakdown, or destruction, of alveoli. This results in the loss of interdependence or the effect of radial traction on the airways and leads to *airway collapse* (see Figure 19.3d). The increase in airway resistance due to this form of lung obstruction is irreversible.

19.8.3 Bronchial smooth muscle tone

Changes in *bronchial smooth muscle tone* are particularly important in the bronchioles compared to the bronchi. Recall that the walls of the bronchioles consist almost entirely of smooth muscle. Contraction and relaxation of this muscle have a marked effect on the internal radius of the airway. An increase in bronchial smooth muscle tone, or *bronchoconstriction*, narrows the lumen of the airway and increases the resistance to airflow.

The activation of irritant receptors in the trachea and large bronchi by airborne pollutants, smoke, and noxious chemicals elicits reflex bronchoconstriction. This reflex is mediated by the *parasympathetic nervous system*, specifically, by the vagus nerve. Acetylcholine released from the vagus nerve stimulates muscarinic receptors on the bronchial smooth muscle to cause bronchoconstriction. This parasympathetic reflex is meant to be a protective response, limiting the penetration of toxic substances deep into the lungs. Parasympathetic stimulation of the lungs also enhances mucus production in an effort to trap inhaled particles.

Bronchoconstriction is also elicited by several endogenous chemicals released from mast cells during an allergy or asthmatic attack. These substances, including histamine and the leukotrienes, may also promote the inflammatory response and edema formation.

A decrease in bronchial smooth muscle tone, or *bronchodilation*, widens the lumen of the airway and decreases the resistance to airflow. *Sympathetic nervous stimulation* causes bronchodilation. The adrenergic receptors found on the airway smooth muscle are β_2-adrenergic receptors. Recall that norepinephrine has a very low affinity for these receptors. Therefore, direct sympathetic stimulation of the airways has little effect. Epinephrine, released from the adrenal medulla, causes most of this bronchodilation. Epinephrine has a strong affinity for β_2-adrenergic receptors. Therefore, during a mass sympathetic discharge, as occurs during exercise or the "fight-or-flight" response, epinephrine-induced bronchodilation minimizes airway resistance and maximizes airflow.

PHARMACY APPLICATION: PHARMACOLOGICAL TREATMENT OF ASTHMA

Bronchial asthma is defined as a chronic inflammatory disease of the lungs. It affects an estimated 9 to 12 million individuals in the United States. Furthermore, its prevalence has been increasing in recent years. Asthma is characterized by reversible airway obstruction (in particular, bronchospasm), airway inflammation, and increased airway responsiveness to a variety of bronchoactive stimuli. There are many factors that may induce an asthmatic attack, including allergens, respiratory infections, hyperventilation, cold air, exercise, various drugs and chemicals, emotional upset, and airborne pollutants (smog, cigarette smoke).

The desired outcome in the pharmacological treatment of asthma is to prevent or relieve the reversible airway obstruction and the airway hyperresponsiveness caused by the inflammatory process. Therefore, categories of medications include bronchodilators and antiinflammatory drugs.

A commonly prescribed class of bronchodilators is the β_2-adrenergic receptor agonists (e.g., albuterol, metaproterenol). These drugs cause relaxation of bronchial smooth muscle and relieve the congestion of the bronchial mucosa. The β_2-adrenergic receptor agonists are useful during an acute asthmatic attack. Furthermore, they are effective when taken prior to exercise in individuals with exercise-induced asthma. These drugs are usually administered by inhalation or by a nebulizer.

Another bronchodilator is ipratropium, an anticholinergic drug that blocks muscarinic receptors on the airway smooth muscle. This results in bronchodilation, particularly in large airways. This agent has no effect on the composition or viscosity of the bronchial mucus. Also used to treat acute asthmatic attacks, ipratropium is administered by inhalation.

Corticosteroids (e.g., beclomethazone, flunisolide, triamcinolone) have both antiinflammatory and immunosuppressant actions. These drugs are used prophylactically to prevent the occurrence of asthma in patients with frequent attacks. Because they are not useful during an acute attack, corticosteroids are prescribed along with maintenance bronchodilators. More recently, a combination medication has been developed. Advair® contains the corticosteroid fluticasone propionate as well as the β_2-adrenergic receptor antagonist salmeterol. Each of these drugs is administered by inhalation.

Cromolyn is another antiinflammatory agent that is used prophylactically to prevent an asthmatic attack. The exact mechanism of action of cromolyn is not fully understood; however, it is likely to involve the stabilization of mast cells. This prevents the release of the inflammatory

mast cell mediators involved in inducing an asthmatic attack. Cromolyn has proven effective in patients with exercise-induced asthma.

Cromolyn is another antiinflammatory agent that is used prophylactically to prevent an asthmatic attack. The exact mechanism of action of cromolyn is not fully understood; however, it is likely to involve the stabilization of mast cells. This prevents the release of the inflammatory mast cell mediators involved in inducing an asthmatic attack. Cromolyn has proven effective in patients with exercise-induced asthma.

A more recent development in the treatment of asthma involves the leukotriene-receptor antagonists and leukotriene-synthesis inhibitors. Receptor antagonists include zafirlukast (Accolate®) and montelukast (Singulair®). Zileuton (Zyflo®) inhibits the enzyme 5-lipoxygenase which synthesizes the leukotrienes from arachidonic acid. These drugs have proven to be effective prophylactic treatment for mild asthma.

19.9 Ventilation

Ventilation is the exchange of air between the external atmosphere and the alveoli. It is typically defined as the volume of air entering the alveoli per minute. A complete understanding of ventilation requires the consideration of lung volumes.

19.9.1 Standard lung volumes

The size of the lungs and, therefore, the lung volumes depend upon an individual's height, weight or body surface area, age, and gender. This discussion will include the typical values for a 70-kg adult male. There are *four standard lung volumes* (see Figure 19.4 and Table 19.1):

1. Tidal Volume
2. Residual Volume
3. Expiratory Reserve Volume
4. Inspiratory Reserve Volume

The *tidal volume* (V_T) is the volume of air that enters the lungs per breath. During normal, quiet breathing, tidal volume is 500 ml per breath. This volume increases significantly during exercise. The *residual volume* (RV) is the volume of air remaining in the lungs following a maximal forced expiration. Dynamic compression of the airways causes collapse and the trapping of air in the alveoli. Residual volume is normally 1.5 liters. It can be much greater in patients with emphysema because of the increased tendency for airway collapse. *Expiratory reserve volume* (ERV) is the volume of air expelled from the lungs during a maximal forced expiration beginning at the end

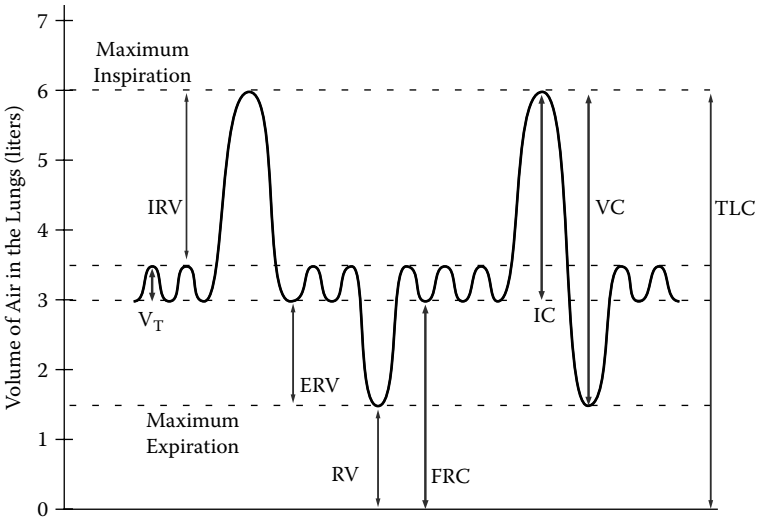


Figure 19.4 Standard lung volumes and lung capacities. Illustrated are the typical values for a 70-kg adult male. Tidal volume (V_T) is 500 ml during normal, quiet breathing. Inspiratory reserve volume (IRV) is obtained with a maximal inspiration and is 2.5 L. Tidal volume and IRV together determine the inspiratory capacity (IC), which is 3.0 L. Expiratory reserve volume (ERV) is obtained with a maximal expiration and is 1.5 L. The volume of air remaining in the lungs following a maximal expiration is the residual volume (RV), which is 1.5 L. The functional residual capacity (FRC) is the volume of air remaining in the lungs following a normal expiration and is 3.0 L. Vital capacity (VC) is obtained with the deepest inspiration and the most forceful expiration and is 4.5 L. The maximum volume to which the lungs can be expanded is the total lung capacity (TLC) and is approximately 6.0 L in adult males and 5.0 L in adult females.

of a normal expiration. The ERV is normally about 1.5 liters. The *inspiratory reserve volume* (IRV) is the volume of air inhaled into the lungs during a maximal forced inspiration beginning at the end of a normal inspiration. The IRV is normally about 2.5 liters. It is determined by the strength of contraction of the inspiratory muscles and the inward elastic recoil of the lungs.

There are *four standard lung capacities*, which consist of two or more lung volumes in combination (see Figure 19.4):

1. Functional Residual Capacity
2. Inspiratory Capacity
3. Total Lung Capacity
4. Vital Capacity

The *functional residual capacity* (FRC) is the volume of air remaining in the lungs at the end of a normal expiration. The FRC consists of the residual

Table 19.1 Lung Volumes and Lung Capacities in a Healthy Adult Male

Lung Volume	Value	Definition
Tidal Volume (V_T)	500 ml	Volume of air that enters the lungs per breath
Residual Volume (RV)	1.5 L	Volume of air remaining in the lungs following a maximal forced expiration
Expiratory Reserve Volume (ERV)	1.5 L	Volume of air expelled from the lungs during a maximal forced expiration (beginning at the end of a normal expiration)
Inspiratory Reserve Volume (IRV)	2.5 L	Volume of air inhaled into the lungs during a maximal forced inspiration (beginning at the end of a normal inspiration)
Lung Capacity	**Volume**	**Definition**
Functional Residual Capacity (FRC)	3.0 L	Volume of air remaining in the lungs at the end of a normal expiration (RV + ERV)
Inspiratory Capacity (IC)	3.0 L	Volume of air that enters the lungs during a maximal forced inspiration beginning at the end of a normal expiration (V_T + IRV)
Total Lung Capacity (TLC)	6.0 L	Volume of air in the lungs following a maximal forced inspiration (RV + ERV + V_T + IRV)
Vital Capacity (VC)	4.5 L	Volume of air expelled from the lungs during a maximal forced expiration following a maximal forced inspiration (IRV + V_T + ERV)

volume and the expiratory reserve volume and is equal to 3 liters. The *inspiratory capacity* (IC) is the volume of air that enters the lungs during a maximal forced inspiration beginning at the end of a normal expiration (FRC). The IC consists of the tidal volume and the inspiratory reserve volume and is equal to 3 liters. The *total lung capacity* (TLC) is the volume of air in the lungs following a maximal forced inspiration. In other words, it is the maximum volume to which the lungs can be expanded. It is determined by the strength of contraction of the inspiratory muscles and the inward elastic recoil of the lungs. The TLC consists of all four lung volumes and is equal to about 6 liters in a healthy adult male and about 5 liters in a healthy adult female. The *vital capacity* (VC) is the volume of air expelled from the lungs during a maximal forced expiration following a maximal forced inspiration. In others words, it consists of the tidal volume as well as the inspiratory and expiratory reserve volumes. Vital capacity is approximately 4.5 liters.

These lung volumes and capacities are measured using a spirometer. When the patient inhales, the pen moves upward on the paper affixed to a rotating drum. When the patient exhales, the pen moves downward on the paper. Pulmonary function tests are useful in the diagnosis of pulmonary disease.

19.9.2 Total ventilation

The *total ventilation* (*minute volume*) is the volume of air that enters the lungs per minute. It is determined by tidal volume and breathing frequency:

$$\begin{aligned}\text{Total ventilation} &= \text{tidal volume} \times \text{breathing frequency}\\ &= 500\ \text{ml/breath} \times 12\ \text{breaths/min}\\ &= 6000\ \text{ml/min}\end{aligned}$$

With an average tidal volume of 500 ml/breath and breathing frequency of 12 breaths/min, 6000 ml or 6 liters of air move in and out of the lungs per minute. These values apply to conditions of normal, quiet breathing. Both tidal volume and breathing frequency increase substantially during exercise.

19.9.3 Alveolar ventilation

Alveolar ventilation is less than the total ventilation because the last portion of each tidal volume remains in the conducting airways. Therefore, that air does not participate in gas exchange. As mentioned at the beginning of the chapter, the volume of the conducting airways is referred to as *anatomical dead space*. The calculation of alveolar ventilation includes the tidal volume adjusted for the anatomical dead space and includes only the air that actually reaches the respiratory zone:

$$\begin{aligned}\text{Alveolar ventilation} &= (\text{tidal volume} - \text{anatomical dead space})\\ &\quad \times \text{breathing frequency}\\ &= (500\ \text{ml/breath} - 150\ \text{ml dead space}) \times 12\ \text{breaths/min}\\ &= 4200\ \text{ml/min}\end{aligned}$$

During exercise, the working muscles need to obtain more oxygen and eliminate more carbon dioxide. Alveolar ventilation is increased accordingly. Interestingly, the increase in tidal volume is greater than the increase in breathing frequency. This is the most efficient mechanism by which to enhance alveolar ventilation. Using the values above, a twofold increase in breathing frequency, from 12 breaths/min to 24 breaths/min, results in an alveolar ventilation of 8400 ml/min. In other words, alveolar ventilation also increases by a factor of two. However, a twofold increase in tidal volume, from 500 ml/breath to 1000 ml/breath, results in an alveolar ventilation of 10,200 ml/min. Alveolar ventilation is enhanced more in this case because a greater percentage of the tidal volume reaches the alveoli. At a tidal volume

Table 19.2 Effect of Tidal Volume and Breathing Frequency on Alveolar Ventilation

Patient	Tidal Volume (ml/breath)	Frequency (breaths/min)	Total Ventilation (ml/min)	Alveolar Ventilation (V_T – ADS) (ml/min)
A	500	12	6,000	4,200
B	500	24	12,000	8,400
C	1,000	12	12,000	10,200

of 500 ml/breath and an anatomical dead space of 150 ml, 30% of the inspired air is wasted as it does not reach the alveoli to participate in gas exchange. However, when the tidal volume is 1000 ml/breath, only 15% of the inspired air remains in the anatomical dead space (see Table 19.2).

19.9.4 Dead space

Anatomical dead space is equal to the volume of the conducting airways. This is determined by the physical characteristics of the lungs, as, by definition, these airways do not contain alveoli to participate in gas exchange.

Alveolar dead space is the volume of air that enters unperfused alveoli. In other words, these alveoli receive airflow but no blood flow. With no blood flow to the alveoli, gas exchange cannot take place. Therefore, alveolar dead space is based on functional considerations rather than anatomical factors. Healthy lungs have little or no alveolar dead space. Various pathological conditions, such as low cardiac output, may result in alveolar dead space.

The anatomical dead space combined with the alveolar dead space is referred to as *physiological dead space*:

Physiological dead space = anatomical dead space + alveolar dead space

Physiological dead space is determined by measuring the amount of carbon dioxide in the expired air. Therefore, it is based on the functional characteristics of the lungs as only perfused alveoli can participate in gas exchange and eliminate carbon dioxide.

19.10 Diffusion

Oxygen and carbon dioxide cross the blood–gas interface by way of *diffusion*. The factors that determine the rate of diffusion of each gas are described by *Fick's Law of Diffusion*:

$$V_{gas} \alpha \frac{A \times D \times (\Delta P)}{T}$$

PHARMACY APPLICATION:
DRUG-INDUCED HYPOVENTILATION

Hypoventilation is defined as a reduction in the rate and depth of breathing. Inadequate ventilation results in hypoxemia, or a decrease in the oxygen content of the arterial blood. Hypoventilation may be induced inadvertently by various pharmacological agents, including opioid analgesics such as morphine. These medications cause hypoventilation by way of their effects on the respiratory centers in the brainstem. Doses of morphine too small to alter a patient's consciousness may cause discernible respiratory depression. This inhibitory effect on the respiratory drive increases progressively as the dose of morphine is increased. In fact, in humans, death due to morphine poisoning is almost always due to respiratory arrest. Although therapeutic doses of morphine decrease tidal volume, the decrease in breathing frequency is the primary cause of decreased minute volume.

Diffusion is proportional to the surface area of the blood–gas interface (A), the diffusion constant (D), and the partial pressure gradient of the gas (ΔP). Diffusion is inversely proportional to the thickness of the blood–gas interface (T).

The *surface area* of the blood–gas interface is about 70 m^2 in a healthy adult at rest. Specifically, 70 m^2 of the potential surface area for gas exchange in the lungs is both ventilated and perfused. The amount of this surface area may be altered under various conditions. For example, during exercise, an increased number of pulmonary capillaries are perfused (due to increased cardiac output and, therefore, blood flow through the lungs). As a result, a larger percentage of the alveoli are both ventilated and perfused, which increases the surface area for gas exchange. Conversely, a fall in the cardiac output reduces the number of perfused capillaries and, therefore, reduces the surface area for gas exchange. Another pathological condition affecting surface area is emphysema. This pulmonary disease, usually associated with cigarette smoking, causes destruction of alveoli.

The *diffusion constant* for a gas is proportional to the solubility of the gas and inversely proportional to the square root of the molecular weight of the gas:

$$D \alpha \frac{\text{solubility}}{\sqrt{MW}}$$

Both oxygen and carbon dioxide are small molecules with low molecular weights. However, carbon dioxide is 20 times more soluble than oxygen. Therefore, the value of the diffusion constant for carbon dioxide is larger

than that of oxygen which facilitates the exchange of carbon dioxide across the blood–gas interface.

The *thickness* of the blood–gas interface is normally less than 0.5 µ. This extremely thin barrier promotes the diffusion of gases. The thickness may increase, however, under conditions of interstitial fibrosis, interstitial edema, and pneumonia. Fibrosis involves the excess production of collagen fibers by fibroblasts in the interstitial space. Edema is the movement of fluid from the capillaries into the interstitial space. Pneumonia causes inflammation and alveolar flooding. In each case, the thickness of the barrier between the air and the blood is increased, and diffusion is impaired.

The diffusion of oxygen and carbon dioxide also depends on their *partial pressure gradients*. Oxygen diffuses from an area of high partial pressure in the alveoli to an area of low partial pressure in the pulmonary capillary blood. Conversely, carbon dioxide diffuses down its partial pressure gradient from the pulmonary capillary blood into the alveoli.

According to *Dalton's Law*, the partial pressure of a gas (P_{gas}) is equal to its fractional concentration (% total gas) multiplied by the total pressure (P_{tot}) of all of the gases in a mixture:

$$P_{gas} = \% \text{ total gas} \times P_{tot}$$

The atmosphere is a mixture of gases containing 21% oxygen and 79% nitrogen. Due to the effects of gravity, this mixture exerts a total atmospheric pressure (barometric pressure) of 760 mmHg at sea level. Using these values of fractional concentration and total pressure, the partial pressures for oxygen (PO_2) and nitrogen (PN_2) can be calculated:

$$PO_2 = 0.21 \times 760 \text{ mmHg} = 160 \text{ mmHg}$$

$$PN_2 = 0.79 \times 760 \text{ mmHg} = 600 \text{ mmHg}$$

The PO_2 of the atmosphere at sea level is 160 mmHg and the PN_2 is 600 mmHg. The total pressure (760 mmHg) is equal to the sum of the partial pressures.

Under normal, physiological conditions, the partial pressure gradient for oxygen between the alveoli and the pulmonary capillary blood is quite substantial. However, this gradient may be diminished under certain conditions, such as ascent to altitude and hypoventilation. Altitude has no effect on the concentration of oxygen in the atmosphere. However, the effects of gravity on barometric pressure progressively decrease as elevation increases. For example, at an elevation of 17,000 feet, which is the height of Pike's Peak in the Rocky Mountains of Colorado, the barometric pressure is only 380 mmHg. Therefore, the PO_2 of the atmosphere at this altitude is 80 mmHg ($PO_2 = 0.21 \times 380$ mmHg = 80 mmHg). This results in a marked decrease in

the partial pressure gradient between the alveoli and the pulmonary capillary blood. Consequently, diffusion is significantly impaired.

Hypoventilation decreases the rate of oxygen uptake into the alveoli. Once again, the partial pressure gradient and the rate of diffusion are reduced. Conditions resulting in impaired diffusion lead to the development of *hypoxemia*, or decreased oxygen in the arterial blood.

19.11 Partial pressures of oxygen and carbon dioxide

As explained in the previous section, the PO_2 of the atmosphere is 160 mmHg. The partial pressure of carbon dioxide (PCO_2) is negligible (see Table 19.3 and Figure 19.5).

Table 19.3 Partial Pressures of Oxygen and Carbon Dioxide

Location	PO_2 (mmHg)	PCO_2 (mmHg)
Atmosphere	160	0
Conducting Airways (inspired)	150	0
Alveolar Gas	100	40
Arterial Blood	100	40
Tissues	40	45
Mixed Venous Blood	40	45

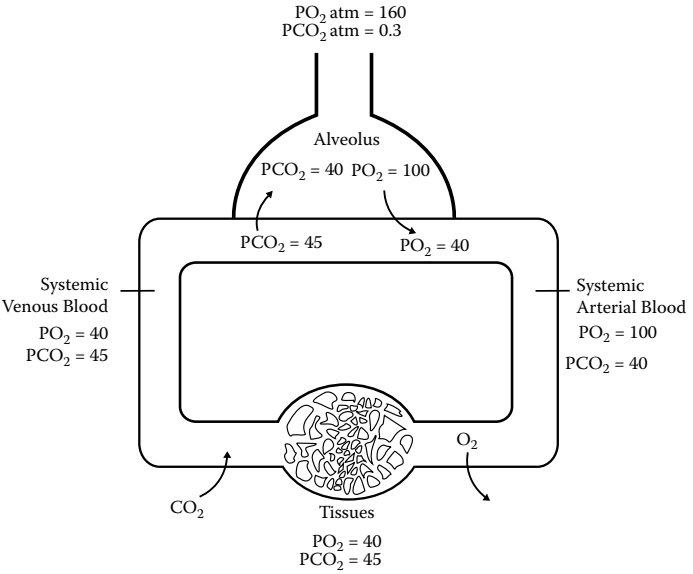


Figure 19.5 Partial pressures of oxygen and carbon dioxide throughout the respiratory and cardiovascular system.

As air is inspired, it is warmed and humidified as it flows through the conducting airways. Therefore, water vapor is added to the gas mixture. This is accounted for in the calculation of PO_2 in the conducting airways:

$$PO_2 \text{ inspired air} = 0.21 \times (760 \text{ mmHg} - 47 \text{ mmHg})$$

$$= 150 \text{ mmHg}$$

The partial pressure of the water vapor is 47 mmHg. As a result, the PO_2 in the conducting airways during inspiration is slightly decreased to 150 mmHg. The PCO_2 remains at 0 mmHg.

By the time the air reaches the alveoli, the PO_2 has decreased to about 100 mmHg. The PO_2 of the alveolar gas is determined by two processes:

1. The rate of replenishment of oxygen by ventilation.
2. The rate of removal of oxygen by the pulmonary capillary blood.

The primary determinant of alveolar PO_2 is the rate of replenishment of oxygen by ventilation. As mentioned previously, hypoventilation causes a decrease in alveolar PO_2. The rate of removal of oxygen by the pulmonary capillary blood is determined largely by the oxygen consumption of the tissues. As metabolic activity and oxygen consumption increase, the PO_2 of the mixed venous blood decreases. As a result, the partial pressure gradient for oxygen between the alveoli and the blood is increased, and the diffusion of oxygen is enhanced.

The PCO_2 of the alveoli is about 40 mmHg. The alveolar PCO_2 is also determined by two processes:

1. The rate of delivery of carbon dioxide to the lungs from the tissues.
2. The rate of elimination of carbon dioxide by ventilation.

As cellular metabolism increases, the rate of production of carbon dioxide also increases. Typically, increased activity is associated with an increase in ventilation so that the increased amounts of carbon dioxide delivered to the lungs are readily eliminated. On the other hand, hypoventilation impairs the elimination of carbon dioxide and causes an increase in alveolar PCO_2.

Assuming that oxygen diffuses down its partial pressure gradient from the alveoli into the pulmonary capillary blood until equilibration is reached, then the PO_2 of this blood reaches 100 mmHg. This is the pressure exerted by oxygen molecules in their *gas phase.* In other words, the PO_2 of the blood is determined by the amount of oxygen *dissolved in plasma* as opposed to the amount of oxygen bound to hemoglobin (a concept discussed later in this chapter). This blood flows back to the left side of the heart and into the systemic circulation. Therefore, the PO_2 of the arterial blood is 100 mmHg. Likewise, assuming that carbon dioxide diffuses down its partial pressure

gradient from the pulmonary capillary blood into the alveoli until equilibration is reached, then the PCO_2 of the blood leaving these capillaries should be 40 mmHg. Therefore, the PCO_2 of the arterial blood is 40 mmHg.

The arterial blood, which is high in oxygen and low in carbon dioxide, is delivered to the tissues. Within the tissues, oxygen is consumed by metabolism and carbon dioxide is produced. Under typical resting conditions, the PO_2 of the tissues is 40 mmHg. Therefore, oxygen diffuses down its concentration gradient from the systemic capillary blood into the cells of the tissues until equilibration is reached. The PO_2 of the venous blood leaving the tissues is also 40 mmHg. The PCO_2 of the tissues is 45 mmHg. Therefore, carbon dioxide diffuses down its concentration gradient from the tissues into the blood until equilibration is reached. The PCO_2 of the venous blood leaving the tissues is 45 mmHg.

The mixed venous blood, which is low in oxygen and high in carbon dioxide, flows back to the lungs to obtain oxygen and eliminate carbon dioxide. Note that the partial pressure gradient for oxygen between the alveoli (100 mmHg) and the mixed venous blood (40 mmHg) is 60 mmHg. The partial pressure gradient for carbon dioxide between the mixed venous blood (45 mmHg) and the alveoli (40 mmHg) is 5 mmHg. According to Fick's Law of Diffusion, the small partial pressure gradient for carbon dioxide would tend to reduce the exchange of this gas. However, the relatively high solubility and diffusion constant of carbon dioxide allows it to diffuse quite readily across the blood–gas interface.

19.12 Ventilation–perfusion matching

In order to optimize gas exchange, the uptake of oxygen from the alveolar gas into the pulmonary blood and the elimination of carbon dioxide from the pulmonary blood into the alveolar gas, a given lung unit must be equally well ventilated and perfused. In other words, the air and the blood must be brought together for the exchange of gases to occur efficiently. This is referred to as *ventilation–perfusion* (*V/Q*) *matching*. The most effective conditions for gas exchange occur when the V/Q ratio is equal to one, or when the amount of ventilation in a lung unit is balanced, or matched, by the amount of perfusion. In this region, the mixed venous blood entering the pulmonary capillaries has a PO_2 of 40 mmHg and a PCO_2 of 45 mmHg. The alveolar gas has a PO_2 of 100 mmHg and a PCO_2 of 40 mmHg. Ventilation–perfusion matching results in efficient gas exchange and a PO_2 of 100 mmHg and a PCO_2 of 40 mmHg in the blood leaving the capillaries and returning to the heart (see Figure 19.6a).

Airway obstruction leads to a reduction in the V/Q ratio to a value less than one. In this lung unit, perfusion is greater than ventilation (see Figure 19.6b). Complete airway obstruction leads to *shunt*. Shunt refers to blood that enters the arterial system without passing through a region of ventilated lung. In other words, mixed venous blood travels through the pulmonary circulation

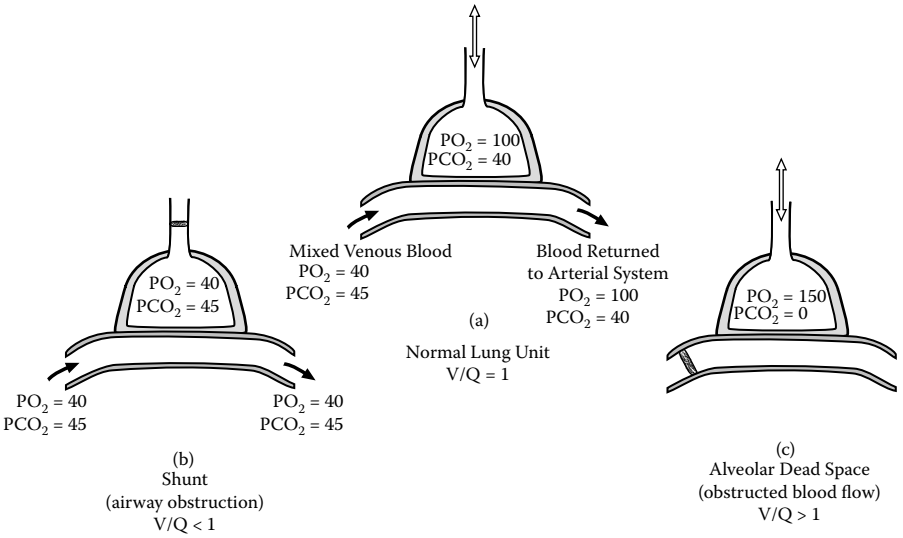


Figure 19.6 Ventilation–perfusion matching. (a) Normal lung unit. Ventilation and perfusion are matched so that the V/Q ratio is equal to one and gas exchange is optimized. Mixed venous blood that is low in oxygen and high in carbon dioxide enters the pulmonary capillaries. As the blood flows through these capillaries in the walls of the alveoli, oxygen is obtained and carbon dioxide is eliminated. Blood returning to the heart and the arterial system is high in oxygen and low in carbon dioxide. (b) Shunt. Airway obstruction with normal perfusion can lead to shunt (V/Q < 1). Blood flows through the lungs without obtaining oxygen or eliminating carbon dioxide. This V/Q mismatch causes hypoxemia. (c) Alveolar dead space. Obstructed blood flow with normal ventilation causes the development of alveolar dead space (V/Q > 1). The partial pressures of oxygen and carbon dioxide in the alveoli are similar to those of the conducting airways. This ventilation is wasted as it does not participate in gas exchange.

without participating in gas exchange. This blood enters the pulmonary capillaries with a PO_2 of 40 mmHg and a PCO_2 of 45 mmHg. If the lung unit is not ventilated, then this blood exits the capillaries and returns to the heart with the partial pressures of oxygen and carbon dioxide unchanged. The addition of blood that is low in oxygen to the rest of the blood returning from the lungs causes *hypoxemia*. The degree of hypoxemia is determined by the magnitude of the shunt. As airway obstruction increases throughout the lungs, then this widespread decrease in the V/Q ratio results in a greater volume of poorly oxygenated blood returning to the heart and a greater degree of hypoxemia. As discussed, airway obstruction may be caused by many factors including bronchoconstriction, excess mucus production, airway collapse, and alveolar flooding.

Obstructed blood flow leads to an increase in the V/Q ratio to a value greater than one. In this lung unit, ventilation is greater than perfusion (see

Figure 19.6c). Complete loss of blood flow leads to *alveolar dead space*. In this lung unit, the air enters the alveoli with partial pressures of oxygen and carbon dioxide equal to those of the conducting airways (PO_2 of 150 mmHg and PCO_2 of 0 mmHg). With no perfusion, oxygen is not taken up from this mixture, and carbon dioxide is not added to the mixture to be eliminated. Alveolar dead space may be caused by pulmonary thromboembolism. This is when a pulmonary blood vessel is occluded by a blood clot. Alveolar dead space may also occur when alveolar pressure is greater than pulmonary capillary pressure. This leads to compression of the capillaries and a loss of perfusion. Alveolar pressure may be increased by positive pressure mechanical ventilation. Pulmonary capillary pressure may be decreased by hemorrhage and a decrease in cardiac output.

Ventilation–perfusion mismatch leads to hypoxemia. Reduced ventilation caused by obstructed airflow or reduced perfusion caused by obstructed blood flow leads to impaired gas exchange. Interestingly, each of these conditions may be minimized by *local control mechanisms* that attempt to match airflow and blood flow in a given lung unit.

Bronchiolar smooth muscle is sensitive to changes in carbon dioxide levels. Excess carbon dioxide causes bronchodilation, and reduced carbon dioxide causes bronchoconstriction. Pulmonary vascular smooth muscle is sensitive to changes in oxygen levels. Excess oxygen causes vasodilation, and insufficient oxygen (hypoxia) causes vasoconstriction. The changes in bronchiolar and vascular smooth muscle tone alter the amount of ventilation and perfusion in a lung unit to return the V/Q ratio to one.

In a lung unit with high blood flow and low ventilation (airway obstruction), the level of carbon dioxide is increased and the level of oxygen is decreased. The excess carbon dioxide causes bronchodilation and an increase in ventilation. The reduced oxygen causes vasoconstriction and a decrease in perfusion. In this way, the V/Q ratio is brought closer to one and overall pulmonary gas exchange is improved.

In a lung unit with low blood flow and high ventilation (alveolar dead space), the level of carbon dioxide is decreased and the level of oxygen is increased. The reduced carbon dioxide causes bronchoconstriction and a decrease in ventilation. The excess oxygen causes vasodilation and an increase in perfusion. Once again, the V/Q ratio is brought closer to one, and overall pulmonary gas exchange is improved.

19.13 Gas transport in the blood

Once the oxygen has diffused from the alveoli into the pulmonary circulation, it must be carried, or transported, in the blood to the cells and tissues that need it. Furthermore, once the carbon dioxide has diffused from the tissues into the systemic circulation, it must be transported to the lungs where it can be eliminated. This section considers the mechanisms by which these gases are transported.

19.13.1 Transport of oxygen

Oxygen is carried in the blood in two forms:

1. Physically dissolved
2. Chemically combined with hemoglobin

Oxygen is poorly soluble in plasma. At a PO_2 of 100 mmHg, only 3 ml of oxygen is *physically dissolved* in 1 liter of blood. Assuming a blood volume of 5 liters, a total of 15 ml of oxygen is in the dissolved form. A normal rate of oxygen consumption at rest is about 250 ml/min. During exercise, oxygen consumption may increase to 3.5 to 5.5 L/min. Therefore, the amount of dissolved oxygen is clearly insufficient to meet the needs of the tissues.

Most of the oxygen in the blood (98.5%) is transported *chemically combined with hemoglobin*. A large, complex molecule, hemoglobin consists of four polypeptide chains (globin portion), each of which contains a ferrous iron atom (heme portion). Each iron atom can bind reversibly with an oxygen molecule:

$$O_2 + Hb \leftrightarrow HbO_2$$

Therefore, the hemoglobin molecule exists in one of two forms: *oxyhemoglobin (HbO_2)* or *deoxyhemoglobin (Hb)*. The binding of oxygen to hemoglobin follows the *Law of Mass Action*, such that as the PO_2 increases, as it does in the lungs, more will combine with hemoglobin. When the PO_2 decreases, as it does in the tissues that are consuming it, the reaction moves to the left, and the hemoglobin releases the oxygen.

Each gram of hemoglobin can combine with up to 1.34 ml of oxygen. In a healthy individual, there are 15 grams of hemoglobin/100 ml of blood. Therefore, the oxygen content of the blood is 20.1 ml O_2/100 ml blood:

$$\frac{15 \text{ g Hb}}{100 \text{ ml blood}} \times \frac{1.34 \text{ ml } O_2}{\text{g Hb}} = \frac{20.1 \text{ ml } O_2}{100 \text{ ml blood}}$$

It is important to note that oxygen bound to hemoglobin has no effect on the PO_2 of the blood. Once again, the *PO_2 of the blood* is determined by the amount of oxygen dissolved in the plasma. The amount of oxygen bound to hemoglobin determines *oxygen content* of the blood. In a normal healthy individual, the percent of hemoglobin saturation directly reflects the number of oxygen molecules present in the blood. However, as will be discussed, under certain pathophysiological conditions, the percent of hemoglobin saturation may not accurately reflect the oxygen content of the blood.

The PO_2 of the blood is the major factor determining the amount of oxygen chemically combined with hemoglobin, or the *percent of hemoglobin saturation*. The relationship between these two variables is illustrated graphically by

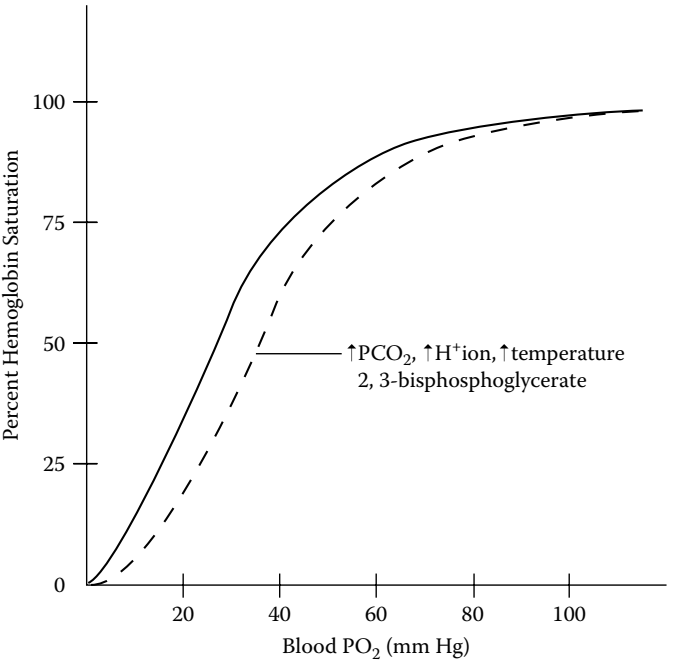


Figure 19.7 Oxyhemoglobin dissociation curve. The percent hemoglobin saturation depends upon the PO_2 of the blood. The PO_2 of the blood in the pulmonary capillaries is 100 mmHg. Consequently, in the lungs, the hemoglobin loads up with oxygen and becomes 97.5% saturated. The average PO_2 of the blood in the systemic capillaries is 40 mmHg. Therefore, at the level of the tissues, hemoglobin releases oxygen and the saturation falls to 75%. Increased PCO_2, H^+ ion concentration, temperature, and 2,3-bisphosphoglycerate shifts the oxyhemoglobin dissociation curve to the right. As a result, at any given PO_2, the hemoglobin releases more oxygen to the tissue.

the *oxyhemoglobin dissociation curve* (see Figure 19.7). As the PO_2 of the blood increases, the combination of oxygen and hemoglobin also increases. However, this relationship is not linear. The amount of oxygen carried by hemoglobin increases steeply up to a PO_2 of about 60 mmHg. Beyond this point, the curve becomes much flatter, such that there is little change in the percent of hemoglobin saturation as PO_2 continues to increase. At a PO_2 of 100 mmHg, which is the normal PO_2 of the alveoli and, therefore, the arterial blood, the hemoglobin is 97.5% saturated with oxygen.

Each region of the curve, the steep portion and the flat plateau portion, has important physiological significance. The *steep portion of the curve*, between 0 mmHg and 60 mmHg, is the PO_2 range found in the cells and tissues. This region of the curve is ideal for *unloading* oxygen to the tissues. On average, the PO_2 of the tissues and, therefore, the mixed venous blood is about 40 mmHg at rest. At a PO_2 of 40 mmHg, the hemoglobin is 75% saturated

with oxygen. In other words, as the blood flows through the systemic capillaries, the hemoglobin releases 22.5% of its oxygen to the tissues. An increase in the metabolic activity of a tissue and, therefore, an increase in oxygen consumption, will decrease the PO_2 in that tissue. The fall in PO_2 in this region of the oxyhemoglobin dissociation curve has a profound effect on the percent of hemoglobin saturation. At a PO_2 of 15 mmHg, the hemoglobin is only 25% saturated with oxygen. In this case, the hemoglobin has released 72.5% of its oxygen to the tissue. Therefore, a relatively small drop in PO_2 (from 40 mmHg down to 15 mmHg) results in a marked increase in the unloading of oxygen (more than three times as much oxygen has been released to the tissue that needs it).

The *plateau portion of the curve*, between 60 mmHg and 100 mmHg, is the PO_2 range found in the alveoli. As the mixed venous blood flows through the pulmonary capillaries in the walls of the alveoli, the hemoglobin loads up with oxygen. As mentioned above, at a normal alveolar PO_2 of 100 mmHg, the hemoglobin becomes almost fully saturated with oxygen (97.5%). Interestingly, at a PO_2 of 60 mmHg, the hemoglobin still becomes 90% saturated with oxygen. In other words, the hemoglobin remains quite saturated with oxygen even with a marked fall in PO_2 (40 mmHg). This provides a good margin of safety for the oxygen-carrying capacity of the blood. Therefore, if an individual ascends to some altitude above sea level or has pulmonary disease such that the alveolar PO_2 falls, the oxygen content of the blood remains high.

19.13.2 Factors affecting the transport of oxygen

There are several factors that affect the transport of oxygen:

- PCO_2, pH, and temperature
- 2,3-bisphosphoglycerate
- Anemia
- Carbon monoxide

An *increase in the PCO_2*, a *decrease in pH,* and an *increase in temperature* all shift the oxyhemoglobin dissociation curve to the right. As a result, at any given PO_2, the hemoglobin releases more oxygen to the tissue (see Figure 19.7). Both carbon dioxide and hydrogen ion can bind to hemoglobin. The binding of these substances changes the conformation of the hemoglobin and reduces its affinity for oxygen. This phenomenon is referred to as the *Bohr effect.* An increase in temperature also reduces the affinity of hemoglobin for oxygen. This effect is of benefit to a metabolically active tissue. As the rate of metabolism increases, as it does during exercise, oxygen consumption and, therefore, the demand for oxygen increase. In addition, the carbon dioxide, the hydrogen ions, and the heat produced by the tissue are all increased. These

products of metabolism facilitate the release of oxygen from the hemoglobin to the tissue that needs it.

2,3-Bisphosphoglycerate (2,3-BPG) is produced by red blood cells. This substance also binds to hemoglobin, shifting the oxyhemoglobin dissociation curve to the right. Once again, the rightward shift of the curve reduces the affinity of hemoglobin for oxygen so that more oxygen is released to the tissues. Levels of 2,3-BPG are increased when the hemoglobin in the arterial blood is chronically undersaturated or, in other words, during *hypoxemia*. Decreased arterial PO_2 may occur at altitude or as the result of various cardiovascular or pulmonary diseases. The rightward shift of the curve is beneficial at the level of the tissues because more of the oxygen that is bound to the hemoglobin is released to the tissues. However, the shift of the curve may be detrimental in the lungs as the loading of hemoglobin may be impaired.

Levels of 2,3-BPG may be decreased in blood stored in a blood bank for as little as one week. A decrease in 2,3-BPG shifts the oxyhemoglobin dissociation curve to the left. In this case, at any given PO_2, the unloading of oxygen to the tissues is decreased. The progressive depletion of 2,3-BPG can be minimized by storing the blood with citrate-phosphate-dextrose.

Anemia causes a decrease in the oxygen content of the blood and, therefore, a decrease in the supply of oxygen to the tissues. It is characterized by a low hematocrit that may be caused by a number of pathological conditions, such as a decreased rate of erythropoiesis (red blood cell production), excessive loss of erythrocytes, or a deficiency of normal hemoglobin in the erythrocytes. Although there is a decrease in the oxygen content of the blood, it is important to note that anemia has no effect on the PO_2 of the blood or on the oxyhemoglobin dissociation curve (see Figure 19.8). Arterial PO_2 is determined only by the amount of oxygen dissolved in the blood, which is unaffected. Furthermore, the affinity of hemoglobin for oxygen has not changed. What has changed is the amount of hemoglobin in the blood. If there is less hemoglobin available to bind with oxygen, then there is less oxygen in the blood.

Anemia does not stimulate ventilation. As will be discussed in a subsequent section, the peripheral chemoreceptors are sensitive to decreases in PO_2, not oxygen content.

Carbon monoxide is a by-product of the combustion of hydrocarbons such as gasoline. It may cause sickness and death due to poisoning because it interferes with the transport of oxygen to the tissues by way of two mechanisms:

1. Formation of carboxyhemoglobin.
2. Leftward shift of the oxyhemoglobin dissociation curve.

Carbon monoxide has a much greater affinity (240 times) for hemoglobin than does oxygen so that *carboxyhemoglobin* is readily formed. Therefore, even

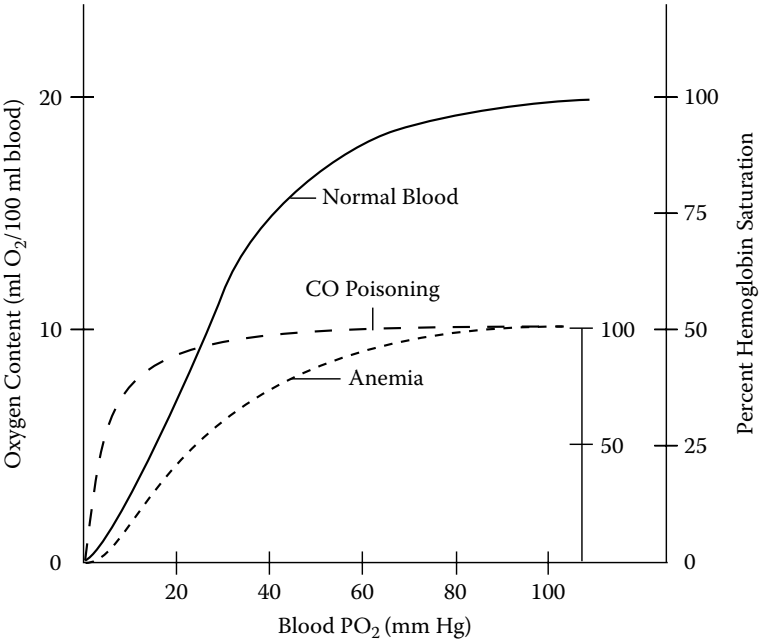


Figure 19.8 Effect of anemia and carbon monoxide poisoning on oxygen transport. Anemia results from a deficiency of normal hemoglobin. The PO_2 of the blood and the percent hemoglobin saturation remain normal. However, because there is less hemoglobin present to transport oxygen, the oxygen content of the blood is decreased. Carbon monoxide impairs the transport of oxygen to the tissues by two mechanisms. First, it binds preferentially with hemoglobin and prevents the hemoglobin from binding with oxygen. As a result, the hemoglobin remains fully saturated (although with carbon monoxide instead of oxygen) and the oxygen content of the blood is decreased. Second, it shifts the oxyhemoglobin dissociation curve to the left and inhibits the release of oxygen from the hemoglobin. As with anemia, the PO_2 of the blood is unaffected.

small amounts of carbon monoxide can tie up the hemoglobin and prevent the loading of oxygen. Furthermore, the formation of carboxyhemoglobin causes a *leftward shift of the oxyhemoglobin dissociation curve* (see Figure 19.8). As a result, at any given PO_2, the unloading of oxygen to the tissues is impaired. Therefore, not only does the hemoglobin carry less oxygen, it does not release this oxygen to the tissues that need it. The concentration of hemoglobin in the blood and the PO_2 of the blood are normal.

Carbon monoxide poisoning is particularly insidious. An individual exposed to carbon monoxide is usually unaware of it as this gas is odorless, colorless, and tasteless. Furthermore, it does not elicit any irritant reflexes resulting in sneezing, coughing, or feelings of dyspnea (difficulty in breathing). Finally, as with anemia, carbon monoxide does not stimulate ventilation.

19.13.3 Transport of carbon dioxide

Carbon dioxide is carried in the blood in three forms:

1. Physically dissolved
2. Carbamino hemoglobin
3. Bicarbonate ions

As with oxygen, the amount of carbon dioxide *physically dissolved* in the plasma is proportional to its partial pressure. However, carbon dioxide is 20 times more soluble in the plasma than is oxygen. Therefore, approximately 10% of the carbon dioxide in the blood is transported in the dissolved form.

Carbon dioxide can combine chemically with the terminal amine groups (NH_2) in blood proteins. The most important of these proteins for this process is hemoglobin. The combination of carbon dioxide and hemoglobin forms *carbamino hemoglobin*:

$$Hb{\cdot}NH_2 + CO_2 \leftrightarrow Hb{\cdot}NH{\cdot}COOH$$

Deoxyhemoglobin can bind more carbon dioxide than oxygenated hemoglobin. Therefore, the unloading of the oxygen in the tissues facilitates the loading of carbon dioxide for transport to the lungs. Approximately 30% of the carbon dioxide in the blood is transported in this form.

The remaining 60% of the carbon dioxide is transported in the blood in the form of *bicarbonate ions*. This mechanism is made possible by the following reaction:

$$H_2O + CO_2 \overset{CA}{\leftrightarrow} H_2CO_3 \leftrightarrow H^+ + HCO_3^-$$

The carbon dioxide produced during cellular metabolism diffuses out of the cells and into the plasma. It then continues to diffuse down its concentration gradient into the red blood cells. Within the red blood cells, the enzyme, *carbonic anhydrase* (CA), facilitates the combination of carbon dioxide and water to form *carbonic acid* (H_2CO_3). The carbonic acid then dissociates into a hydrogen ion (H^+) and a bicarbonate ion (HCO_3^-).

As the bicarbonate ions are formed, they diffuse down their concentration gradient out of the red blood cell and into the plasma. This process is beneficial because bicarbonate ion is far more soluble in the plasma than carbon dioxide. As the negatively charged bicarbonate ions exit the red blood cell, chloride ions, the most abundant anions in the plasma, enter the cells by way of HCO_3^--Cl^- carrier proteins. This process, referred to as the *chloride shift*, maintains electrical neutrality. Many of the hydrogen ions bind with hemoglobin. As with carbon dioxide, deoxyhemoglobin can bind more readily with hydrogen ions than oxygenated hemoglobin.

This entire reaction is reversed when the blood reaches the lungs. Because carbon dioxide is eliminated by ventilation, the reaction is pulled to the left. Bicarbonate ions diffuse back into the red blood cells. The hemoglobin releases the hydrogen ions and is now available to load up with oxygen. The bicarbonate ions combine with the hydrogen ions to form carbonic acid. The carbonic acid then dissociates into carbon dioxide and water. The carbon dioxide diffuses down its concentration gradient from the blood into the alveoli and is exhaled. A summary of the three mechanisms by which carbon dioxide is transported in the blood is illustrated in Figure 19.9.

19.14 Regulation of ventilation

The rate and depth of breathing are perfectly adjusted to meet the metabolic needs of the tissues and to maintain a PO_2 of 100 mmHg, a PCO_2 of 40 mmHg, and a pH of 7.4 in the arterial blood. Breathing is initiated spontaneously by the central nervous system. It occurs in a continuous, cyclical pattern of inspiration and expiration. There are three major components of the regulatory system for ventilation:

1. Medullary respiratory center
2. Receptors and other sources of input
3. Effector tissues (respiratory muscles)

Aggregates of cell bodies within the medulla of the brainstem form the *medullary respiratory center*. There are two distinct functional areas:

1. Dorsal Respiratory Group
2. Ventral Respiratory Group

The aggregate of cell bodies in the dorsal region of the medulla is the *dorsal respiratory group* (DRG). The DRG consists primarily of *inspiratory neurons*. These neurons are self-excitable and repetitively generate action potentials to cause inspiration. The inspiratory neurons descend to the spinal cord where they stimulate neurons that supply the inspiratory muscles, including those of the phrenic nerves and the intercostal nerves. These nerves then stimulate the diaphragm and the external intercostal muscles to contract and cause inspiration. When the inspiratory neurons are electrically inactive, expiration takes place. Therefore, this cyclical electrical activity of the DRG is responsible for the basic rhythm of breathing. Furthermore, the DRG is likely the site of integration of the various sources of input that alter the spontaneous pattern of inspiration and expiration.

The aggregate of cell bodies in the ventral region of the medulla is the *ventral respiratory group* (VRG). The VRG consists of both expiratory and inspiratory neurons. This region is inactive during normal, quiet breathing. (Recall that expiration at this time is a passive process.) However, the VRG

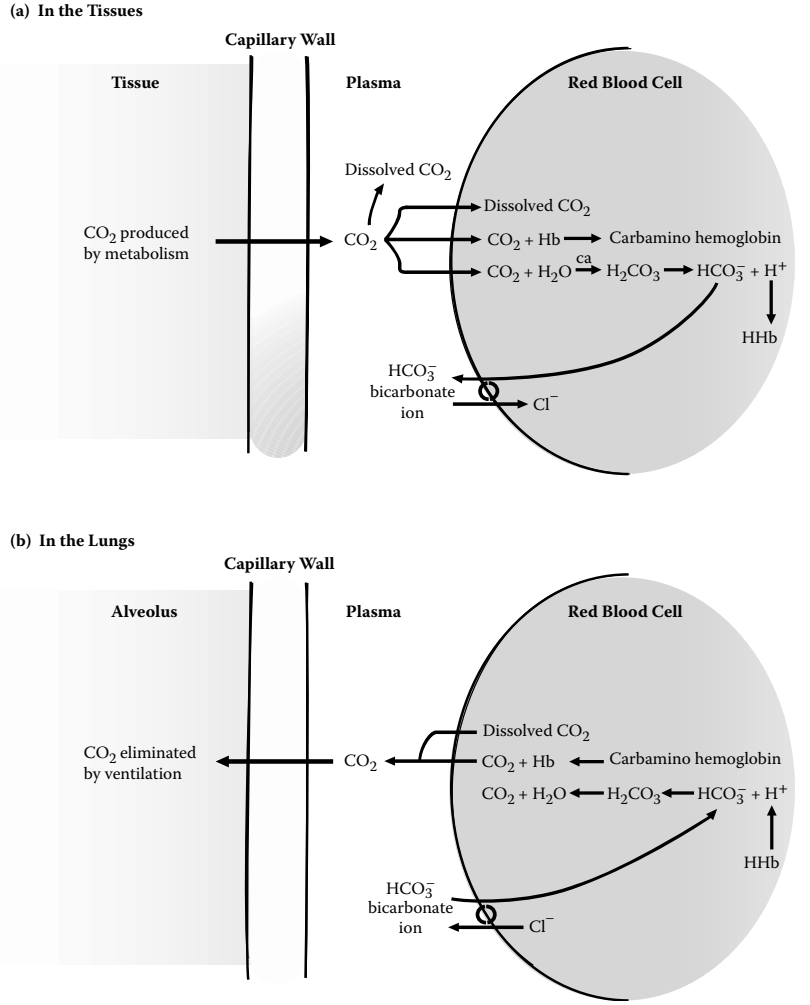


Figure 19.9 Transport of carbon dioxide in the blood. Carbon dioxide (CO_2) is transported in the blood in three forms: dissolved, bound with hemoglobin, and as a bicarbonate ion (HCO_3^-). The carbon dioxide produced by the tissues diffuses down its concentration gradient into the plasma and into red blood cells. A small amount of carbon dioxide remains in the blood in the dissolved form. Within the red blood cell, carbon dioxide can bind with reduced hemoglobin to form carbamino hemoglobin. In addition, due to the presence of the enzyme, carbonic anhydrase (CA), carbon dioxide can combine with water to form carbonic acid (H_2CO_3). Carbonic acid then dissociates into a bicarbonate ion and a hydrogen ion (H^+). The hydrogen ion is picked up by the reduced hemoglobin (HHb). The bicarbonate ion diffuses down its concentration gradient from the red blood cell into the plasma. In order to maintain electrical neutrality, as the bicarbonate ion exits the cell, a chloride ion (Cl^-) enters the cell. This is referred to as the chloride shift. In the lungs, these reactions reverse direction and the carbon dioxide is eliminated by ventilation.

is active when the demands for ventilation are increased, such as during exercise. Under these conditions, action potentials in the *expiratory neurons* cause forced, or active, expiration. These neurons descend to the spinal cord where they stimulate the neurons that supply the expiratory muscles, including those that innervate the muscles of the abdominal wall and the internal intercostal muscles. Contractions of these muscles cause a more rapid and more forceful expiration.

Inspiratory neurons of the VRG augment inspiratory activity. These neurons descend to the spinal cord where they stimulate the neurons that supply the accessory muscles of inspiration, including those that innervate the scalene muscles and the sternocleidomastoid muscles. Contractions of these muscles cause a more forceful inspiration.

In summary, the regulation of ventilation by the medullary respiratory center determines the following:

- The interval between the successive groups of action potentials of the inspiratory neurons, which determines the rate or *frequency of breathing* (as the interval shortens, the breathing rate increases).
- The frequency of action potential generation and the duration of this electrical activity to the motor neurons and, therefore, the muscles of inspiration and expiration, which determines the depth of breathing, or the *tidal volume* (as the frequency and duration of stimulation increases, the tidal volume increases).

The medullary respiratory center receives excitatory and inhibitory inputs from many areas of the brain and the peripheral nervous system, including the following:

- Lung receptors
- Proprioceptors
- Pain receptors
- Limbic system
- Chemoreceptors

Pulmonary stretch receptors are responsible for initiating the *Hering-Breuer reflex*. These stretch receptors are located within the smooth muscle of both large and small airways. They are stimulated when the tidal volume exceeds 1 liter. Nerve impulses are transmitted by the vagus nerve to the medullary respiratory center and inhibit the inspiratory neurons. The primary function of these receptors and the Hering-Breuer reflex is to prevent overinflation of the lungs.

Irritant receptors are located throughout the respiratory system in the nasal mucosa, the upper airways, the tracheobronchial tree, and possibly the alveoli. Mechanical or chemical stimulation of these receptors can cause a reflex cough or sneeze. In either case, a deep inspiration is followed by a

forced expiration against a closed glottis, which causes a marked increase in intrapulmonary pressure. The glottis then opens suddenly, resulting in an explosive expiration. The ensuing high airflow rate is meant to eliminate the irritant from the respiratory tract. In a cough, expiration occurs by way of the mouth, and in a sneeze, expiration is through the nose. Stimulation of irritant receptors also causes hyperpnea (increased ventilation) and bronchoconstriction.

Proprioceptors are located in muscles, tendons, and joints. Stimulation of these receptors causes an increase in ventilation. Proprioceptors are believed to play a role in initiating and maintaining the elevated ventilation associated with exercise.

Pain receptors also influence the medullary respiratory center. Pain may cause a reflex increase in ventilation in the form of a "gasp." Somatic pain typically causes hyperpnea, and visceral pain typically causes apnea, or decreased ventilation.

Breathing is modified during the expression of various emotional states. For example, the normal cyclical pattern of breathing is interrupted during laughing, crying, sighing, or moaning. The modifications in ventilation associated with these activities are elicited by input from the *limbic system* to the medullary respiratory center.

Chemoreceptors provide the most important input to the medullary respiratory center in terms of regulating ventilation to meet the metabolic requirements of the body. Chemoreceptors are sensitive to changes in PO_2, PCO_2, and pH. There are two types of chemoreceptors: *peripheral chemoreceptors* and *central chemoreceptors*.

The *peripheral chemoreceptors* include the carotid bodies and the aortic bodies. The *carotid bodies*, which are more important in humans, are located near the bifurcation of the common carotid arteries. The *aortic bodies* are located in the arch of the aorta. The peripheral chemoreceptors respond to a decrease in PO_2, an increase in PCO_2, and a decrease in pH (increase in H^+ ion concentration) in the arterial blood.

The *central chemoreceptors* are located near the ventral surface of the medulla in close proximity to the respiratory center. These receptors are surrounded by the extracellular fluid (ECF) of the brain. They respond to changes in H^+ ion concentration. The composition of the ECF surrounding the central chemoreceptors is determined by the cerebrospinal fluid (CSF), local blood flow, and local metabolism.

A summary of the responses of the peripheral and the central chemoreceptors to reduced arterial oxygen, increased arterial carbon dioxide, and increased arterial hydrogen ion concentration is found in Table 19.4.

19.14.1 Chemoreceptor response to decreased arterial PO_2

Hypoxia has a direct depressant effect on the central chemoreceptors as well as the medullary respiratory center. In fact, hypoxia tends to inhibit activity

Table 19.4 Chemoreceptor Responses to Changes in Arterial PO_2, PCO_2, and H^+ Ion Concentration

Change in Arterial Blood	Effect on Peripheral Chemoreceptors	Effect on Central Chemoreceptors
↓ Arterial PO_2	Stimulates (PO_2 < 60 mmHg)	Depresses
↑ Arterial PCO_2	Weakly stimulates	Strongly stimulates
↑ Arterial H^+ ion	Stimulates	No effect

in all regions of the brain. Therefore, the ventilatory response to hypoxemia is elicited only by the peripheral chemoreceptors.

A decrease in arterial PO_2 causes stimulation of the peripheral chemoreceptors. The ensuing elevation in ventilation increases the uptake of oxygen and returns PO_2 back to its normal value of 100 mmHg. However, this stimulatory effect does not occur until the arterial PO_2 falls below 60 mmHg. The physiological explanation for this delayed response to hypoxemia is provided by the shape of the oxyhemoglobin dissociation curve (see Figure 19.7). The plateau portion of the curve illustrates that hemoglobin remains quite saturated (90%) at a PO_2 of 60 mmHg. Therefore, when the PO_2 of the arterial blood is between 60 mmHg and 100 mmHg, the oxygen content of the blood is still very high. An increase in ventilation at this point is not critical. However, below a PO_2 of 60 mmHg, the saturation of hemoglobin and, therefore, the oxygen content of the blood decrease rapidly. Stimulation of the peripheral chemoreceptors in order to increase ventilation and enhance the uptake of oxygen is now essential to meet the metabolic needs of the body. Let it be clear, however, that this ventilatory response to hypoxemia is due to the change in PO_2, not oxygen content. For example, as discussed previously, anemia and carbon monoxide poisoning, which decrease oxygen content but not PO_2, do not cause an increase in ventilation.

It is important to note that a decrease in PO_2 is not the primary factor in the minute-to-minute regulation of ventilation. This is because the peripheral chemoreceptors are not stimulated until the PO_2 falls to life-threatening levels. A decrease of this magnitude would likely be associated with abnormal conditions, such as pulmonary disease, hypoventilation, or ascent to extreme altitude.

19.14.2 Chemoreceptor response to increased arterial PCO_2

An increase in arterial PCO_2 causes weak stimulation of the peripheral chemoreceptors. The ensuing mild increase in ventilation contributes to the elimination of carbon dioxide and the decrease in PCO_2 back toward its normal value of 40 mmHg. The response of the peripheral chemoreceptors to changes in arterial PCO_2 is much less important than that of the central chemoreceptors. In fact, less than 20% of the ventilatory response to an increase in arterial PCO_2 is elicited by the peripheral chemoreceptors.

An increase in arterial PCO_2 results in the marked stimulation of the central chemoreceptors. In fact, this is the most important factor in the regulation of ventilation. As we are all aware, it is impossible to hold one's breath indefinitely. As the carbon dioxide accumulates in the arterial blood, the excitatory input to the respiratory center from the central chemoreceptors overrides the voluntary inhibitory input and breathing resumes. Furthermore, this occurs well before the arterial PO_2 falls low enough to stimulate the peripheral chemoreceptors.

Interestingly, the central chemoreceptors are insensitive to carbon dioxide. However, they are very sensitive to changes in the H^+ ion concentration in the ECF surrounding them. How does an increase in arterial PCO_2 cause an increase in H^+ ion concentration in the brain? Carbon dioxide readily crosses the blood–brain barrier. As the arterial PCO_2 increases, carbon dioxide diffuses down its concentration gradient into the ECF of the brain from the cerebral blood vessels. Due to the presence of the enzyme carbonic anhydrase, the following reaction takes place in the ECF of the brain:

$$CO_2 + H_2O \overset{CA}{\leftrightarrow} H_2CO_3 \leftrightarrow H^+ + HCO_3^-$$

Carbonic anhydrase (CA) facilitates the formation of carbonic acid (H_2CO_3) from carbon dioxide and water. The carbonic acid then dissociates to liberate hydrogen ion (H^+) and bicarbonate ion (HCO_3^-). The hydrogen ions strongly stimulate the central chemoreceptors to increase ventilation. The ensuing elimination of the excess carbon dioxide from the arterial blood returns the PCO_2 back to its normal value.

Conversely, a decrease in the arterial PCO_2 due to hyperventilation results in a decrease in the H^+ ion concentration in the ECF of the brain. Decreased stimulation of the central chemoreceptors and, therefore, a decrease in the excitatory input to the medullary respiratory center, cause a decrease in ventilation. Continued metabolism allows the carbon dioxide to accumulate in the blood such that the PCO_2 returns to its normal value.

19.14.3 Chemoreceptor response to increased arterial hydrogen ion concentration

An increase in arterial hydrogen ion concentration, or a decrease in arterial pH, stimulates the peripheral chemoreceptors and enhances ventilation. This response is important in maintaining acid–base balance. For example, under conditions of *metabolic acidosis*, caused by the accumulation of acids in the blood, the enhanced ventilation eliminates carbon dioxide and, thereby, reduces the concentration of the H^+ ions in the blood. Metabolic acidosis may occur in patients with uncontrolled diabetes mellitus or when tissues become hypoxic and produce lactic acid.

An increase in arterial hydrogen ion concentration has no effect on the central chemoreceptors. Hydrogen ions are unable to cross the blood–brain barrier.

19.15 Ventilatory response to exercise

Exercise results in an increase in oxygen consumption and an increase in carbon dioxide production by the working muscles. In order to meet the metabolic demands of these tissues, ventilation increases accordingly. Minute ventilation increases linearly in response to oxygen consumption and carbon dioxide production up to the level of approximately 60% of an individual's work capacity. During this period of mild to moderate exercise, mean arterial PO_2 and PCO_2 remain relatively constant at their normal values. In fact, the partial pressures of these gases may even improve (arterial PO_2 is increased, arterial PCO_2 is decreased). Therefore, it does not appear that hypoxic or hypercapnic stimulation of the peripheral chemoreceptors plays a role in the ventilatory response to mild to moderate exercise.

Beyond this point, during more severe exercise associated with anaerobic metabolism, the minute ventilation increases faster than the rate of oxygen consumption, but proportionally to the increase in carbon dioxide production. The mechanism of the ventilatory response to severe exercise involves the *metabolic acidosis* caused by anaerobic metabolism. The lactic acid produced under these conditions liberates the H^+ ion, which effectively stimulates the peripheral chemoreceptors to increase ventilation.

During exercise, the increase in minute ventilation results from increases in both tidal volume and breathing frequency. Initially, the increase in tidal volume is greater than the increase in breathing frequency. As discussed earlier in this chapter, increases in tidal volume increase alveolar ventilation more effectively. Subsequently, however, as metabolic acidosis develops, the increase in breathing frequency predominates.

The mechanisms involved with the ventilatory response to exercise remain quite unclear. No single factor, or combination of factors, can fully account for the increase in ventilation during exercise. Therefore, much of this response remains unexplained. Factors that appear to play a role include the following:

- Impulses from the cerebral cortex
- Impulses from proprioceptors
- Body temperature
- Epinephrine

At the beginning of exercise, there is an immediate increase in ventilation. This increase is thought to be caused by two mechanisms involving the cerebral cortex. Neurons of the primary motor cortex stimulate alpha motor neurons in the spinal cord to cause skeletal muscle contraction. In addition, *impulses from the motor cortex*, transmitted through collateral interconnections

to the medullary respiratory center, stimulate ventilation. The motor cortex is also involved in the stimulation of the cardiovascular system during exercise. These adjustments, which occur before any homeostatic factors (e.g., blood gases) have changed, are referred to as *anticipatory adjustments*. The immediate increase in ventilation may account for as much as 50% of the total ventilatory response to exercise. A *conditioned reflex*, or a learned response to exercise, may also be involved. Once again, impulses from the cerebral cortex provide input to the medullary respiratory center.

Proprioceptors originating in muscles and joints of the exercising limbs provide substantial input to the medullary respiratory center. In fact, even passive movement of the limbs causes an increase in ventilation. Therefore, the mechanical aspects of exercise also contribute to the ventilatory response.

The increased metabolism associated with exercise increases *body temperature*. The increase in body temperature further contributes to the increase in ventilation during exercise. (Not surprisingly, ventilation is also enhanced in response to a fever.)

Exercise is associated with a mass sympathetic discharge. As a result, epinephrine release from the adrenal medulla is markedly increased. *Epinephrine* is believed to stimulate ventilation.

Medical terminology

Aerobic (ĕr-ō′bĭk): Requiring oxygen.

Alveolar dead space (ăl-vē′ō-lăr dĕd spās): Volume of the poorly perfused alveoli; the resulting lack of gas exchange contributes to physiological dead space.

Anatomical dead space (ăn″ă-tŏm′ĭk-ăl dĕd spās): Volume of the conducting airways where there is no gas exchange.

Anemia (ă-nē′mē-ă): Condition resulting from a decrease in the number of circulating red blood cells or a decrease in hemoglobin.

Apnea (ăp-nē′ă): Temporary cessation of breathing.

Atelectasis (ăt″ĕ-lĕk′tă-sĭs): Collapse of the airways.

Bohr effect (bor ē-fĕkt′): Hydrogen ions or an acid environment changes the structure of hemoglobin causing it to release more oxygen in the tissues.

Chemoreceptor (kē″mō-rē-sĕp′tor): In the respiratory system, a receptor that is stimulated by a decrease in oxygen, an increase in carbon dioxide, or an increase in hydrogen ion resulting in an increase in ventilation.

Compliance (kŏm-plī′ăns): In the respiratory system, the distensibility of the lungs.

Dyspnea (dĭsp-nē′ă): Difficulty breathing.

Exogenous (ĕks-ŏj′ĕ-nŭs): Originating outside of the body or an organ.

Hering-Breuer reflex (hĕr′ĭng-broo′ĕr rē′flĕks): Reflex inhibition of inspiration due to stimulation of stretch receptors in the lungs.

Hypercapnia (hī″pĕr-kăp′nē-ă): Increased carbon dioxide in the blood.

Hyperpnea (hī″pĕrp′nē-ă): Increased ventilation due to an increase in tidal volume, an increase in breathing frequency, or both.
Hypoxemia (hī-pŏks-ē′mē-ă): Decreased oxygen in the arterial blood.
Insidious (ĭn-sĭd′ē-ŭs): Gradual, subtle, or indistinct in onset.
Intrapleural (ĭn″tră-ploo′răl): Within the pleural cavity surrounding the lungs.
Nebulizer (nĕb′ū-lī″zĕr): Apparatus that produces a fine mist for inhalation.
Neonate (nē′ō-nāt): Newborn infant up to 1 month of age.
Noncompliant (nŏn-kŏm-plī′ănt): In the respiratory system, stiff, poorly distensible lungs that are difficult to inflate.
Phrenic nerve (frĕn′ĭk nĕrv): Portion of the somatic motor nervous system that innervates the diaphragm.
Physiological dead space (fĭz″ē-ō-lŏj′ĭ-kăl dĕd spās): Sum of the anatomical dead space and the alveolar dead space; represents all of the regions of the lungs that do not participate in gas exchange and do not eliminate carbon dioxide.
Pleura (ploo′ră): Double-walled membrane that encloses each lung.
Pneumothorax (nū-mō-thō′răks): Entry of air into the pleural cavity causing collapse of the lung.
Proprioceptors (prō″prē-ō-sĕp′tors): Receptors in muscles and joints that provide an awareness of body position and movement.
Recoil (rē′koyl): Spring back; in the respiratory system, the return of the lungs to their preinspiratory volume.
Shunt (shŭnt): In the respiratory system, blood that perfuses unventilated alveoli.
Surfactant (sŭr-făk′tănt): In the respiratory system, a substance secreted by alveolar type II cells which decreases the surface tension of the fluid lining the alveoli.
Transpulmonary (trănz″pŭl′mō-nĕ-rē): Across the surface of the lungs.
Transudation (trănz″-ū-dā′shŭn): Movement of fluid across the wall of the capillary and into the interstitial space.

Bibliography

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.
Gutstein, H. B., and Akil, H., Opioid Analgesics, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics*, 9th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 21.
Levitsky, M. G., *Pulmonary Physiology*, 5th ed., McGraw-Hill, New York, 1999.
Marieb, E. N., and Hoehn, K., *Anatomy and Physiology,* 3rd ed., Pearson Benjamin Cummings, San Francisco, CA, 2008.
Porth, C. M., *Pathophysiology, Concepts of Altered Health States*, 7th ed., J. B. Lippincott, Philadelphia, PA, 2005.
Robergs, R. A., and Roberts, S. O., *Exercise Physiology, Exercise, Performance, and Clinical Applications,* Mosby, St. Louis, MO, 1997.
Schmann, L., Obstructive Pulmonary Disorders, in *Perspectives on Pathophysiology*, Copstead, L. -E. C., Ed., W.B. Saunders, Philadelphia, PA, 1995, chap. 20.

Schmann, L., Ventilation and Respiratory Failure, in *Perspectives on Pathophysiology*, Copstead, L. -E. C., Ed., W.B. Saunders, Philadelphia, PA, 1995, chap. 22.
Schwartzstein, R. M., and Parker, M. J., *Respiratory Physiology: A Clinical Approach,* Lippincott Williams & Wilkins, Philadelphia, PA 2006.
Sherwood, L., *Human Physiology, From Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.
Silverthorn, D. U., *Human Physiology, An Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.
Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Baltimore, MD, 2005.
Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.
Ward, J. P. T., Ward, J., Wiener, C. M., and Leach, R. M., *The Respiratory System at a Glance*, Blackwell Science, Oxford, UK, 2002.
West, J. B., *Pulmonary Physiology and Pathophysiology, An Integrated, Case-Based Approach*, Lippincott Williams & Wilkins, Philadelphia, PA, 2001.
West, J. B., *Respiratory Physiology, The Essentials,* 7th ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005.
Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.
www.healthinfotranslations.com/incentive_spirometer_441477.php.

chapter twenty

The digestive system

Study objectives

- Describe the anatomical and functional characteristics of each of the four layers of the digestive tract wall: mucosa, submucosa, muscularis externa, and serosa
- Distinguish between the two types of gastrointestinal motility: segmentation and peristalsis
- Explain how each of the three types of sensory receptors within the digestive tract are stimulated: chemoreceptors, osmoreceptors, and mechanoreceptors
- Explain how the following mechanisms regulate the activity of the digestive system: intrinsic nerve plexuses, extrinsic autonomic nerves, and gastrointestinal hormones
- List the components of saliva and their functions
- Describe how salivary secretion is regulated
- Explain how swallowing takes place
- For each of the following organs, esophagus, stomach, small intestine, and large intestine, describe:
  - Specialized anatomical modifications
  - The type of motility and how it is regulated
  - The types of secretions and how they are regulated
  - The digestive processes that take place
  - The absorptive processes that take place
  - Other functions

20.1 Introduction

The overall function of the digestive system is to make ingested food available to the cells of the body. Most ingested food is in the form of very large molecules that must be broken down by mechanical and biochemical processes into their smaller components (see Table 20.1). These smaller units are then absorbed across the wall of the digestive tract and distributed throughout the body. Not all ingested materials may be completely digested and absorbed by the human gastrointestinal tract. For example, cellulose, the fibrous form of plant carbohydrates, is indigestible by humans. Normally, about 95% of

Table 20.1 Ingested and Absorbable Molecules for the Three Major Categories of Nutrients

Ingested Form of Nutrient Molecules	Absorbable Form of Nutrient Molecules
Carbohydrates	Monosaccharides
• Polysaccharides (Starch)	• Glucose
• Disaccharides	• Galactose
• Sucrose (Table Sugar)	• Fructose
• Lactose (Milk Sugar)	
Proteins	Amino Acids
Fats	Monoglycerides
• Triglycerides	Free Fatty Acids

ingested food materials are made available for use by the body. Interestingly, as long as food remains within the digestive tract, it is technically outside of the body. Not until the materials have crossed the epithelium that lines the tract are they considered to have "entered" the body.

The digestive system consists of the gastrointestinal tract and the accessory digestive organs.

The *gastrointestinal tract* is essentially a tube that runs through the center of the body from the mouth to the anus. This tube consists of the following organs:

- Mouth
- Pharynx
- Esophagus
- Stomach
- Small intestine
- Large intestine

Although these organs are continuous with one other, each has important anatomical modifications that allow them to carry out their specific functions.

The *accessory digestive organs* exist outside of the gastrointestinal tract; however, each of these organs empties secretions into the tract which contribute to the process of digestion. These accessory digestive organs include the following:

- Salivary glands
- Liver
- Gallbladder
- Pancreas

20.2 The digestive tract wall

The digestive tract wall has the same basic structure from the esophagus through and including the colon. There are four major layers within the wall:

1. Mucosa
2. Submucosa
3. Muscularis externa
4. Serosa

20.2.1 Mucosa

The innermost layer of the wall is the *mucosa*. It consists of a mucous membrane, the lamina propria, and the muscularis mucosa. The *mucous membrane* provides important protective, absorptive, and secretory functions for the digestive tract. The nature of the epithelial cells lining the tract varies from one region to the next. Rapidly dividing stem cells continually produce new cells to replace worn out epithelial cells. The average life span of these epithelial cells is only a few days. Goblet cells, which secrete mucus, are found in the mucosa throughout much of the gastrointestinal tract. The *lamina propria* is a thin middle layer of connective tissue. This region contains the capillaries and small lymphatic vessels that take up the digested nutrient molecules. It also contains numerous lymph nodules that provide protection against infection by microbes. The *muscularis mucosa* is a thin layer of smooth muscle. Contraction of this muscle may alter the effective surface area for absorption in the lumen.

20.2.2 Submucosa

The *submucosa* is a thick middle layer of connective tissue. This tissue provides the digestive tract wall with its distensibility and elasticity as nutrient materials move through the system.

20.2.3 Muscularis externa

The outer layer of the wall is the *muscularis externa*. In most regions of the tract, it consists of two layers of muscle: an inner circular layer and an outer longitudinal layer. Contraction of the circular layer narrows the lumen of the tube. Contraction of the longitudinal layer causes the tube to shorten.

The muscle of the digestive tract consists of *single-unit smooth muscle*. Within each layer, the muscle cells are connected by gap junctions forming a syncytium. In this way, action potentials generated at a given site may travel along the muscle layer. The smooth muscle undergoes slow but continuous electrical activity producing rhythmic contractions of the digestive tract wall. The cycles of depolarization and repolarization in the smooth muscle are referred

to as *slow-wave potentials*. Interestingly, recent studies have suggested that the generation of the slow waves in gastrointestinal smooth muscle actually involves specialized cells referred to as *interstitial cells of Cajal* (*ICC*). These cells have long processes as well as gap junctions. As a result, the ICCs may communicate with each other as well as smooth muscle cells, which permits the spread of depolarization from one cell to the next.

Slow-wave potentials do not reach threshold during each cycle so that contraction does not necessarily occur with each depolarization. Smooth muscle contraction will take place only when the slow wave actually depolarizes all of the way to threshold. At this point, voltage-gated Ca^{++} channels open, Ca^{++} ions enter the muscle cells, and one or more action potentials are generated. The influx of Ca^{++} ions contributes to both the depolarization phase of the action potential as well as the muscle contraction. These action potentials result in *phasic contractions*. The force and duration of muscle contraction are determined by the number of action potentials generated. Typically, phasic contractions last only a few seconds. Tissues that exhibit phasic contractions include the body of the esophagus, the gastric antrum, and the small intestine. Other tissues, including the lower esophageal sphincter, the upper region of the stomach, and many sphincters located throughout the gastrointestinal tract, exhibit *tonic contractions.* These contractions are slow and sustained and may last minutes to hours.

Muscular activity, or *gastrointestinal motility*, is enhanced by stretching the muscle, as occurs with the presence of food materials and distension of the digestive tract wall. It is also enhanced by parasympathetic nervous stimulation. Motility is inhibited by sympathetic nervous stimulation, circulating epinephrine and by several specific gastrointestinal hormones.

There are two basic forms of gastrointestinal motility: segmentation and *peristalsis*. The contents of the digestive tract must be constantly churned and mixed. In this way, the materials are exposed to digestive enzymes, and they come into contact with the wall of the tract for absorption. This mixing is carried out by *segmentation*, or stationary muscular contractions. This form of motility divides some portion of the tract into alternating constricted regions and unconstricted regions. Segmentation contractions move back and forth so that a previously constricted region relaxes and a previously relaxed region contracts. This activity results in the thorough mixing of the contents with digestive enzymes and other secretions. This is the more important form of motility in the small intestine where most digestion and absorption take place.

The contents of the tract must also be continually moved along so that they can be acted upon by the sequential regions of the tract. *Peristalsis* is a muscular contraction that produces a ring of contraction that moves along the length of the tract. This wave-like contraction causes propulsion and

forces the contents forward. Peristalsis is more important in the pharynx, the esophagus, and the stomach.

Gastrointestinal sphincters are formed where the circular layer of smooth muscle is thickened. Sphincters occur at several points along the tract. Their function is to limit the movement of food materials from one region to another. For example, the pyloric sphincter is found between the stomach and the duodenum of the small intestine. This sphincter plays an important role in limiting the rate of gastric emptying. Sphincters undergo *tonic contractions* that may be sustained for minutes or hours.

20.2.4 Serosa

The connective tissue membrane that surrounds the wall of the digestive tract is the *serosa*. This membrane secretes a watery fluid that provides lubrication and prevents friction between the digestive organs as they move about in the abdomen. The serosa is continuous with the *peritoneum,* which is the serous membrane lining the abdominal cavity. The peritoneum also forms sheets of tissue, or *mesentery*, which suspend the digestive organs from the wall of the abdomen. The mesentery acts as a sling which, while offering structural support for the organs, also provides for the range of movement needed during the digestive process.

20.3 Regulation of gastrointestinal function

The digestive tract contains three types of sensory receptors that are sensitive to chemical or mechanical changes within the system:

1. Chemoreceptors
2. Osmoreceptors
3. Mechanoreceptors

Chemoreceptors respond to various chemical components within the gastrointestinal lumen. For example, chemoreceptors in the duodenum of the small intestine are stimulated by excessive amounts of hydrogen ion secreted by the stomach. *Osmoreceptors* are sensitive to the osmolarity of the contents within the lumen. As the digestive process progresses, large nutrient molecules are split into their smaller components. This increases the number of molecules and, therefore, increases the osmolarity of the material being processed. Excessive osmolarity may suggest that absorption is not keeping pace with digestion. *Mechanoreceptors* respond to stretch or distension of the gastrointestinal tract wall.

Receptor stimulation may lead to the activation of any or all of the following regulatory mechanisms within the tract:

- Enteric nervous system
  - Intrinsic nerve plexuses
  - Extrinsic autonomic nerves
- Gastrointestinal hormones

The *enteric nervous system* consists of submucosal (Meissner's) and myenteric (Auerbach's) plexuses. These plexuses within the wall of the intestine contain 100 million neurons (about the same number of neurons as found in the spinal cord). The enteric nervous system includes both intrinsic neurons and extrinsic neurons.

20.3.1 Intrinsic nerve plexuses

The *intrinsic nerve plexuses* are interconnecting networks of nerve cells located entirely within the gastrointestinal tract. These plexuses are responsible for *intratract reflexes*. The stimulation of a receptor in one region of the tract neurally influences the activity of another region of the tract. These reflexes occur directly, independent of the central nervous system. Intratract reflexes provide a mechanism for self-regulation of the tract and help to coordinate the activity of the organs within it. Examples of such reflexes include the following:

- *Enterogastric reflex*: Where receptor stimulation in the duodenum of the small intestine elicits neural activity that regulates both muscle contraction and glandular secretion in the stomach.
- *Gastroileal reflex*: Where increased gastric activity causes increased activity in the ileum and increased movement of chyme through the ileocecal sphincter.
- *Ileogastric reflex*: Where distension of the ileum causes a decrease in gastric motility.
- *Intestino-intestinal reflex*: Where overdistension of a given segment of the intestine causes relaxation throughout the rest of the intestine.

20.3.2 Extrinsic autonomic nerves

Gastrointestinal activity is also modified by *extrinsic autonomic nerves*. The tract is innervated by both the parasympathetic and the sympathetic divisions of the autonomic nervous system. Parasympathetic innervation is provided primarily by the vagus nerves (esophagus, stomach, pancreas, gallbladder, small intestine, and upper large intestine) and the pelvic nerves (the rest of the large intestine). Sympathetic innervation is provided by pathways that pass through the celiac, inferior mesenteric, and superior mesenteric ganglia. The effects of these two divisions tend to oppose each other. The parasympathetic system stimulates most digestive activities, and the sympathetic system inhibits them. Interestingly, the autonomic nerves to

the digestive system, especially the vagus nerve of the parasympathetic system, can be discretely activated. In this way, digestive activity can be modified without affecting tissue function in other regions of the body.

20.3.3 Gastrointestinal hormones

A third factor contributing to the regulation of digestive activity is the secretion of *gastrointestinal hormones*. These hormones may be released in one region of the tract, travel in the circulatory system to other regions of the tract, and influence the activity of effector cells in that region. A summary of the source, the stimulus for release, and the actions of several important hormones is found in Table 20.2.

A summary of these three mechanisms that regulate the activity of the digestive system is illustrated in Figure 20.1. A local change in the tract may lead to the stimulation of one or more of the three types of receptors present in the tract wall. Receptor stimulation may then activate any or all of the three regulatory mechanisms. These mechanisms then alter the activity of the effector tissues within the digestive system, including smooth muscle, exocrine glands, and endocrine glands.

The following sections will discuss each region of the digestive system separately. Where appropriate, the basic digestive processes will be considered: motility, secretion, digestion, and absorption.

20.4 Mouth

The *mouth* is the region from the lips to the pharynx. The first step in the digestive process is chewing or *mastication*. This initial mechanical breakdown of the food facilitates its movement to the stomach. The mouth is lined with *stratified squamous epithelium* that provides extra protection from injury by coarse food materials.

Three pairs of salivary glands secrete saliva into the oral cavity:

1. *Parotid Glands*: Located between the angle of the jaw and the ear.
2. *Sublingual Glands*: Located below the tongue.
3. *Submandibular Glands*: Located below the jaw.

Saliva contains the following:

- Water
- Mucus
- Lysozyme
- Salivary amylase
- Lingual lipase

Table 20.2 Digestive Hormones

Hormone	Source	Stimulus for Release	Actions of Hormone
Gastrin	G cells in pyloric region of the stomach	Protein in stomach; vagal stimulation	Stimulates parietal cells (HCl) and chief cells (pepsinogen) in the stomach; enhances gastric motility
Secretin	Endocrine cells in mucosa of duodenum	Acid in duodenum	Inhibits gastric emptying; inhibits gastric secretion; stimulates secretion of bicarbonate from the pancreas; stimulates secretion of bicarbonate-rich bile from the liver
Cholecystokinin	Endocrine cells in mucosa of duodenum	Breakdown products of lipid and, to a small extent, protein digestion in duodenum	Inhibits gastric emptying Inhibits gastric secretion Stimulates contraction of the gallbladder Stimulates secretion of digestive enzymes from the pancreas
Glucose-dependent insulinotropic peptide (GIP); formerly referred to as Gastric Inhibitory Peptide	Endocrine cells in mucosa of duodenum	Glucose, lipids, acid, and hyperosmotic chyme in duodenum; distension of duodenum	Inhibits gastric emptying; inhibits gastric secretion; stimulates secretion of insulin from pancreas

Approximately 99.5% of saliva is *water*. Swallowing is facilitated by the moistening of food materials. Furthermore, water serves as a solvent for molecules that stimulate the taste buds. The presence of *mucus*, which is thick and slippery, lubricates the mouth and the food and assists in swallowing. *Lysozyme* is an enzyme that lyses or kills many types of bacteria that may be ingested with the food.

Saliva begins the process of chemical digestion with *salivary amylase*. This enzyme splits starch molecules into fragments. Specifically, polysaccharides, or starches, are broken down into maltose, a disaccharide consisting of two glucose molecules. Salivary amylase may account for up to 75% of starch digestion before it is denatured by hydrochloric acid in the stomach.

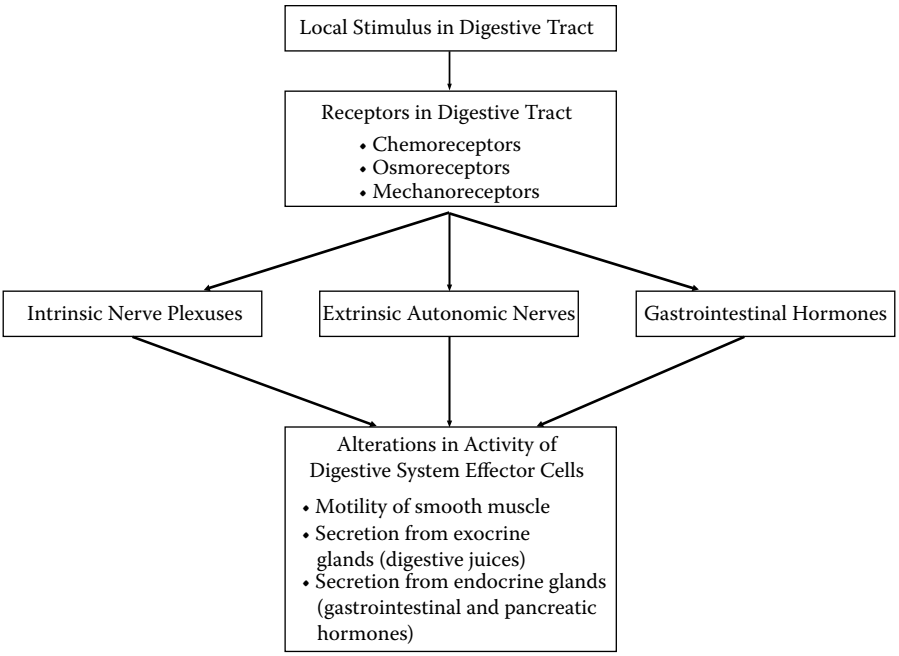


Figure 20.1 Summary of the regulatory mechanisms influencing gastrointestinal function.

A small amount of *lingual lipase* is also present and plays a role in the breakdown of dietary lipid. This enzyme is optimally active at an acidic pH and, therefore, remains active through the stomach and into the intestine.

Due to parasympathetic stimulation of the salivary glands, saliva is secreted continuously at a basal rate of approximately 0.5 ml/min. Secretion may be enhanced by two types of reflexes:

1. Simple or unconditioned salivary reflex
2. Acquired or conditioned salivary reflex

The *simple or unconditioned salivary reflex* occurs when food is present within the oral cavity causing stimulation of chemoreceptors and pressure receptors. These receptors then transmit impulses to the *salivary center* in the medulla of the brainstem. Parasympathetic efferent impulses are transmitted back to the salivary glands and secretion is enhanced.

The *acquired or conditioned salivary reflex* is elicited in response to the thought, sight, smell, or sound of food. As demonstrated with Pavlov's dog, these stimuli result in a learned response. Another stimulus that enhances salivation is nausea. Salivary secretion is inhibited by fatigue, sleep, fear, and dehydration. Overall, 1 to 2 liters of saliva may be produced per day.

PHARMACY APPLICATION: EFFECTS OF ANTICHOLINERGIC DRUGS ON THE DIGESTIVE SYSTEM

In addition to their therapeutic actions, many drugs have undesirable side effects that may influence the digestive system. An example of such a drug is scopolamine, one of the most effective agents used for the prevention of motion sickness. This drug may be administered transdermally in a multilayered adhesive unit, or "patch" form. Its mechanism of action likely involves the inhibition of muscarinic receptors in the vestibular apparatus of the inner ear. This interrupts the transmission of nerve impulses from the inner ear to the emetic center in the medulla of the brainstem. As a result, vomiting in response to motion is inhibited. However, the salivary glands are also quite sensitive to the activity of muscarinic receptor antagonists. In fact, scopolamine and other anticholinergic agents may severely inhibit the copious, watery secretion of the salivary glands. In this case, the mouth becomes dry, and swallowing and speaking may become difficult. Other anticholinergic agents may be used to (1) reduce muscle rigidity and muscle tremor in Parkinson's disease (benztropine mesylate); (2) reduce bronchospasm and airway mucus secretion in asthma and chronic obstructive pulmonary disease (COPD) (ipratropium); and (3) reduce the accumulation of secretions in the trachea and the possibility of laryngospasm prior to the administration of general anesthesia. In each case, these medications also inhibit salivary secretion and cause "dry mouth."

20.5 Pharynx

The *pharynx* is the cavity at the rear of the throat and links the mouth with the esophagus. It serves as a common passageway for both the respiratory and the digestive systems. The *swallowing reflex* takes place largely in the pharynx. This is an example of an all-or-none reflex where, once the process has begun, it cannot be stopped. Swallowing may be initiated voluntarily when the tongue pushes a bolus of food toward the back of the mouth and into the pharynx. The stimulation of pressure receptors in the pharynx results in the transmission of nerve impulses to the *swallowing center* in the medulla of the brainstem. This elicits a coordinated, involuntary reflex that involves the contraction of muscles in the appropriate sequence. A wave of contraction sweeps down the constrictor muscles of the pharynx. The epiglottis moves downward over the larynx to seal off the trachea, and the upper esophageal sphincter relaxes, allowing the bolus of food to enter the esophagus. Once the food bolus enters the esophagus, the upper esophageal sphincter closes in order to prevent the swallowing of air. This phase of the swallowing reflex is referred to as the *pharyngeal stage* and lasts approximately 1 second.

20.6 Esophagus

The *esophagus*, located behind the trachea, is a muscular tube connecting the pharynx and the stomach. It is lined with *stratified squamous epithelium*. The only substance secreted by the esophagus is *mucus*. The protective mucus provides lubrication for the passage of food, and it helps to prevent damage to the esophageal wall by coarse food materials. The esophagus is sealed off by two sphincters, one at either end of the tube: the *upper esophageal sphincter* (UES) and the *lower esophageal sphincter* (LES). Each of these sphincters is normally closed except during the process of swallowing. The normal function of the respiratory system creates a subatmospheric pressure in the thoracic cavity. If, indeed, the esophagus were open to the atmosphere, this pressure gradient would pull air into the esophagus and stomach during each inspiration. Therefore, the closure of these sphincters prevents large volumes of air from entering the digestive tract. In addition, the LES prevents the reflux of the corrosive gastric contents into the esophagus.

The *esophageal stage* of the swallowing reflex involves a *primary peristaltic wave* of contraction that is initiated by the swallowing center and mediated by the vagus nerve. This wave that begins at the UES moves slowly down the esophagus at a rate of 2 to 6 cm/sec until it reaches the LES. Therefore, it may take a bolus of food 10 seconds to pass through the esophagus. Some food particles that are particularly large or sticky may remain in the esophagus after the primary peristaltic wave. The distension of the esophagus by the presence of these particles elicits *secondary peristaltic waves* that do not involve the swallowing center. The smooth muscle of the LES relaxes immediately prior to the arrival of the peristaltic contraction to allow for the movement of the food into the stomach.

20.7 Stomach

The *stomach*, located on the left side of the abdominal cavity just below the diaphragm, lies between the esophagus and the small intestine. It is the most distensible portion of the gastrointestinal tract. As with the esophagus, it has a sphincter at either end; the previously mentioned LES is located at the entrance to the stomach, and the *pyloric sphincter* is located at the exit of the stomach leading to the duodenum of the small intestine. The LES is normally closed except during swallowing. The pyloric sphincter is subject to tonic contraction that keeps it almost, but not completely, closed. In this way, fluids may easily pass through it. The movement of food materials through this sphincter requires strong gastric contractions. Even then, only a few milliliters are pushed through at a time. Gastric contractions mash the food materials and thoroughly mix them with the gastric secretions. This produces a thick, semifluid mixture referred to as *chyme*.

The stomach is divided into three regions:

1. *Fundus*: Uppermost region of the stomach located above the junction with the esophagus.
2. *Body*: Middle or main portion of the stomach.
3. *Antrum*: Terminal region of the stomach leading to the gastroduodenal junction.

The stomach performs several important functions:

- Stores ingested food until it can be processed by the remainder of the digestive tract.
- Mechanically mashes ingested food and mixes it with gastric secretions.
- Begins the process of protein digestion.

Food is stored in the body of the stomach, which may expand to hold as much as 1 liter of chyme. As food enters the stomach, it undergoes reflex relaxation. This reflex is referred to as *receptive relaxation*. It enhances the ability of the stomach to accommodate a marked increase in volume with only a small increase in stomach pressure. The fundus does not typically store food, as it is located above the esophageal opening into the stomach. Instead, it usually contains a pocket of gas.

20.7.1 Gastric motility

In addition to the circular and longitudinal layers of smooth muscle, there is an extra layer of smooth muscle in the stomach. The *oblique layer* of smooth muscle begins at the UES and fans out across the anterior and posterior surfaces of the stomach. It fuses with the circular layer in the lower region of the stomach. This extra layer of muscle enhances gastric motility and, therefore, the mixing and mashing of the food.

Contraction of gastric smooth muscle occurs in the form of *peristalsis*. Peristaltic contractions begin in the body of the stomach and proceed in a wave-like fashion toward the duodenum. These contractions are weak in the upper portion of the stomach where the muscle layers are relatively thin. The contractions become much stronger in the lower portion of the stomach as the muscle layers become thicker. As the wave of contraction sweeps through the antrum, a small amount of chyme is pushed through the partially open pyloric sphincter. When the peristaltic contraction actually reaches the pyloric sphincter, the sphincter closes, and the rest of the chyme in this region is forced back toward the body of the stomach where more mixing and mashing takes place. This propulsion of chyme back into the stomach is referred to as *retropropulsion.*

It may take many hours for the contents of the stomach to be processed and moved into the small intestine. Several factors influence gastric motility and, therefore, the rate of gastric emptying:

- Volume of chyme in the stomach.
- Fluidity of the chyme.
- Volume and chemical composition of the chyme in the duodenum.

The major gastric factor that affects motility and the rate of emptying is the *volume of chyme in the stomach*. As the volume of chyme increases, the wall of the stomach becomes distended and mechanoreceptors are stimulated. This elicits reflexes that enhance gastric motility by way of the intrinsic nerves and the vagus nerve. The release of the hormone gastrin from the antral region of the stomach further contributes to enhanced motility.

The degree of *fluidity of the chyme* also affects the rate of gastric emptying. Ingested liquids move through the pyloric sphincter and begin to empty almost immediately. Ingested solids must first be converted into a semifluid mixture of uniformly small particles. Only particles about 1 mm^3 or smaller move readily into the duodenum. The faster the necessary degree of fluidity is achieved, the more rapidly the contents of the stomach may empty into the duodenum.

The most important factors that regulate gastric motility and the rate of emptying of the stomach involve the *volume and chemical composition of the chyme in the duodenum*. Receptors in the duodenum are sensitive to the following:

- Distension
- Lipids
- Acid
- Osmolarity of the chyme

The ultimate goal of these duodenal factors is to maintain a rate of gastric emptying that is consistent with the proper digestion and absorption of nutrient molecules in the small intestine. In other words, emptying must be regulated so that the duodenum has adequate opportunity to process the chyme that it already contains before it receives more from the stomach. Regulation occurs by way of the *enterogastric reflex* that inhibits gastric motility, increases contraction of the pyloric sphincter, and, therefore, decreases the rate of gastric emptying. This reflex is mediated through the intrinsic nerves and the vagus nerve. Regulation also occurs by way of a *hormonal response* that involves the release of the *enterogastrones* from the duodenum. These hormones include *secretin*, *cholecystokinin*, and *glucose-dependent insulinotropic peptide* (GIP).

As the volume of the chyme in the duodenum increases, it causes *distension* of the duodenal wall and the stimulation of *mechanoreceptors*. This receptor stimulation elicits reflex inhibition of gastric motility mediated through the intrinsic nerves and the vagus nerve. Distension also causes the release of GIP from the duodenum, which contributes to the inhibition of gastric contractions.

Duodenal receptors are also sensitive to the chemical composition of the chyme and are able to detect the presence of lipids, excess hydrogen ion, and hyperosmotic chyme. These conditions also elicit the enterogastric reflex and the release of the enterogastrones in order to decrease the rate of gastric emptying.

Of the three major categories of nutrients, *lipids* are the slowest to be digested and absorbed. Furthermore, these processes take place only in the small intestine. Therefore, in order to ensure complete lipid digestion and absorption, the rate of movement of lipid from the stomach to the duodenum must be carefully regulated. The presence of lipid in the duodenum stimulates intestinal chemoreceptors. This receptor stimulation elicits reflex inhibition of gastric motility and slows the addition of more lipids from the stomach. Lipid also causes the release of cholecystokinin from the duodenum. This hormone contributes to the inhibition of gastric contractions. The significance of the inhibitory effect of lipid is illustrated by the comparison between a high-fat meal (up to 6 hours for gastric emptying) and a meal consisting of carbohydrates and protein (3 hours for gastric emptying). Therefore, a fatty meal is "more filling" than a low-fat meal due its effect on gastric motility.

An important gastric secretion is *hydrochloric acid*, which performs a number of functions in the stomach. This acid from the stomach is neutralized by pancreatic bicarbonate ion in the duodenum. Excess acid in the chyme stimulates chemoreceptors in the duodenum. This receptor stimulation elicits reflex inhibition of gastric motility. Excess acid also causes the release of secretin from the duodenum. This hormone contributes to the inhibition of gastric contractions. In this way, the neutralization process may be completed before additional acid arrives in the chyme from the stomach.

Chyme within the duodenum has, by this point, undergone some degree of carbohydrate and protein digestion. Salivary amylase has fragmented starch molecules and, as will be discussed, pepsin from the stomach has fragmented proteins. Therefore, the number of disaccharides and small peptides has increased, which leads to an increase in the *osmolarity of the chyme*. The rate of absorption of these smaller molecules must keep pace with the rate of digestion of the larger molecules. If not, the stimulation of *osmoreceptors* in the duodenum by the hyperosmotic chyme will inhibit gastric motility and gastric emptying. This effect is mediated through reflex inhibition.

20.7.2 Gastric secretion

The human stomach secretes 2 to 4 liters of gastric juice per day. The gastric mucosa, which produces these secretions, is divided into two functional regions: the *oxyntic gland area* and the *pyloric gland area*.

The *oxyntic gland area* is located in the proximal 80% of the stomach. These glands consist of three types of cells:

1. Mucous neck cells
2. Parietal cells
3. Chief cells

The *pyloric gland area* is located in the remaining distal 20% of the stomach.

In addition to a large amount of water, secretions of the stomach include the following:

- Hydrochloric acid
- Pepsinogen
- Mucus
- Intrinsic factor
- Gastrin

Hydrochloric acid (HCl) is produced by the *parietal cells*. This is a strong acid that dissociates into H^+ ion and Cl^- ion. These ions are actively transported into the lumen of the stomach by the *proton pump* (H^+/K^+ ATPase). These pumps transport H^+ ions uphill against a million-to-one concentration gradient into the lumen of the stomach while they transport K^+ ions in the opposite direction. Functions of HCl include the following:

- Activating pepsinogen, the precursor for the enzyme, pepsin.
- Assisting in the breakdown of connective tissue and muscle fibers within the ingested food.
- Killing of most types of microorganisms ingested with the food.

Pepsinogen is produced by the *chief cells*. Within the lumen of the stomach, this precursor molecule is split by HCl to form the active enzyme *pepsin*. Optimally active at an acidic pH (pH = 2), pepsin begins protein digestion by fragmenting the proteins into smaller peptide chains.

Mucus is produced by the *mucus neck cells* and by the *surface epithelial cells* of the stomach wall. A thick layer of mucus adheres to the wall of the stomach, forming the *gastric mucosal barrier*. The function of this barrier is to protect the gastric mucosa from injury, specifically, from the corrosive actions of HCl and pepsin. Together with bicarbonate ion released into the lumen of the stomach, mucus neutralizes the acid and maintains the mucosal surface at a nearly neutral pH.

PHARMACY APPLICATION:
DRUG-INDUCED GASTRIC DISEASE

In addition to their beneficial effects, some medications may actually cause cellular injury and disease. An example of this phenomenon involves nonsteroidal antiinflammatory drugs (NSAIDS). These drugs include aspirin (a derivative of salicylic acid), ibuprofen (arylpropionic acid; Advil), and acetaminophen (para-aminophenol derivative; Tylenol®). Because of their beneficial pharmacological effects, the consumption of these agents has increased significantly in recent years. NSAIDS have the ability to treat fever, pain, acute inflammation, and chronic inflammatory diseases such as arthritis. They are also used prophylactically to prevent heart disease, stroke, and colon cancer.

Unfortunately, frequent exposure to NSAIDS may also cause two detrimental effects. These agents inhibit the activity of cyclooxygenase, an important enzyme in the synthesis of gastroprotective prostaglandins. More importantly, NSAIDS may cause breaks in the gastric mucosal barrier. The normal gastric mucosa is relatively impermeable to H^+ ion. When the gastric mucosal barrier is weakened or damaged, H^+ ion leaks into the mucosa in exchange for Na^+ ion. As H^+ ion accumulates in the mucosa, intracellular buffer systems become saturated, the pH decreases, and cell injury and cell death occur. These damaged cells then secrete more HCl, which causes more injury, and so on, resulting in a positive feedback cycle. An ulcer may form when injury from the gastric secretions, HCl and pepsin, overwhelms the ability of the mucosa to protect itself and replace damaged cells. Local capillaries are also damaged, causing bleeding or hemorrhage into the gastric lumen.

Intrinsic factor is produced by the *parietal cells*. Within the stomach, it combines with *vitamin* B_{12} forming a complex that is necessary for the absorption of this vitamin in the ileum of the small intestine. Vitamin B_{12} is an essential factor in the formation of red blood cells. Individuals unable to produce intrinsic factor cannot absorb vitamin B_{12} and, therefore, red blood cell production is impaired. This condition, referred to as *pernicious anemia*, occurs as the result of an autoimmune disorder that involves the destruction of parietal cells.

Gastrin is a hormone produced by gastric endocrine tissue, specifically, the *G cells* in the pyloric gland area. It is released into the blood and carried back to the stomach. The major function of gastrin is to enhance acid secretion by directly stimulating both parietal cells (HCl) and chief cells (pepsinogen). Gastrin also stimulates the local release of histamine from *enterochromaffin-like cells* in the wall of the stomach. *Histamine* stimulates parietal cells to release HCl.

There are three major phases of gastric secretion:

1. *Cephalic phase*: 20% to 30% of the gastric secretory response to a meal.
2. *Gastric phase*: 60% to 70% of the gastric secretory response to a meal.
3. *Intestinal phase*: Approximately 10% of the gastric secretory response to a meal.

The *cephalic phase* of gastric secretion occurs before food even enters the stomach. Thoughts of food; sensory stimuli, such as the smell, sight, or taste of food; and activities, such as chewing and swallowing, all enhance gastric secretion. The cephalic phase is mediated by the vagus nerve and gastrin, which is released in response to vagal stimulation. These mechanisms promote the secretion of HCl and pepsinogen.

The *gastric phase* is elicited by the presence of food in the stomach. Distension of the stomach wall as well as the presence of protein, caffeine, and alcohol all enhance gastric secretion. This phase is mediated by the intrinsic nerves, the vagus nerve, and gastrin. Each of these mechanisms promotes the secretion of HCl and pepsinogen. Other factors that influence gastric acid secretion include histamine (enhances) and somatostatin (inhibits).

PHARMACY APPLICATION:
PHARMACOLOGICAL TREATMENT OF GASTRIC ULCERS

The pharmacological treatment of ulcers involves the inhibition of gastric acid secretion. However, more than one approach may be used to accomplish this goal: H_2-receptor antagonists and proton pump inhibitors.

Histamine does not play a role in normal acid production. However, histamine may stimulate the release of HCl under pathological conditions. In the case of an ulcer, when H^+ ion enters the gastric mucosa, it stimulates the release of histamine from enterochromaffin-like cells. The histamine then stimulates H_2-receptors on the parietal cells to release more HCl. Therefore, excess acid release may be prevented with the administration of H_2-receptor antagonists, such as cimetidine (Tagamet®) and famotidine (Pepcid®). However, the inhibition of histamine-induced acid secretion is not adequate in all patients. More recently, proton pump inhibitors, such as omeprazole (Prilosec®) and lansoprazole (Prevacid®) have been used in the treatment of ulcers and gastroesophageal reflux disease (GERD). These drugs bind irreversibly to the proton pump (H^+, K^+-ATPase), which is found only in the parietal cell. This causes permanent inhibition of enzyme activity. As a result, the secretion of H^+ ions into the lumen of the stomach is inhibited. The secretion of acid resumes only after new molecules of H^+, K^+-ATPase are inserted into the gastric mucosa.

The *intestinal phase* has two components that influence gastric secretion: the *excitatory component* and the *inhibitory component*.

The *excitatory component* involves the release of *intestinal gastrin*. This occurs in response to the presence of the products of protein digestion in the duodenum. Intestinal gastrin travels in the blood to the stomach where it enhances the secretion of HCl and pepsinogen. The magnitude of this effect is very small, however, as it accounts for approximately 10% of the acid secretory response to a meal.

In contrast to the excitatory component, the *inhibitory component* of the intestinal phase has a very strong influence on gastric secretion. As with gastric motility, the volume and composition of the chyme in the duodenum affect gastric secretion. Distension of the duodenal wall as well as the presence of lipids, acid, and hyperosmotic chyme all inhibit secretion by way of the enterogastric reflex and the release of the enterogastrones.

20.8 Liver

The *liver* is the largest internal organ, weighing more than 1.5 kg (3.5 to 4.0 lbs) in the adult. The blood flow to the liver is 1350 ml/min (27% of the cardiac output) on average and comes from two sources:

1. Hepatic artery
2. Hepatic portal vein

The *hepatic artery* supplies the liver with 300 ml/min of oxygenated blood from the aorta. The remaining 1050 ml/min of blood flow is delivered by the *hepatic portal vein*. This blood comes directly from the digestive tract. It is low in oxygen but contains a high concentration of nutrients absorbed from the intestines.

The liver performs many important functions:

- Storage of blood
- Filtration of blood
- Storage of vitamins and iron
- Formation of blood coagulation factors
- Metabolism and excretion of certain drugs, bilirubin, and hormones
- Metabolism of carbohydrates, proteins, and lipids
- Formation of bile

The liver is a large and distensible organ. As such, large quantities of blood may be stored in its blood vessels providing a *blood reservoir* function. Under normal physiological conditions, the hepatic veins and hepatic sinuses contain approximately 450 ml of blood, or almost 10% of the blood volume. When needed, a significant portion of this blood may be mobilized to increase venous return and cardiac output.

Blood flowing from the intestines to the liver through the hepatic portal vein often contains bacteria. *Filtration of this blood* is a protective function provided by the liver. Large phagocytic macrophages, referred to as *Kupffer cells*, line the hepatic venous sinuses. As the blood flows through these sinuses, bacteria are rapidly taken up and digested by the Kupffer cells. This system is very efficient and it removes more than 99% of the bacteria from the hepatic portal blood.

The liver serves as an important *storage site for vitamins and iron*. Sufficient quantities of several vitamins may be stored so as to prevent vitamin deficiency for some period of time:

- Vitamin A: Up to 10 months
- Vitamin D: 3 to 4 months
- Vitamin B_{12}: At least 1 year

Iron is stored in the liver in the form of *ferritin*. When the level of circulating iron becomes low, ferritin releases iron into the blood.

Several substances that contribute to the *blood coagulation* process are formed in the liver. These include fibrinogen, prothrombin, and several of the blood-clotting factors (II, VII, IX, and X). Deficiency in any of these substances leads to impaired blood coagulation.

The liver is capable of *detoxifying or excreting into the bile many drugs*, such as sulfonamides (antibacterial drugs), penicillin, ampicillin, and erythromycin. *Bilirubin*, the major end-product of hemoglobin degradation, is also excreted in the bile. In addition, several *hormones* are metabolized by the liver, including thyroid hormone and all of the steroid hormones, such as estrogen, cortisol, and aldosterone.

In terms of nutrients, the liver is the most important metabolic organ in the body. It receives a large volume of nutrient-rich blood directly from the digestive tract, which provides an abundant amount of substrates for metabolism. *Metabolic processes involving carbohydrates* include the following:

- Storage of a significant amount of glycogen
- Conversion of galactose and fructose into glucose
- Gluconeogenesis

Metabolic processes involving proteins include the following:

- Deamination of amino acids
- Formation of urea (for removal of ammonia from the body fluids)
- Formation of plasma proteins
- Conversion of amino acids into other amino acids and other essential compounds

Most cells in the body metabolize lipids; however, some processes of *fat metabolism* occur mainly in the liver. These include:

- Oxidation of fatty acids to supply energy for other body functions.
- Synthesis of cholesterol, phospholipids, and lipoproteins.
- Synthesis of fat from proteins and carbohydrates.

Another important product of liver metabolism is *bile*, which is necessary for the digestion and absorption of dietary lipids. Bile is an aqueous, alkaline fluid consisting of a complex mixture of organic and inorganic components. The major organic constituents of bile are the *bile salts*, which account for approximately 50% of the solid components. Derived from cholesterol, bile salts are *amphipathic molecules*. In other words, these molecules have a hydrophilic region and a hydrophobic region. Inorganic ions are also present in the bile and include Na^+, K^+, Ca^{++}, Cl^-, and HCO_3^- ions. The total number of cations exceeds the total number of anions.

Bile is produced continuously by the liver. The bile salts are secreted by the hepatocytes, and the water, sodium bicarbonate, and other inorganic salts are added by the cells of the bile ducts within the liver. The bile is then transported by way of the *common bile duct* to the duodenum. Bile facilitates fat digestion and absorption throughout the length of the small intestine. In the terminal region of the ileum, the final segment of the small intestine, the bile salts are actively reabsorbed into the blood. They are then returned to the liver by way of the hepatic portal system and resecreted into the bile. This recycling of the bile salts from the small intestine back to the liver is referred to as the *enterohepatic circulation*.

Bile secretion by the liver is stimulated by the following:

- Bile salts
- Secretin
- Parasympathetic stimulation

The return of the *bile salts* to the liver from the small intestine is the most potent stimulus of bile secretion. In fact, these bile salts may cycle two to five times during each meal. The intestinal hormone *secretin*, which is released in response to acid in the duodenum, enhances the aqueous alkaline secretion by the liver. Secretin has no effect on the secretion of bile salts. During the cephalic phase of digestion, before food even reaches the stomach or intestine, *parasympathetic stimulation*, by way of the vagus nerve, promotes bile secretion from the liver.

20.9 Gallbladder

The gallbladder is attached to the inferior surface of the liver. This pear-shaped organ is about 7 to 10 cm long. During a meal, bile enters the duodenum

from the common bile duct through the *Sphincter of Oddi*. Between meals, this sphincter is closed to prevent the bile from entering the small intestine. As a result, much of the bile secreted from the liver is backed up the common bile duct into the *cystic duct* and into the gallbladder. Because the gallbladder is a small organ, it can accommodate only 35 to 100 ml of bile when full.

Within the gallbladder, sodium is actively removed from the bile. Chloride follows the sodium down its electrical gradient and water follows osmotically. As a result, the organic constituents of the bile are concentrated five- to twentyfold. During a meal, when the bile is needed for digestion, the gallbladder contracts, and the bile is squeezed out and into the duodenum. Contraction is elicited by *cholecystokinin*, an intestinal hormone released in response to the presence of chyme, especially lipids, in the duodenum.

20.10 Pancreas

The pancreas, which is about 12 to 15 cm long, is located behind the stomach along the posterior abdominal wall. Exocrine glands within the pancreas secrete an aqueous fluid referred to as pancreatic juice. The pancreas secretes approximately 1 liter of this fluid per day. Pancreatic juice is alkaline and contains a high concentration of bicarbonate ion. It is transported to the duodenum by the pancreatic duct. Pancreatic juice neutralizes the acidic chyme entering the duodenum from the stomach. Neutralization not only prevents damage to the duodenal mucosa, it creates a neutral or slightly alkaline environment that is optimal for the function of pancreatic enzymes. The pancreas also secretes several enzymes that are involved in the digestion of carbohydrates, proteins, and lipids.

There are three major phases of pancreatic secretion:

1. *Cephalic Phase*: Approximately 20% of the pancreatic secretory response to a meal.
2. *Gastric Phase*: 5% to 10% of the pancreatic secretory response to a meal.
3. *Intestinal Phase*: Approximately 70% to 80% of the pancreatic secretory response to a meal.

As with gastric secretion, both nervous stimulation and hormones regulate secretion from the pancreas. During the *cephalic phase* and the *gastric phase*, the pancreas secretes a low-volume, enzyme-rich fluid. This secretion is mediated by the vagus nerve.

Most pancreatic secretion takes place during the *intestinal phase*. The intestinal hormone *secretin* stimulates the release of a large volume of pancreatic juice with a high concentration of bicarbonate ion. Secretin, "nature's antacid," is released in response to acidic chyme in the duodenum (maximal release at pH <3). The intestinal hormone, *cholecystokinin*, is released in response to the

presence of the products of protein and lipid digestion. Cholecystokinin then stimulates the release of digestive enzymes from the pancreas.

20.11 Transport of bile and pancreatic juice to the small intestine

Bile is secreted by the liver, stored in the gallbladder, and used in the small intestine. It is transported toward the small intestine by the *hepatic duct* (from the liver) and the *cystic duct* (from the gallbladder), which join together to form the *common bile duct*. Pancreatic juice is transported toward the small intestine by the *pancreatic duct*. The common bile duct and the pancreatic duct join together to form the *hepatopancreatic ampulla* which empties into the duodenum. The entrance to the duodenum is surrounded by the *Sphincter of Oddi*. This sphincter is closed between meals in order to prevent bile and pancreatic juice from entering the small intestine. The Sphincter of Oddi relaxes in response to the intestinal hormone cholecystokinin, thus allowing biliary and pancreatic secretions to flow into the duodenum.

20.12 Small intestine

The small intestine is the longest (>3 meters in a living human) and most convoluted organ in the digestive system. It is divided into three segments:

1. *Duodenum*: First 20 to 30 cm.
2. *Jejunum*: Next two-fifths of the small intestine.
3. *Ileum*: Remaining three-fifths of the small intestine.

The small intestine is the region where most digestion and absorption take place. As such, the mucosa of the small intestine is well-adapted for these functions with the following anatomical modifications:

- Plicae circulares
- Villi
- Microvilli

The *plicae circulares*, or circular folds, form internal rings around the circumference of the small intestine. These rings are found along the length of the small intestine. They are formed from inward foldings of the mucosal and submucosal layers of the intestinal wall. The plicae circulares are particularly well developed in the duodenum and jejunum and increase the absorptive surface area of the mucosa about threefold.

Each plica is covered with millions of smaller projections of mucosa referred to as *villi*. Two types of epithelial cells cover the villi: the *goblet cells* and the *absorptive cells*. The *goblet cells* produce mucus. The *absorptive*

cells, found in a single layer covering the villi, are far more abundant. Taken together, the villi increase the absorptive surface area another tenfold.

Microvilli are microscopic projections found on the luminal surface of the absorptive cells. Each absorptive cell may have literally thousands of microvilli forming the *brush border*. These structures increase the surface area for absorption another twentyfold. All together, these three anatomical adaptations of the intestinal mucosa, plicae circulares, villi, and microvilli, increase the surface area as much as 600-fold which has a profound positive effect on the absorptive process.

20.12.1 Motility of the small intestine

Both segmentation and peristalsis take place in the small intestine. *Segmentation* mixes the chyme with the digestive juices and exposes the chyme to the intestinal mucosa for absorption. This form of motility causes only a small degree of forward movement of the chyme along the small intestine. *Peristalsis*, the wave-like form of muscle contraction, primarily moves chyme along the intestine and causes only a small amount of mixing. These contractions are weak and slow in the small intestine. In this way, there is sufficient time for complete digestion and absorption of the chyme as it moves forward. Intestinal peristaltic contractions are normally limited to short distances (1 to 4 cm).

Segmentation contractions occur as the result of the *basic electrical rhythm* (*BER*) of the pacemaker cells in the small intestine. This form of muscular activity is slight or absent between meals. The motility of the small intestine may be enhanced during a meal by the following:

- Distension of the small intestine
- Gastrin
- Extrinsic nerve stimulation

During a meal, segmentation occurs initially in the duodenum and the ileum. The movement of chyme into the intestine and the *distension of the duodenum* elicit segmentation contractions in this segment of the small intestine. Segmentation of the empty ileum is caused by *gastrin* released in response to distension of the stomach. This mechanism is referred to as the *gastroileal reflex*. *Parasympathetic stimulation*, by way of the vagus nerve, further enhances segmentation. *Sympathetic stimulation* inhibits this activity.

20.12.2 Digestion and absorption in the small intestine

Most digestion and absorption of carbohydrates, proteins, and lipids occurs in the small intestine. A summary of the digestive enzymes involved in these processes is found in Table 20.3.

Table 20.3 Digestive Enzymes

Nutrient Molecule	Enzyme	Action of Enzyme	Source of Enzyme	Site of Action
		Carbohydrate		
Polysaccharide (Starch)	Amylase	Fragment polysaccharides into disaccharides (maltose)	Salivary glands, Pancreas	Mouth, Stomach, Small intestine
Disaccharides	Disaccharidases (maltase, lactase, sucrase)	Hydrolyze disaccharides into monosaccharides (glucose, galactose, fructose)	Absorptive cells of small intestine	Brush border of absorptive cells
		Protein		
Protein (Long Peptide Chain)	Pepsin	Fragment proteins into smaller peptides	Stomach chief cells	Stomach
Peptides	Trypsin, chymotrypsin, carboxypeptidase	Hydrolyze peptides into di- and tripeptides	Pancreas	Small intestine
Di- and Tripeptides	Aminopeptidases	Hydrolyze di- and tripeptides into amino acids	Absorptive cells of small intestine	Brush border of absorptive cells
		Lipid		
Triglyceride	Lingual lipase[a]	Hydrolyze triglycerides into monoglycerides and free fatty acids	Salivary glands	Mouth, stomach
	Pancreatic lipase	Hydrolyze triglycerides into monoglycerides and free fatty acids	Pancreas	Small intestine

[a] The role of lingual lipase in the digestion of dietary lipids is minor as it accounts for less than 10% of the enzymatic breakdown of triglycerides.

20.12.3 Carbohydrates

Approximately 50% of the human diet is composed of starch. Other major dietary carbohydrates include the disaccharides, sucrose (30%, table sugar, composed of glucose and fructose), and lactose (6%, milk sugar, composed of glucose and galactose). Starch is initially acted upon by amylase. *Salivary amylase* breaks down starch molecules in the mouth and the stomach. *Pancreatic amylase* carries on this activity in the small intestine. Amylase fragments polysaccharides into disaccharides (maltose, composed of two glucose molecules).

The disaccharide molecules, primarily maltose, are presented to the brush border of the absorptive cells. As the disaccharides are absorbed, *disaccharidases* (maltase, sucrase, and lactase) split these nutrient molecules into monosaccharides (glucose, fructose, and galactose).

Glucose and galactose enter the absorptive cells by way of *secondary active transport*. Cotransport carrier molecules associated with the disaccharidases in the brush border transport both the monosaccharide and an Na^+ ion from the lumen of the small intestine into the absorptive cell. This process is referred to as "secondary" because the cotransport carriers operate passively and do not require energy. However, they do require a concentration gradient for the transport of Na^+ ions into the cell. This gradient is established by the active transport of Na^+ ions out of the absorptive cell at the basolateral surface.

Fructose enters the absorptive cells by way of facilitated diffusion. All monosaccharide molecules exit the absorptive cells by way of facilitated diffusion and enter the blood capillaries.

The ability of the human small intestine to absorb free sugars is quite remarkable. It has been estimated that hexoses equivalent to 22 pounds of sucrose can be absorbed daily. There appears to be little physiologic regulation of sugar absorption.

20.12.4 Proteins

Protein digestion begins in the stomach by the action of the gastric enzyme *pepsin*. This enzyme fragments large protein molecules into smaller peptide chains. Digestion is continued in the small intestine by the pancreatic enzymes *trypsin*, *chymotrypsin,* and *carboxypeptidase*. Similar to pepsin (pepsinogen), these enzymes are secreted as inactive precursors (trypsinogen, chymotrypsinogen, and procarboxypeptidase). The intestinal enzyme *enterokinase* activates trypsin at the brush border. Trypsin then activates chymotrypsin and carboxypeptidase. These pancreatic enzymes hydrolyze the peptide chains into amino acids (40%), as well as dipeptides and tripeptides (60% combined).

Similar to glucose and galactose, the amino acids enter the absorptive cells by way of secondary active transport. Once again, energy is expended to pump Na^+ ions out of the absorptive cells, creating a concentration gradient for the cotransport of amino acids and Na^+ ions into the cell.

Dipeptides and tripeptides are also presented to the brush border of the absorptive cells. As the nutrient molecules are absorbed, aminopeptidases split them into their constituent amino acids. The activity of the *aminopeptidases* accounts for approximately 60% of protein digestion. The amino acid molecules then exit the absorptive cells by way of facilitated diffusion and enter the blood capillaries.

20.12.5 Lipids

Dietary fat consists primarily of triglycerides. Fat digestion begins in the mouth and stomach by the action of the salivary enzyme, lingual lipase. However, the role of this enzyme is minor as it accounts for less than 10% of the enzymatic breakdown of triglycerides. *Gastric lipase* is also responsible for a very small degree of lipid digestion.

Lipids are digested primarily in the small intestine. The first step in this process involves the action of the *bile salts* contained in the bile. Bile salts cause *emulsification*, which is the dispersal of large fat droplets into a suspension of smaller droplets (<1 μm). This process creates a significantly increased surface area upon which fat-digesting enzymes can act.

Intact triglycerides are too large to be absorbed. Therefore, *pancreatic lipase* acts on the lipid droplets to hydrolyze the triglyceride molecules into *monoglycerides* and *free fatty acids*. These constituent molecules are water insoluble and would tend to float on the surface of the aqueous chyme. Therefore, they must be transported to the absorptive surface. This process is carried out by *micelles*, which are sphere-like structures formed by the amphipathic bile salts. The bile salts associate with each other such that the polar region of the molecules are oriented outward, making them water soluble. The nonpolar region faces inward away from the surrounding water. The monoglycerides and free fatty acids are carried in this interior region of the micelle. Upon reaching the brush border of the absorptive cells, the monoglycerides and free fatty acids leave the micelles and enter the cells by simple diffusion. Because they are nonpolar, these molecules move passively through the lipid bilayer of the cell membrane. This process takes place primarily in the jejunum and proximal ileum. The bile salts are absorbed in the distal ileum by way of either passive diffusion or secondary active transport.

Within the absorptive cells, the monoglycerides and free fatty acids are transported to the endoplasmic reticulum which contains the necessary enzymes to resynthesize these substances into triglycerides. The newly synthesized triglycerides then move to the Golgi apparatus. Within this organelle, they are packaged in a lipoprotein coat consisting of phospholipids, cholesterol, and apoproteins. These protein-coated lipid globules,

referred to as *chylomicrons*, are now water soluble. Approximately 90% of the chylomicron consists of triglycerides.

Chylomicrons leave the absorptive cell by way of exocytosis. Because they are unable to cross the basement membrane of the blood capillaries, the chylomicrons enter the *lacteals* which are part of the lymphatic system. The vessels of the lymphatic system converge to form the thoracic duct which drains into the venous system near the heart. Therefore, unlike the products of carbohydrate and protein digestion, which are transported directly to the liver by way of the hepatic portal vein, absorbed lipids are diluted in the blood of the circulatory system before they reach the liver. This dilution of the lipids prevents the liver from being overwhelmed with more fat than it can process at one time.

20.12.6 Water and electrolytes

Each day in an average adult, about 5.5 liters of food and fluids move from the stomach to the small intestine as chyme. An additional 3.5 liters of pancreatic and intestinal secretions produce a total of 9 liters of material in the lumen. Most of this (>7.5 liters) is absorbed from the small intestine. The absorption of nutrient molecules, which takes place primarily in the duodenum and jejunum, creates an osmotic gradient for the passive absorption of water.

Sodium may be absorbed either passively or actively. Passive absorption occurs when the electrochemical gradient favors the movement of Na^+ between the absorptive cells through "leaky" tight junctions. Sodium is actively absorbed by way of transporters in the absorptive cell membrane. One type of transporter carries an Na^+ ion and a Cl^- ion into the cell. Another carries an Na^+ ion, a K^+ ion, and two Cl^- ions into the cell.

20.13 Large intestine

The *large intestine* is the region of the digestive tract from the ileocecal valve to the anus. Approximately 1.5 meters in length, this organ has a larger diameter than the small intestine. The mucosa of the large intestine is composed of absorptive cells and mucus-secreting goblet cells. Similar to the small intestine, the mucosal layer of the large intestine contains lymphatic nodules to protect against microbial infection. In contrast to the small intestine, the mucosa in this organ does not form villi.

The large intestine consists of the following structures:

- Cecum
- Appendix
- Colon
- Rectum

The *cecum*, which is the proximal-most portion of the large intestine, receives chyme from the ileum of the small intestine through the *ileocecal valve*. The *appendix*, the small projection at the bottom of the cecum, is a lymphoid tissue. This tissue contains lymphocytes and assists in the defense against bacteria that enter the body through the digestive system.

The largest portion of the large intestine is the *colon*. It consists of four regions: ascending colon (travels upward toward the diaphragm on the right side of the abdomen), transverse colon (crosses the abdomen under the diaphragm), descending colon (travels downward through the abdomen on the left side), and the sigmoid colon (S-shaped region found in the lower abdomen). The sigmoid colon is continuous with the *rectum*, which leads to the external surface of the body through the *anus*.

The large intestine typically receives 500 to 1500 ml of chyme from the small intestine per day. As discussed, most digestion and absorption have already taken place in the small intestine. In fact, there are no digestive enzymes produced by the large intestine. At this point in the human digestive tract, the chyme consists of indigestible food residues (e.g., cellulose), unabsorbed biliary components, and any remaining fluid. Therefore, the two major functions of the large intestine are drying and storage.

The colon absorbs most of the water and salt from the chyme resulting in this "*drying*" or concentrating process. The absorption of water occurs passively down its osmotic gradient following the active transport of ions. As a result, only about 100 ml of water is lost through this route daily. The remaining contents, now referred to as *feces*, are "*stored*" in the large intestine until they can be eliminated by way of defecation.

In addition to water, the large intestine also absorbs electrolytes, several B complex vitamins, and vitamin K. The normally occurring bacteria, also referred to as *microflora*, that reside in the large intestine produce vitamin K and folic acid which are then absorbed.

20.13.1 Motility of the large intestine

The longitudinal layer of smooth muscle in the small intestine is continuous. In the large intestine, this layer of muscle is concentrated into three flat bands referred to as *taniae coli*. Furthermore, the large intestine appears to be subdivided into a chain of pouches or sacs referred to as *haustra*. The haustra are formed because the bands of taniae coli are shorter than the underlying circular layer of smooth muscle and cause the colon to bunch up.

Movements through the large intestine are typically quite sluggish. It will often take 18 to 24 hours for materials to pass through its entire length. The primary form of motility in the large intestine is *haustral contractions*, or *haustrations*. These contractions are produced by the inherent rhythmicity of smooth muscle cells in the colon. Haustrations, which result in the pronounced appearance of the haustra, are similar to segmentation contractions in the small intestine. Haustrations are *nonpropulsive* and serve

primarily to slowly move the contents back and forth, exposing them to the absorptive surface.

In contrast to segmentation contractions in the small intestine (9 to 12 per minute), haustral contractions occur much less frequently (up to 30 minutes between contractions). These very slow movements allow for the growth of *bacteria* in the large intestine. Normally, the bacterial flora in this region is harmless.

A second form of motility in the large intestine is the *mass movement*. Three or four times per day, typically after a meal, a strong propulsive contraction occurs that moves a substantial bolus of chyme forward toward the distal portion of the colon. Mass movements may result in the sudden distension of the rectum that elicits the defecation reflex.

20.13.2 Secretion of the large intestine

The large intestine produces an *alkaline mucus secretion*. The function of this secretion is to protect the mucosa from mechanical or chemical injury. *Mucus* provides lubrication to facilitate the movement of the contents of the lumen. *Bicarbonate ion* neutralizes the irritating acids produced by the local bacterial fermentation. Colonic secretion increases in response to mechanical or chemical stimuli. The mechanism of the enhanced secretion involves both intrinsic nerve and vagal nerve reflexes.

Medical terminology

Ampulla (ăm′pūl-ă): Sac-like dilatation of a duct.
Amylase (ăm′ĭ-lās): Enzyme that hydrolyzes or splits starch into smaller molecules.
Anus (ā′nŭs): Outlet for body waste from the rectum to the body surface.
Bolus (bō′lŭs): A concentrated mass.
Cecum (sē′kŭm): Blind pouch forming the first portion of the large intestine.
Cephalic (sĕ-făl′ĭk): Referring to the cranium.
Cholecystokinin (kō″lĕ-sĭs″tō-kīn′ĭn): Hormone produced in the small intestine in response to the presence of lipids.
Chylomicron (kī″lō-mī′krŏn): Lipoprotein molecule formed in the small intestine for the transport of triglycerides in the blood.
Chyme (kīm): Food materials mixed with gastric juice forming a thick semi-fluid substance.
Colon (kō′lŏn): Large intestine.
Enteric (ĕn-tĕr′ĭk): Referring to the small intestine.
Esophagus (ē-sŏf′ă-gŭs): Muscular tube that transports ingested materials from the pharynx to the stomach.
Feces (fē′sēz): Body waste eliminated from the colon through the anus.
Haustrum (haw′strŭm): Sac-like structure in the large intestine.
Ileocecal (ĭl′ē-ō-sē′kăl): Referring to the ileum and the cecum.

Lacteal (lăk′tē-ăl): Lymphatic capillary found in the villi of the small intestine.
Lipase (lī′pās): Enzyme that hydrolyzes or splits triglycerides into monoglycerides and free fatty acids.
Mastication (măs-tĭ-kā′shŭn): Chewing.
Mesentery (měs′ĕn-tĕr″ē): Serous membrane that attaches organs to the abdominal wall.
Micelle (mī′sĕl): Sphere of bile salt molecules needed to transport fatty acids and monoglycerides to the absorptive cells in the small intestine.
Mucosa (mū-kō′să): Mucous membrane that lines hollow organs.
Peristalsis (pĕr-ĭ-stăl′sĭs): Wave-like muscular contraction that propels chyme along the gastrointestinal tract.
Peritoneum (pĕr″ĭ-tō-nē′ŭm): Serous membrane lining the abdominal cavity.
Pharynx (făr′ĭnks): Common passageway for food and air.
Plexus (plĕks′ŭs): Network of nerves.
Pyloric (pī-lor′ĭk): Referring to the junction between the stomach and the duodenum.
Secretin (sē-krē′tĭn): Hormone produced in the small intestine in response to the presence of acid.
Segmentation (sĕg″mĕn-tā′shŭn): Stationary muscular contraction that mixes chyme with digestive enzymes and exposes chyme to the absorptive surface.

Bibliography

Brown, J. H., and Taylor, P., Muscarinic Receptor Agonists and Antagonists, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 7.

Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.

Ganong, W. F., *Review of Medical Physiology*, 19th ed., Appleton & Lange, Stamford, CT, 1999.

Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., W.B. Saunders, Philadelphia, PA, 2006.

Hoogerwerf, W. A., and Pasricha, P. J., Pharmacotherapy of Gastric Acidity, Peptic Ulcers, and Gastroesophageal Reflux Disease, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 36.

Johnson, L. R., Salivary Secretion, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 7.

Johnson, L. R., Gastric Secretion, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 8.

Johnson, L. R., Pancreatic Secretion, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 9.

Johnson, L. R., Digestion and Absorption, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 11.

Lichtenberger, L. M., Gastrointestinal Physiology, in *Physiology Secrets*, Raff, H., Ed., Hanley and Belfus, Philadelphia, PA, 1999, chap. 5.

Pappano, A. J., and Katzung, B. G., Cholinergic-Blocking Drugs, in *Basic and Clinical Pharmacology*, 8th ed., Katzung, B. G., Ed., Lange Medical/McGraw-Hill, New York, 2001, chap. 8.

Pasricha, P. J., Treatment of Disorders of Bowel Motility and Water Flux; Antiemetics; Agents Used in Biliary and Pancreatic Disease, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 37.

Sherwood, L., *Human Physiology from Cells to Systems,* 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.

Silverthorn, D. U., *Human Physiology, an Integrated Approach,* 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.

Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.

Weisbrodt, N. W., Regulation: Nerves and Smooth Muscle, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 2.

Weisbrodt, N. W., Swallowing, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 3.

Weisbrodt, N. W., Gastric Emptying, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 4.

Weisbrodt, N. W., Bile Production, Secretion, and Storage, in *Gastrointestinal Physiology*, 7th ed., Johnson, L. R., Ed., Mosby, St. Louis, MO, 2007, chap. 10.

Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

chapter twenty one

The renal system

Study objectives

- List the vascular components of the nephron and describe their function
- List the tubular components of the nephron and describe their function
- Distinguish between a cortical nephron and a juxtamedullary nephron
- Define the *three basic renal processes*
- Describe the components of the filtration barrier
- Explain how the filtration coefficient and the net filtration pressure determine glomerular filtration
- Describe the mechanisms by which sodium, chloride, and water are reabsorbed
- Describe how each segment of the tubule handles sodium, chloride, and water
- Distinguish between the vertical osmotic gradient and the horizontal osmotic gradient
- Describe the functions of the vasa recta
- Describe the process by which potassium ions are secreted and the mechanism that regulates this process
- Define *plasma clearance*
- Explain how the plasma clearance of inulin is used to determine glomerular filtration rate
- Explain how the plasma clearance of para-aminohippuric acid is used to determine the effective renal plasma flow
- Explain how the myogenic mechanism and tubuloglomerular feedback are responsible for the autoregulation of renal blood flow
- Explain how sympathetic nerves, angiotensin II, and prostaglandins affect the resistance of the afferent arteriole
- Describe the factors that regulate the release of renin
- Explain how the control of sodium excretion regulates plasma volume
- Describe the mechanisms by which sodium excretion is controlled
- Explain how the control of water excretion regulates plasma osmolarity
- Describe the mechanisms by which water balance is maintained

21.1 Introduction

The kidneys are organs specialized to filter the blood. As such, they contribute importantly to the removal of metabolic waste products as well as to the maintenance of fluid and electrolyte balance. Specific functions of the kidneys include the following:

- Regulation of extracellular fluid volume.
- Regulation of inorganic electrolyte concentration in the extracellular fluid.
- Regulation of the osmolarity of the extracellular fluid.
- Removal of metabolic waste products.
- Excretion of foreign compounds.
- Maintenance of acid–base balance.
- Hormone and enzyme production.

The *regulation of extracellular fluid volume*, in particular, plasma volume, is important in the long-term regulation of blood pressure. An increase in plasma volume leads to an increase in blood pressure, and a decrease in plasma volume leads to a decrease in blood pressure. Plasma volume is regulated primarily by altering the excretion of sodium in the urine. Other *inorganic electrolytes* regulated by the kidneys include chloride, potassium, calcium, magnesium, sulfate, and phosphate.

The kidneys also *regulate the osmolarity of the extracellular fluid*, in particular, plasma osmolarity. The maintenance of plasma osmolarity close to 290 mOsm prevents any unwanted movement of fluid in or out of the body's cells. An increase in plasma osmolarity causes water to leave the cells, leading to cellular dehydration. A decrease in plasma osmolarity causes water to enter the cells, leading to cellular swelling and possibly lysis. Plasma osmolarity is regulated primarily by altering the excretion of water in the urine.

As the major excretory organs in the body, the kidneys are responsible for the *removal of many metabolic waste products*. These include urea and uric acid, which are nitrogenous waste products of amino acid and nucleic acid metabolism, respectively; creatinine, a breakdown product of muscle metabolism; and urobilinogen, a metabolite of hemoglobin which gives urine its yellow color. *Foreign compounds* excreted by the kidneys include drugs (e.g., penicillin, nonsteroidal antiinflammatory drugs), food additives (e.g., saccharin, benzoate), pesticides, and other exogenous nonnutritive materials that have entered the body. If allowed to accumulate, these substances become quite toxic.

Along with the respiratory system, the renal system *maintains acid–base balance* by altering the excretion of hydrogen ions and bicarbonate ions in the urine. When the extracellular fluid becomes acidic and pH decreases, then the kidneys excrete H^+ ions and conserve HCO_3^- ions. Conversely, when the extracellular fluid becomes alkaline and pH increases, then the kidneys

conserve H^+ ions and excrete HCO_3^- ions. Normally, the pH of the arterial blood is 7.4.

Although the kidneys are not considered endocrine glands, per se, they are involved in *hormone production*. Erythropoietin is a peptide hormone that stimulates red blood cell production in the bone marrow. Its primary source is the kidneys. Erythropoietin is secreted in response to renal hypoxia. Chronic renal disease may impair the secretion of erythropoietin, leading to the development of anemia.

The kidneys also *produce enzymes*. The enzyme renin is part of the renin-angiotensin-aldosterone system. As will be discussed, these substances play an important role in the regulation of plasma volume and, therefore, blood pressure. Other renal enzymes are needed for the conversion of vitamin D into its active form, 1,25-dihydroxyvitamin D_3, which is involved with calcium balance.

21.2 Functional anatomy of the kidneys

The kidneys lie outside of the peritoneal cavity in the posterior abdominal wall, one on each side of the vertebral column, slightly above the waistline. In the adult human, each kidney is approximately 11 cm long, 6 cm wide, and 3 cm thick. These organs are divided into two regions: the inner *renal medulla* and the outer *renal cortex*.

The functional unit of the kidney is the *nephron* (see Figure 21.1 and Figure 21.2). There are well over 1 million nephrons in each kidney. The nephron has two components: the *vascular component* and the *tubular component*.

21.2.1 Vascular component

Filtration of the plasma takes place at the *glomerulus* (that is, *glomerular capillaries*), which is located in the cortical region of the kidney. Water and solutes exit the vascular compartment through these capillaries to be processed by the tubular component of the nephron. Blood is delivered to the glomerulus by the *afferent arterioles*. The glomerular capillaries then join together to form a second arteriole referred to as the *efferent arteriole*. All cellular elements of the blood (red blood cells, white blood cells, and platelets) as well as the unfiltered plasma continue through this vessel. The efferent arterioles then lead to a second set of capillaries, the *peritubular capillaries*. These capillaries provide nourishment to the renal tissue and return the substances reabsorbed from the tubule to the vascular compartment. Peritubular capillaries are closely associated with all portions of the renal tubules and wrap around them. These capillaries then join together to form venules and progressively larger veins that remove the blood from the kidneys.

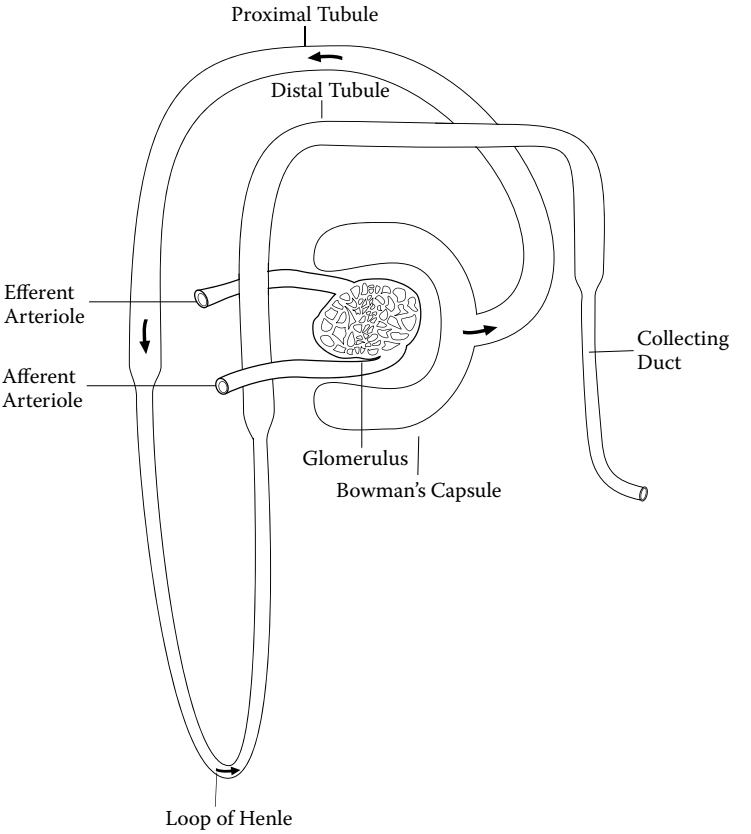


Figure 21.1 The nephron. The functional unit of the kidney is the nephron. It has two components. The vascular component includes the afferent arteriole, which carries blood toward the glomerulus, where filtration of the plasma takes place. The efferent arteriole carries the unfiltered blood away from the glomerulus. The tubular component of the nephron includes Bowman's capsule, which receives the filtrate, the proximal tubule, the Loop of Henle, the distal tubule, and the collecting duct. The tubule processes the filtrate, excreting waste products and reabsorbing nutrient molecules, electrolytes, and water.

21.2.2 Tubular component

Approximately 180 liters of filtrate is processed by the kidneys each day. Depending upon the volume of fluid intake, about 99% of this filtrate must be reabsorbed from the renal tubule back into the vascular compartment. The movement of substances out of the tubule is facilitated by its structure, which consists of *a single layer of epithelial cells*. As will be discussed, each region of the tubule plays a different role in the reabsorption process.

Upon leaving the glomerular capillaries, the filtrate enters the first portion of the tubule, *Bowman's capsule*. The glomerulus is pushed into Bowman's

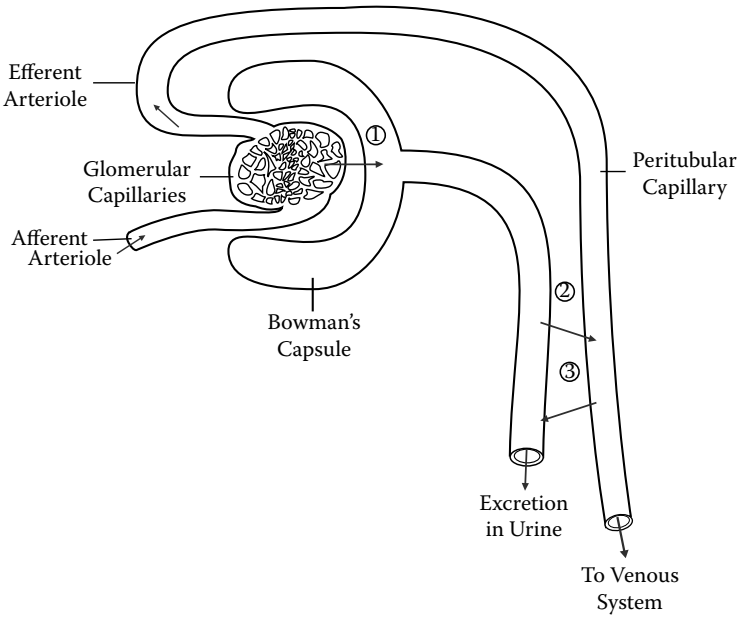


Figure 21.2 Basic renal processes. These processes include filtration, reabsorption, and secretion. (1) Filtration is the movement of fluid and solutes from the glomerular capillaries into Bowman's capsule. (2) Reabsorption, which takes place throughout the nephron, is the movement of filtered substances out of the tubule and into the surrounding peritubular capillaries. (3) Secretion is the movement of selected unfiltered substances from the peritubular capillaries into the renal tubule for excretion. Any substance that is filtered or secreted, but not reabsorbed, is excreted in the urine.

capsule, much like a fist pushed into a balloon or a catcher's mitt. From Bowman's capsule, the filtrate passes through the *proximal tubule*, which is also located in the cortex of the kidney. The next segment of the tubule is the *Loop of Henle*. This portion of the tubule is found in the medulla of the kidney. The descending limb penetrates into the medulla, and the ascending limb returns toward the cortex. From the Loop of Henle, the filtrate passes through the *distal tubule* in the cortex of the kidney. Finally, up to eight distal tubules empty into a *collecting duct*. The collecting ducts run downward through the medulla. Any filtrate remaining within the tubule at the end of the collecting duct drains through the renal pelvis to the ureters and is excreted as urine.

There are two types of nephrons that are distinguished by their anatomical characteristics: the *cortical nephron* and the *juxtamedullary nephron*.

The glomerulus of each *cortical nephron* is located in the outer region of the cortex. Furthermore, the Loops of Henle in these nephrons are short and do not penetrate deeply into the medulla. In humans, 70% to 80% of the nephrons are of the cortical type.

In contrast, the glomerulus of each *juxtamedullary nephron* is located in the inner region of the cortex, close to the medulla. The Loops of Henle in these nephrons are significantly longer, penetrating to the innermost region of the medulla. Within the medulla, the peritubular capillaries of the nephrons are modified to form the *vasa recta*, or straight vessels. Similar to the Loops of Henle, the vasa recta descend deep into the medulla, form a hairpin loop, and then ascend back toward the cortex. In fact, these vessels run parallel, and in close association, with the Loops of Henle and the collecting ducts. The remaining 20% to 30% of the nephrons in the human kidney are of the juxtamedullary type.

21.3 Basic renal processes

There are three basic renal processes performed by the nephron (see Figure 21.2):

1. Filtration
2. Reabsorption
3. Secretion

Filtration is the movement of fluid and solutes from the glomerular capillaries into Bowman's capsule. Filtration is a *nonselective* process, such that everything in the plasma except for the plasma proteins is filtered. Approximately 20% of the plasma is filtered as it passes through the glomerulus. On average, this results in a *glomerular filtration rate* (*GFR*) of 125 ml/min or 180 liters of filtrate per day.

Reabsorption is the movement of filtered substances from the renal tubule into the peritubular capillaries for return to the vascular compartment. This process takes place throughout the tubule. Approximately 178.5 liters of filtrate are reabsorbed, resulting in an average urine output of 1.5 liters per day.

Secretion is the movement of selected unfiltered substances from the peritubular capillaries into the renal tubule for excretion. Any substance that is filtered or secreted, but not reabsorbed, is *excreted* in the urine.

The maintenance of plasma volume and plasma osmolarity occurs through the regulation of the *renal excretion* of sodium, chloride, and water. Each of these substances is freely filtered from the glomerulus and reabsorbed from the tubule. None of these substances are secreted. Because salt and water intake in the diet may vary widely, the renal excretion of these substances is also highly variable. In other words, the kidneys must be able to produce a wide range of urine concentrations and urine volumes. The most dilute urine produced by humans is 65 to 70 mOsm/L, and the most concentrated the urine can be is 1200 mOsm/L (recall that the plasma osmolarity is 290 mOsm/L). The volume of urine produced per day depends largely upon fluid intake. As fluid intake increases, then the urine output

increases to excrete the excess water. Conversely, as fluid intake decreases or as an individual becomes dehydrated, then the urine output decreases in order to conserve water.

On average, 500 mOsm of waste products must be excreted in the urine per day. The minimum volume of water in which these solutes can be dissolved is determined by the ability of the kidney to produce a maximally concentrated urine of 1200 mOsm/L:

$$\frac{500 \text{ mOsm/day}}{1200 \text{ mOsm/L}} = 420 \text{ ml water/day}$$

This volume, referred to as *obligatory water loss*, is 420 ml water/day. In other words, 420 ml of water will be lost in the urine each day in order to excrete metabolic waste products regardless of water intake.

21.4 Glomerular filtration

The first step in the formation of urine is glomerular filtration. The barrier to filtration is designed to facilitate the movement of fluid from the glomerular capillaries into Bowman's capsule without any loss of cellular elements or plasma proteins. There are two advantages to maximizing GFR:

1. Waste products are rapidly removed from the body.
2. All body fluids are filtered and processed by the kidneys several times per day resulting in the precise regulation of volume and composition of these fluids.

21.4.1 Filtration barrier

The *filtration barrier* is composed of three structures:

1. Glomerular capillary wall
2. Basement membrane
3. Inner wall of Bowman's capsule

Like the walls of other capillaries, the *glomerular capillary wall* consists of a single layer of endothelial cells. However, the glomerular endothelial cells are specialized in that they are *fenestrated*. The presence of large pores in these capillaries makes them 100 times more permeable than the typical capillary. These pores are too small, however, to permit the passage of blood cells through them.

The *basement membrane* is an acellular meshwork consisting of collagen and glycoproteins. The *collagen* provides structural support. The negatively charged *glycoproteins* prevent the filtration of plasma proteins into Bowman's capsule.

The *inner wall of Bowman's capsule* consists of specialized epithelial cells referred to as *podocytes*. This layer of epithelial cells is not continuous. Instead, the podocytes have foot-like processes that project outward. The processes of one podocyte interdigitate with the processes of an adjacent podocyte, forming narrow *filtration slits*. These slits provide an ample route for the filtration of fluid.

In summary, the filtrate moves through the pores of the capillary endothelium, through the basement membrane, and, finally, through the filtration slits between the podocytes. This route of filtration is completely acellular.

21.4.2 Determinants of filtration

The glomerular filtration rate is influenced by two factors: the *filtration coefficient* and the *net filtration pressure*.

The *filtration coefficient* is determined by the *surface area* and the *permeability* of the filtration barrier. An increase in the filtration coefficient leads to an increase in GFR, and a decrease in the filtration coefficient leads to a decrease in GFR. However, this factor does not play a role in the daily regulation of GFR, as its value is relatively constant under normal physiological conditions. On the other hand, chronic, uncontrolled hypertension and diabetes mellitus lead to the gradual thickening of the basement membrane and, therefore, a decrease in the filtration coefficient, a decrease in GFR, and impaired renal function.

The *net filtration pressure* is determined by the following forces (see Figure 21.3):

- Glomerular capillary blood pressure
- Plasma colloid osmotic pressure
- Bowman's capsule pressure

Glomerular capillary pressure (P_{GC}) is a hydrostatic pressure that pushes blood out of the capillary. The blood pressure in these capillaries is markedly different from that of typical capillaries. In capillaries elsewhere in the body, the blood pressure at the arteriolar end is about 30 mmHg, and the blood pressure at the venular end is about 10 mmHg (see Chapter 16). These pressures lead to the net filtration of fluid at the inflow end of the capillary and the net reabsorption of fluid at the outflow end of the capillary.

In contrast, blood pressure in the glomerular capillaries is significantly higher and essentially nondecremental. At the inflow end of the capillary near the afferent arteriole, P_{GC} is about 60 mmHg, and at the outflow end of the capillary near the efferent arteriole, P_{GC} is about 58 mmHg. Interestingly, the diameter of the afferent arteriole is larger than that of the efferent arteriole. Therefore, the vascular resistance in the afferent arteriole is comparatively low, and blood flows readily into the glomerular capillaries resulting in a higher pressure. Furthermore, the smaller diameter of the

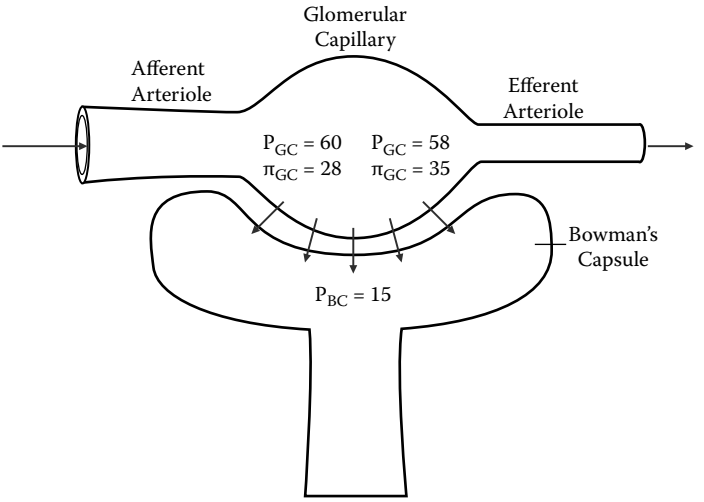


Figure 21.3 Forces determining net filtration pressure. Three forces contribute to the net filtration pressure in the glomerulus. Glomerular capillary blood pressure (P_{GC}) is higher than that of a typical capillary (60 mmHg versus 30 mmHg). Furthermore, P_{GC} remains high throughout the length of the capillary. This is due to the comparatively small diameter of the efferent arteriole, which causes the blood to dam up within the glomerular capillaries. Glomerular capillary pressure promotes filtration along the entire length of the glomerular capillaries. Plasma colloid osmotic pressure (π_{GC}), generated by the plasma proteins, opposes filtration. This force increases from 28 mmHg at the inflow end of the glomerular capillary to 35 mmHg at the outflow end of the capillary. This is due to the concentration of the plasma proteins as the filtration of the plasma fluid progresses. Bowman's capsule pressure (P_{BC}) is generated by the presence of filtered fluid within Bowman's capsule. This pressure opposes filtration with a force of 15 mmHg.

efferent arteriole results in an increase in vascular resistance that limits the flow of blood through this vessel. Consequently, the blood dams up in the glomerular capillaries. The result is a sustained, elevated hydrostatic pressure that promotes the net filtration of fluid along the entire length of the glomerular capillaries.

Plasma colloid osmotic pressure (π_{GC}) is generated by the plasma proteins. These proteins exert an osmotic force on the fluid, which opposes filtration and draws the fluid into the capillary. The π_{GC} is approximately 28 mmHg at the inflow end of the glomerular capillaries. Because 20% of the fluid within the capillaries is filtered into Bowman's capsule, the plasma proteins become increasingly concentrated. Therefore, at the outflow end of the glomerular capillaries, π_{GC} is approximately 35 mmHg.

Bowman's capsule pressure (P_{BC}) is a hydrostatic pressure generated by the presence of filtered fluid within Bowman's capsule. This pressure pushes the fluid out of the capsule and forward toward the remainder of the renal

tubule for processing. Bowman's capsule pressure also tends to oppose filtration. On average, P_{BC} is approximately 15 mmHg.

The net filtration pressure may be summarized as follows:

$$\text{Net filtration pressure} = P_{GC} - \pi_{GC} - P_{BC}$$

Therefore, at the inflow end of the glomerular capillaries,

$$\text{Net filtration pressure} = 60 \text{ mmHg} - 28 \text{ mmHg} - 15 \text{ mmHg} = 17 \text{ mmHg}$$

At the outflow end of the glomerular capillaries:

$$\text{Net filtration pressure} = 58 \text{ mmHg} - 35 \text{ mmHg} - 15 \text{ mmHg} = 8 \text{ mmHg}$$

Under physiological conditions, there is little variation in the values for π_{GC} and P_{BC}. In other words, when plasma protein synthesis is normal, and in the absence of any urinary obstruction that would cause the urine to back up and increase P_{BC}, the primary factor that affects glomerular filtration is P_{GC}. An increase in P_{GC} leads to an increase in GFR, and a decrease in P_{GC} leads to a decrease in GFR.

Glomerular capillary pressure is determined primarily by *renal blood flow* (*RBF*). As RBF increases, then P_{GC} and, therefore, GFR increase. On the other hand, as RBF decreases, then P_{GC} and GFR decrease. Renal blood flow is determined by mean arterial pressure (MAP), and the resistance of the afferent arteriole (aff art):

$$\text{RBF} = \frac{\text{MAP}}{R_{\text{aff art}}}$$

21.5 Tubular reabsorption

The process of *tubular reabsorption* is essential for the conservation of plasma constituents that are important to the body, in particular, electrolytes and nutrient molecules. This process is highly selective in that waste products and substances with no physiological value are not reabsorbed but, instead, are excreted in the urine. Furthermore, the reabsorption of many substances, such as Na^+ ions, H^+ ions, Ca^{++} ions, and water, is physiologically controlled. Consequently, the volume, osmolarity, composition, and pH of the extracellular fluid are precisely regulated.

Throughout its length, the tubule of the nephron is composed of a single layer of epithelial cells. Furthermore, the tubule is in close proximity to the peritubular capillaries. Therefore, reabsorption involves the movement of a substance along the following pathway:

Filtrate within tubular lumen

↓

Across the luminal membrane of the epithelial cell

↓

Through the cytoplasm of the epithelial cell

↓

Across the basolateral membrane of the epithelial cell

↓

Through the interstitial fluid

↓

Across the capillary endothelium

↓

Peritubular capillary blood

This pathway is referred to as *transepithelial transport*.

There are two types of tubular reabsorption: *passive* and *active*. Tubular reabsorption is considered *passive* when each of the steps in transepithelial transport takes place without the expenditure of energy. In other words, the movement of a given substance is from an area of high concentration to an area of low concentration by way of passive diffusion. Water is passively reabsorbed from the tubules back into the peritubular capillaries.

Active reabsorption occurs when the movement of a given substance across either the luminal surface or the basolateral surface of the tubular epithelial cell requires energy. Substances that are actively reabsorbed from the tubule include glucose, amino acids, Na^+ ions, PO_4^{-3} ions, and Ca^{++} ions.

Three generalizations can be made regarding the tubular reabsorption of sodium, chloride, and water:

1. Reabsorption of *Na^+ ions* is an *active* process: 80% of the total energy expended by the kidneys is used for sodium transport out of the tubular epithelial cell.
2. Reabsorption of *Cl^- ions* is a *passive* process: Cl^- ions are reabsorbed according to the electrical gradient created by the reabsorption of Na^+ ions.

3. Reabsorption of *water* is a *passive* process: Water is reabsorbed according to the osmotic gradient created by the reabsorption of Na^+ ions.

In other words, when sodium is reabsorbed, chloride and water follow it.

21.5.1 Sodium reabsorption

Sodium is reabsorbed by different mechanisms as the filtrate progresses through the tubule. Sodium ions leave the filtrate and enter the tubular epithelial cell by way of the following processes (see Figure 21.4):

- Na^+-glucose, Na^+-amino acid, Na^+-phosphate, and Na^+-lactate symporter mechanisms; Na^+-H^+ antiporter mechanism: first half of the proximal tubule.
- Coupled with Cl^- reabsorption by way of both transcellular (through the epithelial cell) and paracellular (in between the epithelial cells) pathways: second half of the proximal tubule.
- Na^+, K^+, $2Cl^-$ symporter mechanism: ascending limb of the Loop of Henle.
- Na^+, Cl^- symporter mechanism: distal tubule.
- Na^+ channels: distal tubule, collecting duct.

More simply, in the early regions of the tubule (proximal tubule and Loop of Henle), Na^+ ions leave the lumen and enter the tubular epithelial cells by way of *facilitated transport mechanisms* that are passive. The diffusion of Na^+ ions is coupled with organic molecules or with other ions that electrically balance the flux of these positively charged ions. In the latter regions of the tubule (distal tubule and collecting duct), Na^+ ions diffuse into the epithelial cells through Na^+ channels.

An essential requirement for the diffusion of Na^+ ions is the creation of a concentration gradient for sodium between the filtrate and the intracellular fluid of the epithelial cells. This is accomplished by the *active transport of Na^+ ions* through the basolateral membrane of the epithelial cells (see Figure 21.4). Sodium is moved across this basolateral membrane and into the interstitial fluid surrounding the tubule by the *Na^+, K^+-ATPase pump*. As a result, the concentration of Na^+ ions within the epithelial cells is reduced, facilitating the diffusion of Na^+ ions into the cells across the luminal membrane. Potassium ions that are transported into the epithelial cells as a result of this pump either diffuse back into the interstitial fluid (proximal tubule and Loop of Henle) or into the tubular lumen for excretion in the urine (distal tubule and collecting duct).

The amount of sodium reabsorbed from the proximal tubule and from the Loop of Henle is held constant:

- Proximal tubule: 65% of the filtered sodium is reabsorbed.
- Ascending limb of the Loop of Henle: 25% of the filtered sodium is reabsorbed.

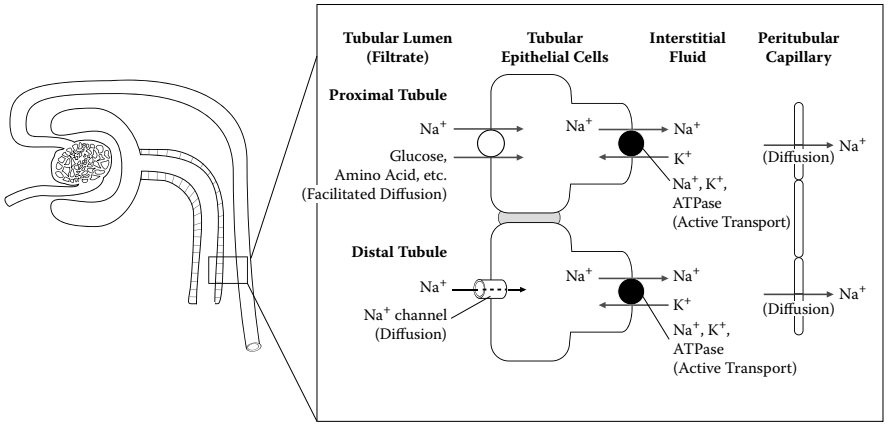


Figure 21.4 Tubular reabsorption of sodium. Sodium ions are actively transported out of the tubular epithelial cell through the basolateral membrane by the Na^+, K^+-ATPase pump. These ions then passively diffuse from the interstitial fluid into the blood of the peritubular capillaries. The active removal of Na^+ ions from the tubular epithelial cells establishes a concentration gradient for the passive diffusion of Na^+ ions into the cells from the tubular lumen. Potassium ions that are actively transported into the epithelial cells of the proximal tubule as the result of this process simply diffuse back into the interstitial fluid through channels located in the basolateral membrane. In the distal tubule and the collecting duct, the K^+ ions diffuse through channels in the luminal membrane into the tubular fluid and are excreted in the urine. The diffusion of sodium may be coupled to the reabsorption of organic molecules, such as glucose or amino acids, in the proximal tubule and the Loop of Henle. It may also occur through Na^+ channels in the distal tubule and the collecting duct.

This reabsorption occurs regardless of the sodium content of the body. In order to make adjustments in the *sodium load*, the reabsorption of the remaining 10% of the filtered Na^+ ions from the distal tubule and the collecting duct is physiologically controlled by two hormones: *aldosterone* and *atrial natriuretic peptide*.

Aldosterone, released from the adrenal cortex, *promotes the reabsorption of sodium* from the distal tubule and the collecting duct. The mechanisms of action of aldosterone include the following:

- Formation of Na^+ channels in the luminal membrane of the tubular epithelial cells (facilitates the passive diffusion of Na^+ ions into the cell).
- Formation of Na^+, K^+-ATPase carrier molecules in the basolateral membrane of the tubular epithelial cells (promotes the extrusion of Na^+ ions from the cells and their movement into the plasma by way of the peritubular capillaries; enhances the concentration gradient for passive diffusion through the Na^+ channels in the luminal membrane).

Atrial natriuretic peptide (*ANP*), released from myocardial cells in the atria of the heart, *inhibits the reabsorption of sodium* from the collecting duct. The mechanisms of action of ANP include:

- Inhibition of aldosterone secretion.
- Inhibition of Na^+ channels in the luminal membrane of the tubular epithelial cells.

Recall that the reabsorption of Na^+ ions is accompanied by the reabsorption of Cl^- ions, which diffuse down their electrical gradient, and by the reabsorption of water, which diffuses down its osmotic gradient. The net result is an expansion of plasma volume and, consequently, an increase in blood pressure. Therefore, the regulation of sodium reabsorption is important in the long-term regulation of blood pressure. As such, aldosterone and ANP, as well as the factors involved in their release, are discussed further in subsequent sections.

21.5.2 Chloride reabsorption

Chloride ions are reabsorbed passively according to the electrical gradient established by the active reabsorption of sodium. Chloride ions move from the tubular lumen back into the plasma by two pathways:

1. *Transcellular*: Through the tubular epithelial cells.
2. *Paracellular*: in between the tubular epithelial cells.

Most of the Cl^- ions diffuse between the tubular epithelial cells.

21.5.3 Water reabsorption

Water is reabsorbed passively by way of osmosis from many regions of the tubule. As with sodium and chloride, 65% of the filtered water is reabsorbed from the proximal tubule. An additional 15% of the filtered water is reabsorbed from the descending limb of the Loop of Henle. This reabsorption occurs regardless of the water content of the body. The water enters the tubular epithelial cells through *water channels*, also referred to as *aquaporins*. These channels are always open in the early regions of the tubule.

In order to make adjustments in the water load, the reabsorption of the remaining 20% of the filtered water from the distal tubule and the collecting duct is physiologically controlled by *antidiuretic hormone* (ADH), also referred to as *vasopressin*. Antidiuretic hormone, synthesized in the hypothalamus and released from the neurohypophysis of the pituitary gland, *promotes the reabsorption of water* from the distal tubule and the collecting duct. The mechanism of action of ADH involves an increase in the permeability of the water channels in the luminal membrane of the tubular epithelial cells.

Water diffuses into these cells and is ultimately reabsorbed back into the plasma by way of the peritubular capillaries.

Recall that the reabsorption of water is important in the regulation of plasma osmolarity. As the levels of ADH increase and more water is reabsorbed from the kidneys, then the plasma is diluted and plasma osmolarity decreases. Conversely, as the levels of ADH decrease and more water is lost in the urine, then the plasma becomes more concentrated and plasma osmolarity increases. Factors involved in the release of ADH are discussed further in subsequent sections.

21.5.4 Production of urine of varying concentrations

In order to effectively regulate plasma volume and osmolarity, the kidneys must be able to alter the volume and the concentration of the urine that is eliminated. Accordingly, the concentration of urine may be varied over a very wide range depending upon the body's level of hydration. The most dilute urine produced by the kidneys is 65 to 70 mOsm/L (when the body is overhydrated), and the most concentrated urine is 1200 mOsm/L (when the body is dehydrated). (Recall that the plasma osmolarity is 290 to 300 mOsm/L.)

An essential factor in the ability to excrete urine of varying concentrations is the presence of a *vertical osmotic gradient* in the medullary region of the kidney (see Figure 21.5). The osmolarity of the interstitial fluid in the cortical region of the kidney is about 300 mOsm/L. However, the osmolarity of the interstitial fluid in the medulla increases progressively, from 300 mOsm/L in the outer region of the medulla near the cortex to 1200 mOsm/L in the innermost region of the medulla. The increase in osmolarity is due to the accumulation of Na^+ ions and Cl^- ions in the interstitial fluid. This vertical osmotic gradient is created by the Loops of Henle of the juxtamedullary nephrons. Recall that the Loops of Henle in these nephrons penetrate deeply into the medulla. The gradient is then utilized by the collecting ducts, along with ADH, to alter the concentration of urine. The following is a summary of the reabsorption of sodium, chloride, and water by each region of the nephron.

Plasma is freely filtered from the *glomerulus*, such that everything in the plasma, except for the plasma proteins, is filtered. Therefore, the initial osmolarity of the filtrate is not different from that of the plasma and is about 300 mOsm/L (see Figure 21.5). Approximately 125 ml/min of the plasma is filtered. As the filtrate flows through the *proximal tubule*, 65% of the filtered Na^+ ions are actively reabsorbed, and 65% of the filtered Cl^- ions and water are passively reabsorbed. Because the water follows the sodium by way of osmosis, there is no change in the osmolarity of the filtrate; only a change in volume. At the end of the proximal tubule, approximately 44 ml of filtrate with an osmolarity of 300 mOsm/L remain in the tubule.

The *descending limb of the Loop of Henle* is permeable to water only. As this region of the tubule passes deeper into the medulla, water leaves the filtrate

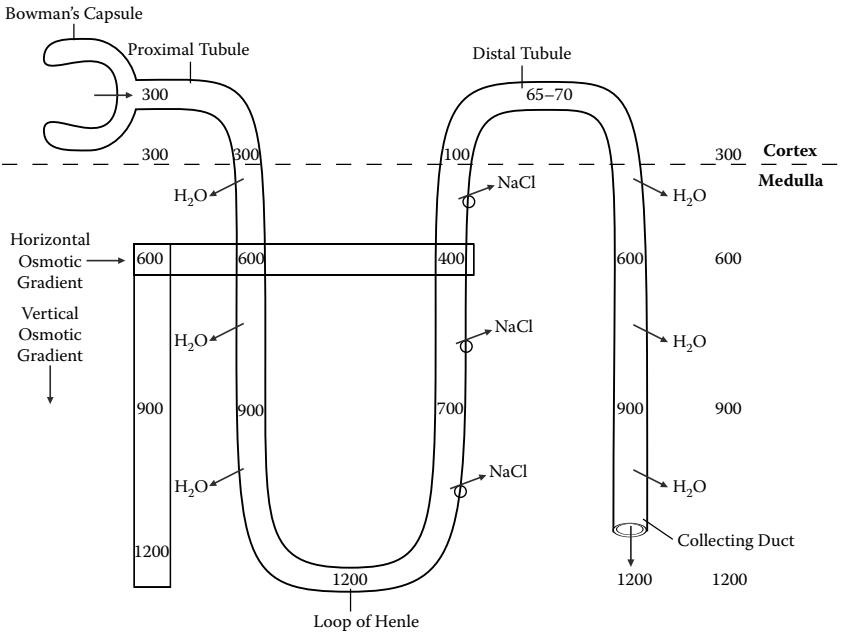


Figure 21.5 Production of urine of varying concentrations. The kidneys are capable of producing urine as dilute as 65 to 70 mOsm/L and as concentrated as 1200 mOsm/L. The concentration of the urine is determined by the body's level of hydration. Sodium ions are actively transported from the ascending limb of the Loop of Henle into the interstitial fluid. This active process is used to accumulate Na^+ ions and Cl^- ions in the medulla. As a result, a vertical osmotic gradient is established, where the interstitial fluid becomes increasingly concentrated. This gradient is necessary for the reabsorption of water from the collecting duct. Furthermore, a horizontal osmotic gradient of 200 mOsm/L is developed between the filtrate of the ascending limb of the Loop of Henle and the interstitial fluid. As a result, the osmolarity of the filtrate at the end of the Loop of Henle is 100 mOsm/L. Consequently, the kidney may now excrete a urine that is significantly more dilute than plasma. In this way, when the body is overhydrated, excess water is eliminated. The presence of aldosterone promotes additional reabsorption of Na^+ ions from the distal tubule and the collecting duct, which further dilutes the filtrate to 65 to 70 mOsm/L. The presence of ADH promotes the reabsorption of water from the distal tubule and the collecting duct. Water diffuses out of the collecting duct down its concentration gradient into the interstitial fluid. High levels of ADH may concentrate the filtrate to 1200 mOsm/L. In this way, when the body is dehydrated, water is conserved.

down its osmotic gradient until it equilibrates with the increasingly concentrated interstitial fluid (see Figure 21.5). As a result, the filtrate also becomes increasingly concentrated. At the tip of the Loop of Henle, the filtrate has an osmolarity of 1200 mOsm/L.

The *ascending limb of the Loop of Henle* is permeable to NaCl only. As the filtrate flows upward through this region of the tubule back toward the cortex,

Na^+ ions are continuously and actively pumped out of the filtrate and into the interstitial fluid. Chloride ions passively follow the sodium. As a result, the filtrate becomes increasingly dilute. At the end of the ascending limb of the Loop of Henle, approximately 25 ml of filtrate with an osmolarity of 100 mOsm/L remain in the tubule.

Because the transport of sodium is an active process, it is used to accumulate NaCl in the interstitial fluid of the medulla. In fact, this activity is involved in the initial establishment of the vertical osmotic gradient. Furthermore, sodium is actively transported out of the tubular epithelial cells up its concentration gradient until the filtrate is 200 mOsm/L less concentrated than the surrounding interstitial fluid. This difference between the filtrate and the interstitial fluid is referred to as the *horizontal osmotic gradient*. Because the filtrate at the end of the Loop of Henle has an osmolarity of 100 mOsm/L, the kidneys have the ability to produce urine that is significantly more dilute than the plasma.

As the filtrate progresses through the *distal tubule* and the *collecting duct*, the remaining NaCl (10% of that which was filtered) and water (20% of that which was filtered) are handled. As discussed, the presence of aldosterone enhances the reabsorption of sodium from these regions. As a result, the filtrate may become as dilute as 65 to 70 mOsm/L. The presence of ADH enhances the reabsorption of water from these regions. In particular, as the filtrate flows through the collecting duct, it enters a region of increasing osmolarity. The increased permeability of water due to ADH allows the water to diffuse out of the collecting duct and into the interstitial fluid down its concentration gradient. When the levels of ADH are high, the water may continue to leave the tubule until the filtrate equilibrates with the surrounding interstitial fluid. In this case, the filtrate may become as concentrated as 1200 mOsm/L, and a small volume of urine is produced. When the levels of ADH are low, water remains in the collecting duct and a large volume of urine is produced.

PHARMACY APPLICATION: PHYSIOLOGICAL ACTION OF DIURETICS

Diuretics are drugs that cause an increase in urine output. It is important to note that, except for the osmotic diuretics, these drugs typically enhance the excretion of both solute and water. Therefore, the net effect of most diuretics is to decrease plasma volume but cause little change in plasma osmolarity. Five classes of diuretics and their major sites of action are as follows:

1. *Osmotic diuretics*: Proximal tubule and descending limb of the Loop of Henle.
2. *Loop diuretics*: Ascending limb of the Loop of Henle.

3. *Thiazide diuretics*: Distal tubule.
4. *Potassium-sparing diuretics*: Cortical collecting duct.
5. *Carbonic anhydrase inhibitors*: Proximal tubule.

Osmotic diuretics, such as mannitol, act on the proximal tubule and, in particular, the descending limb of the Loop of Henle, portions of the tubule that are permeable to water. These drugs are freely filtered at the glomerulus but are not reabsorbed. Therefore, the drug remains in the tubular filtrate, increasing the osmolarity of this fluid. This increase in osmolarity keeps the water within the tubule, causing water diuresis. Because they primarily affect water and not sodium, the net effect is a reduction in total body water content more than cation content. Osmotic diuretics are poorly absorbed and must be administered intravenously. These drugs may be used to treat patients in acute renal failure and with dialysis disequilibrium syndrome. The latter disorder is caused by the excessively rapid removal of solutes from the extracellular fluid by hemodialysis.

Loop diuretics, such as furosemide, act on the ascending limb of the Loop of Henle, a portion of the tubule that is permeable to sodium and chloride. The mechanism of action of these diuretics involves the inhibition of the Na^+, K^+, $2Cl^-$ symporter in the luminal membrane. By inhibiting this transport mechanism, loop diuretics reduce the reabsorption of both NaCl and K^+ ions. Recall that the reabsorption of NaCl from the ascending limb of the Loop of Henle generates and maintains the vertical osmotic gradient in the medulla. Without the reabsorption of NaCl, then this gradient is diminished, and the osmolarity of the interstitial fluid in the medulla is decreased. When the osmolarity of the medulla is decreased, then the reabsorption of water from the descending limb of the Loop of Henle and the collecting duct is significantly reduced. The net result of the loop diuretics includes reduced NaCl and water reabsorption and, therefore, enhanced NaCl and water loss in the urine. The most potent diuretics available (up to 25% of the filtered Na^+ ions may be excreted), the loop diuretics, may cause hypovolemia. These drugs are often used to treat acute pulmonary edema, chronic congestive heart failure, and the edema and ascites of liver cirrhosis.

Thiazide diuretics, such as chlorothiazide, act on the distal tubule, a portion of the tubule that is permeable to sodium. The mechanism of action of these diuretics involves the inhibition of NaCl reabsorption by blocking the Na^+, Cl^- symporter in the luminal membrane. The thiazide diuretics are only moderately effective due to the location of their site of action. Approximately 90% of the filtered Na^+ ions have already been reabsorbed when the filtrate reaches the distal tubule. These drugs may be used for the treatment of edema associated with

heart, liver, and renal disease. Thiazide diuretics are also widely used for the treatment of hypertension.

Potassium-sparing diuretics act on the late portion of the distal tubule and on the cortical collecting duct. As a result of their site of action, the potassium-sparing diuretics also have a limited effect on diuresis compared to the loop diuretics (3% of the filtered Na^+ ions may be excreted). However, the clinical advantage of these drugs is that the reabsorption of K^+ ions is enhanced, reducing the risk of hypokalemia.

There are two types of potassium-sparing diuretics with different mechanisms of action. Agents of the first type, which include spironolactone, are also known as aldosterone antagonists. These drugs bind directly to the aldosterone receptor and prevent this hormone from exerting its effects. Agents of the second type, which include amiloride, are inhibitors of the tubular epithelial Na^+ channels. Acting on the Na^+ channels in the luminal membrane, these drugs prevent the movement of Na^+ ions from the filtrate into the epithelial cell. Because this transport of Na^+ ions into the cell is coupled to the transport of K^+ ions out of the cell, then less potassium is lost to the filtrate and, therefore, the urine.

Potassium-sparing diuretics are often coadministered with the thiazide or loop diuretics in the treatment of edema and hypertension. In this way, the edema fluid is lost to the urine while K^+ ion balance is better maintained. The aldosterone antagonists are particularly useful in the treatment of primary hyperaldosteronism.

Carbonic anhydrase inhibitors, such as acetazolamide, act in the proximal tubule. These drugs prevent the formation of H^+ ions, which are transported out of the tubular epithelial cell in exchange for Na^+ ions. These agents have limited clinical usefulness as they result in the development of metabolic acidosis.

21.6 Vasa recta

The *vasa recta* are modified peritubular capillaries. As with the peritubular capillaries, the vasa recta arise from the efferent arterioles. However, these vessels are associated only with the juxtamedullary nephrons and are found only in the medullary region of the kidney. The vasa recta pass straight through to the inner region of the medulla, form a hairpin loop, and return straight toward the cortex. This structure allows these vessels to lie parallel to the Loops of Henle and the collecting ducts.

The vasa recta perform several important functions:

- Provide oxygen and nourishment to the tubules of the medullary region of the kidneys.
- Return the NaCl and water reabsorbed from the Loops of Henle and the collecting ducts back to the general circulation.

- Deliver substances to the tubules for secretion.
- Maintain the vertical osmotic gradient within the interstitial fluid of the medulla.

Blood entering the vasa recta has an osmolarity of about 300 mOsm/L. As the vessels travel through the increasingly concentrated medulla, the osmolarity of the blood within them equilibrates with that of the surrounding interstitial fluid. In other words, the blood also becomes increasingly concentrated. Water leaves the vasa recta down its concentration gradient, and NaCl enters the vasa recta down its concentration gradient. Therefore, at the innermost region of the medulla, the osmolarity of the blood is 1200 mOsm/L. If the process were to be interrupted at this point, all of the NaCl that had initially created the vertical gradient would eventually be washed away, or removed from the medulla, by the blood flowing through it. However, like the Loops of Henle, the vasa recta form a hairpin loop and travel back toward the cortex through an increasingly dilute interstitial fluid. Once again, the osmolarity of the blood within them equilibrates with that of the surrounding interstitial fluid. In other words, the blood now becomes increasingly dilute. Water enters the vasa recta down its concentration gradient, and NaCl leaves the vasa recta down its concentration gradient. Consequently, when this blood has reached the cortex, its osmolarity has returned to 300 mOsm/L. Therefore, the blood leaving the vasa recta has an osmolarity similar to that of the blood that entered the vasa recta. What does change is the volume of blood that leaves the vasa recta. Once again, the excess NaCl and water reabsorbed from the tubules within the medulla have been picked up by these vessels and returned to the general circulation. It is important to note that this process has been performed without disrupting the vertical medullary gradient.

21.7 Tubular secretion

Tubular secretion is the transfer of substances from the peritubular capillaries into the renal tubule for excretion in the urine. This process is particularly important for the regulation of potassium and hydrogen ions in the body. It is also responsible for the removal of many organic compounds from the body. These may include metabolic wastes as well as foreign compounds, including drugs, such as penicillin. Most substances are secreted by secondary active transport.

21.7.1 Potassium ion secretion

Potassium ions are secreted in the distal tubule and the collecting duct. These ions diffuse down their concentration gradient from the peritubular capillaries into the interstitial fluid. They are then actively transported up their concentration gradient into the tubular epithelial cells by way of the Na^+, K^+

pump in the basolateral membrane. Finally, potassium ions exit the epithelial cells by passive diffusion through K^+ channels in the luminal membrane and enter the tubular fluid to be excreted in the urine.

Potassium secretion is enhanced by aldosterone. As the concentration of K^+ ions in the extracellular fluid increases, then the secretion of aldosterone from the adrenal cortex also increases. The mechanism of action of aldosterone involves an increase in the activity of the Na^+, K^+ pump in the basolateral membrane. Furthermore, aldosterone enhances the formation of K^+ channels in the luminal membrane.

21.7.2 Hydrogen ion secretion

Hydrogen ions are secreted in the proximal tubule, the distal tubule, and the collecting duct. The secretion of hydrogen ions is an important mechanism in acid–base balance. The normal pH of the arterial blood is 7.4. When the plasma becomes acidic, then H^+ ion secretion increases. Conversely, when the plasma becomes alkalotic, then the secretion of H^+ ions is reduced.

21.8 Plasma clearance

Plasma clearance is defined as the volume of plasma from which a substance is completely cleared by the kidneys per unit time (ml/min). The calculation of the plasma clearance of certain substances can be used to determine the volume of plasma filtered per minute (GFR) and the effective renal plasma flow (ERPF).

In order to measure the plasma clearance of a substance, the following variables must be determined:

- Rate of urine formation (V; ml/min).
- Concentration of the substance in the urine (U; mg/ml).
- Concentration of the substance in the arterial plasma (P; mg/ml).

The plasma clearance of a substance is calculated as follows:

$$\text{Plasma clearance} = \frac{\text{V (ml/min)} \times \text{U (mg/ml)}}{\text{P (mg/ml)}}$$

In order to use the *plasma clearance of a substance to determine GFR*, several criteria regarding the substance must be met. The substance must:

- Be freely filtered at the glomerulus.
- Not be reabsorbed.
- Not be secreted.

- Not be synthesized or broken down by the tubules.
- Not alter the GFR.

A substance that fulfills these criteria is *inulin*, a polysaccharide found in plants. Inulin is administered intravenously to a patient at a rate that results in a constant plasma concentration over the course of at least 1 hour. The urine is collected, and its volume and its concentration of inulin are measured.

Consider the following example where, at the end of 1 hour, 60 ml of urine are produced, the concentration of inulin in the urine is 20 mg/ml, and the concentration of inulin in the plasma is 0.16 mg/ml:

$$\text{Plasma clearance of insulin} = \frac{1\ \text{ml/min} \times 20\ \text{mg/ml}}{0.16\ \text{mg/ml}}$$

$$= 125\ \text{ml/min}$$

Because inulin is neither reabsorbed nor secreted, all of the inulin in the urine was filtered at the glomerulus. Therefore, the plasma clearance of inulin is equal to the GFR.

Although the measurement of GFR with inulin is quite accurate, it is inconvenient, because it requires the continuous infusion of this exogenous substance for several hours. More often, in clinical situations, the plasma clearance of *creatinine* is used to estimate GFR. Creatinine, an end-product of muscle metabolism, is released into the blood at a fairly constant rate. Consequently, only a single blood sample and a 24-hour urine collection are needed. The measurement of the plasma clearance of creatinine provides only an *estimate* of GFR. In fact, this measurement slightly overestimates GFR. A small amount of creatinine is secreted into the urine (about 10% on average). In other words, the concentration of creatinine in the urine is the result of the amount that is filtered (as determined by GFR) *plus* the small amount secreted.

In order to use the *plasma clearance of a substance to determine the effective rate of plasma flow* (ERPF) through the kidneys, several criteria regarding the substance must be met. The substance must:

- Be freely filtered at the glomerulus.
- Not be reabsorbed.
- Be secreted into the tubules.

A substance that fulfills these criteria is *para-aminohippuric acid* (PAH). All of the PAH that is not filtered at the glomerulus is secreted by the proximal tubule. The net effect is that all of the plasma that flows through the nephrons is completely cleared of PAH. It is important to note that about 10% to

15% of the total renal plasma flow supplies regions of the kidneys that are not involved with filtration or secretion. Consequently, this plasma cannot be cleared of PAH. Therefore, the plasma clearance of PAH provides a measurement of the *effective* renal plasma flow — that is, the volume of plasma that actually flows through the nephrons. The ERPF is normally about 625 ml/min. (This value is based on a renal blood flow of about 1.1 L/min and a hematocrit of about 42.)

The *filtration fraction* is the percent of the plasma flowing through the nephrons that is filtered into the tubules. It is calculated using the plasma clearance of inulin (GFR) and the plasma clearance of PAH (ERPF).

$$\text{Filtration fraction} = \frac{\text{GFR}}{\text{ERPF}}$$

$$= \frac{125 \text{ ml/min}}{625 \text{ ml/min}}$$

$$= 20\%$$

On average, 20% of the plasma that flows through the glomerulus is filtered into the tubules.

21.9 Renal blood flow

The kidneys receive a disproportionate fraction of the cardiac output (CO). Although the combined weight of the kidneys accounts for less than 1% of the total body weight, these organs receive 20% to 25% of the cardiac output. This magnitude of blood flow, which is in profound excess to their metabolic needs, enables them to carry out their multiple homeostatic functions more efficiently. Assuming a resting cardiac output of 5 L/min, the renal blood flow (RBF) is approximately 1.1 L/min.

Renal blood flow has a direct effect on GFR, which, in turn, has a direct effect on urine output. As RBF increases, then GFR increases and urine output increases. Conversely, as RBF decreases, the GFR decreases and urine output decreases. Furthermore, any change in urine output affects plasma volume and blood pressure. Therefore, the regulation of RBF and GFR is an important consideration.

According to Ohm's Law ($Q = \Delta P/R$), RBF is determined by mean arterial pressure (MAP) and the resistance of the afferent arteriole ($R_{\text{aff art}}$):

$$\text{RBF} = \frac{\text{Mean Arterial Pressure}}{R_{\text{afferent arteriole}}}$$

21.9.1 Autoregulation

The equation for RBF predicts that an increase in MAP will increase blood flow through the kidneys, and a decrease in MAP will decrease blood flow through the kidneys. Physiologically, this response is not always desired. For example, during exercise, MAP increases in order to increase blood flow to the working skeletal muscles. However, a corresponding increase in RBF would lead to an increase in GFR and an undesired loss of water and solutes in the urine. On the other hand, a profound decrease in MAP could decrease RBF and GFR. In this case, the elimination of wastes would be impaired. Therefore, there are physiological conditions in which maintaining a relatively constant RBF and GFR, even when MAP changes, is advantageous.

Interestingly, RBF remains rather constant when MAP changes in the range of 90 to 180 mmHg. This ability to maintain a constant blood flow in spite of changes in MAP is referred to as *autoregulation*. The mechanism of autoregulation involves corresponding changes in the resistance of the afferent arteriole. For example, when there is an increase in MAP, then the resistance of the afferent arteriole increases proportionately so that RBF remains unchanged. It is important to note that the major site of autoregulatory changes is the *afferent arteriole*. As this arteriole constricts, then the glomerular capillary pressure and, therefore, the GFR are reduced back toward their normal values.

Autoregulation of RBF is an *intrarenal response*. In other words, the mechanisms responsible for autoregulation function entirely within the kidney and rely on no external inputs. There are two mechanisms that elicit this response: the *myogenic mechanism* and the *tubuloglomerular feedback*.

21.9.2 Myogenic mechanism

As discussed in Chapter 17, the *myogenic mechanism* involves the contraction of vascular smooth muscle in response to stretch. For example, an increase in MAP would tend to increase RBF. This leads to an increase in pressure within the afferent arteriole and the distension, or stretch, of the vessel wall. Consequently, the vascular smooth muscle of the afferent arteriole contracts, increases the resistance of the vessel, and decreases RBF back toward normal.

21.9.3 Tubuloglomerular feedback

Tubuloglomerular feedback involves the activity of the *juxtaglomerular apparatus* (see Figure 21.1). This structure is located where the distal tubule comes into contact with the afferent and efferent arterioles, adjacent to the glomerulus. The juxtaglomerular apparatus is composed of the following: the *macula densa* and *granular cells*.

The *macula densa* consists of specialized cells of the distal tubule. These tubular cells are adapted to monitor GFR. In other words, they are sensitive to changes in the rate of filtrate flow through the distal tubule. *Granular cells* are specialized smooth muscle cells of the arterioles, in particular, the afferent arteriole. These cells are adapted to monitor RBF. In other words, they are sensitive to changes in blood flow and blood pressure in the afferent arteriole. As such, they are also referred to as *intrarenal baroreceptors*. The granular cells of the juxtaglomerular apparatus secrete renin. Further discussion of granular cell function can be found in Section 21.9.4.2. Tubuloglomerular feedback involves the function of the macula densa. This mechanism may be summarized with the following example where there is an increase in MAP:

↑ Mean arterial pressure

↓

↑ Renal blood flow

↓

↑ Glomerular capillary pressure

↓

↑ Glomerular filtration rate

↓

↑ Rate of fluid flow through the distal tubule

↓

↑ Reabsorption of Na^+ and Cl^- ions by the cells of the macula densa

↓

↑ Release of vasoconstrictor substances

↓

↑ Resistance of the afferent arteriole

↓

↓ *Renal blood flow to normal*

↓

↓ Glomerular capillary pressure

↓

↓ *Glomerular filtration rate to normal*

An increase in MAP leads to an increase in RBF, an increase in P_{GC}, and an increase in GFR. As a result, the rate of fluid flow through the distal tubule increases. This leads to an increase in the reabsorption of Na^+ and Cl^- ions by the cells of the macula densa in the distal tubule. Consequently, these cells release vasoconstrictor substances, primarily *adenosine*. The subsequent increase in the resistance of the nearby afferent arteriole decreases RBF back to normal. As a result, P_{GC} and, therefore, GFR decrease back to normal. In this way, the distal tubule regulates its own filtrate flow.

21.9.4 Resistance of the afferent arteriole

Many physiological conditions warrant a change in RBF and GFR, even when MAP is within the autoregulatory range. For example, volume overload is resolved with an increase in RBF, an increase in GFR, and an increase in urine output. In this way, excess water and solutes are eliminated. Conversely, volume depletion, such as occurs with hemorrhage or dehydration, is resolved with a decrease in RBF, a decrease in GFR, and a decrease in urine output. In this way, water and solutes are conserved.

The resistance of the afferent arteriole is influenced by several factors:

- Sympathetic nerves
- Angiotensin II
- Prostaglandins

21.9.4.1 Sympathetic nerves

The afferent and efferent arterioles are densely innervated by the sympathetic nervous system. Either norepinephrine, released directly from the nerves, or circulating epinephrine, released from the adrenal medulla, stimulate α_1 adrenergic receptors to cause vasoconstriction. The predominant site of regulation is the afferent arteriole. Under normal resting conditions, there is little sympathetic tone to these vessels so that RBF is comparatively high. As discussed previously, this facilitates glomerular filtration. However, the degree of sympathetic stimulation to the kidneys is altered under various physiological and pathophysiological conditions. For example, consider the case where an individual is volume depleted due to hemorrhage or dehydration:

↓ Plasma volume

↓

↓ Mean arterial pressure

↓

Baroreceptor reflex

↓

↑ Activity of sympathetic nervous system → [↑ CO, ↑ TPR → ↑ MAP]

↓

↑ Activity of renal sympathetic nerves

↓

↑ Resistance of the afferent arteriole

↓

↓ *Renal blood flow*

↓

↓ Glomerular capillary pressure

↓

↓ *Glomerular filtration rate*

↓

↓ Urine output

↓

↓ Renal excretion of sodium and water

↓

↑ *Plasma volume to normal*

↓

↑ Mean arterial pressure to normal

A loss of plasma volume leads to a decrease in MAP. Baroreceptors located in the aortic and carotid sinuses detect this fall in MAP and elicit reflex responses that include an increase in the overall activity of the sympathetic nervous system. Sympathetic stimulation of the heart and blood vessels leads to an increase in cardiac output (CO) and an increase in total peripheral resistance (TPR). These adjustments, which increase MAP, are responsible for the *short-term regulation of blood pressure*. Although increases in CO and TPR are effective in the temporary maintenance of MAP and blood flow to the vital organs, these activities cannot persist indefinitely. Ultimately, plasma volume must be returned to normal.

An overall increase in sympathetic nervous activity includes an increase in sympathetic input to the kidneys. Consequently, resistance of the afferent

arteriole increases, which leads to a decrease in RBF. As discussed, this results in a decrease in P_{GC}, a decrease in GFR, and a decrease in urine output. As such, the renal excretion of sodium and water is decreased. In other words, sodium and water are conserved by the body, which increases plasma volume and MAP back toward normal. These changes are responsible for the *long-term regulation of blood pressure*.

Sympathetic stimulation also increases the resistance of the efferent arteriole. This leads to a decrease in the blood pressure in the peritubular capillaries. This fall in pressure facilitates the movement of sodium and water from the tubules into these capillaries.

21.9.4.2 Angiotensin II

Angiotensin II also increases the resistance of the renal arterioles. Consequently, it decreases RBF and GFR.

Angiotensin II is synthesized by the following pathway:

Angiotensinogen

↓

Angiotensin I

↓

Angiotensin II

Angiotensinogen, which is synthesized by the liver, is an inactive plasma protein. *Renin*, a hormone secreted by the granular cells of the juxtaglomerular apparatus, promotes the conversion of circulating angiotensinogen into *angiotensin I*. As angiotensin I travels in the blood through the lungs, it is exposed to *angiotensin-converting enzyme* (ACE), which is located in the endothelial cells lining the blood vessels of the pulmonary circulation. This enzyme converts angiotensin I into *angiotensin II*.

Angiotensin II has multiple effects throughout the body, all of which directly or indirectly increase MAP. For example, angiotensin II causes the secretion of aldosterone, which enhances sodium reabsorption, expands plasma volume, and increases MAP. Angiotensin II is also a potent vasoactive substance that causes widespread vasoconstriction and, therefore, an increase in TPR, which increases MAP. Furthermore, angiotensin II causes powerful vasoconstriction in the renal arterioles, in particular, which leads to a decrease in RBF and GFR. Consequently, urine output is reduced, and water and solutes are conserved by the body. This leads to an increase in plasma volume and, therefore, MAP. Taken together, these effects demonstrate that the production of angiotensin II is beneficial when there has been a fall in blood pressure or blood volume.

The formation of angiotensin II requires the release of renin from the granular cells. Therefore, the following factors that affect renin release must be considered:

- Renal sympathetic nerves
- Intrarenal baroreceptors
- Macula densa
- Atrial natriuretic peptide
- Angiotensin II

The sympathetic nervous system increases blood pressure through multiple mechanisms including an increase in cardiac activity and vasoconstriction. Furthermore, stimulation of *β_1-adrenergic receptors* on the granular cells of the afferent arterioles, through the activity of *renal sympathetic nerves* or by circulating epinephrine, has a *direct stimulatory effect on renin secretion*. The enhanced formation of angiotensin II also increases blood pressure through the mechanisms listed above. Specifically, angiotensin II constricts the afferent arteriole, decreases RBF, decreases GFR, decreases urine output, and increases plasma volume and blood pressure. Conversely, a decrease in sympathetic activity results in a decrease in the secretion of renin.

The granular cells that secrete renin also serve as *intrarenal baroreceptors*. These cells monitor blood volume and blood pressure in the afferent arterioles. There is an inverse relationship between arteriolar pressure and renin secretion. In other words, an increase in blood volume causes an increase in arteriolar blood pressure, increased stimulation of the intrarenal baroreceptors, and *decreased secretion of renin*. With less angiotensin-II-induced vasoconstriction of the afferent arteriole, RBF, GFR, and urine output will all increase so that blood volume returns to normal.

The *macula densa*, which is involved in tubuloglomerular feedback, is also a factor in the regulation of renin secretion. In fact, this mechanism involving the macula densa is thought to be important in the maintenance of arterial blood pressure under conditions of decreased blood volume. For example, a decrease in blood volume leads to a decrease in RBF, a decrease in GFR, and a decrease in filtrate flow through the distal tubule. The resulting decrease in the delivery of NaCl to the macula densa *stimulates the secretion of renin*. Increased formation of angiotensin II serves to increase MAP and maintain blood flow to the tissues.

Atrial natriuretic peptide is released from myocardial cells in the atria of the heart in response to an increase in atrial filling, or an increase in plasma volume. This hormone *inhibits the release of renin*. With less angiotensin-II-induced vasoconstriction of the afferent arteriole, RBF, GFR, and urine output all increase. The increased loss of water and solutes decreases blood volume back toward normal.

Angiotensin II directly *inhibits the secretion of renin* from the granular cells. This negative feedback mechanism enables angiotensin II to limit its own formation.

21.9.4.3 Prostaglandins

The third important factor that influences the resistance of the afferent arterioles is the *prostaglandins*, specifically, PGE_2 and PGI_2. Produced by the kidney, these prostaglandins function as local *vasodilators* that decrease the resistance of the arterioles and increase RBF without changing GFR. Interestingly, the synthesis of PGE_2 and PGI_2 is stimulated by increased activity of the renal sympathetic nerves and by angiotensin II. The vasodilator prostaglandins then oppose the vasoconstrictor effects of norepinephrine and angiotensin II. The net result is a smaller increase in the resistance of the afferent arterioles. This "dampening" effect is important in that it prevents an excessive reduction in RBF which could lead to ischemia and potential damage of the renal tissues.

21.10 Control of sodium excretion: regulation of plasma volume

Sodium is the major extracellular cation. Because of its osmotic effects, changes in sodium content in the body have an important influence on extracellular fluid volume, including plasma volume. For example, excess sodium leads to the retention of water and an increase in plasma volume. An increase in plasma volume then causes an increase in blood pressure. Conversely, sodium deficit leads to water loss and a decrease in plasma volume. A decrease in plasma volume then causes a decrease in blood pressure. Therefore, homeostatic mechanisms involved in the regulation of plasma volume and blood pressure involve the regulation of sodium content, or *sodium balance*, in the body.

Sodium balance is achieved when salt intake is equal to salt output. The intake of salt in the average American diet (10 to 15 g/day) is far in excess of what is required physiologically. Only about 0.5 g/day of salt is lost in the sweat and the feces. The remaining ingested salt must be excreted in the urine.

The *amount of sodium excreted* by the renal system is determined by the amount of sodium filtered at the glomerulus and the amount of sodium reabsorbed from the tubules.

Sodium is freely filtered at the glomerulus. Therefore, any factor that affects GFR will also affect sodium filtration. As discussed previously, GFR is directly related to RBF. In turn, RBF is determined by blood pressure and the resistance of the afferent arteriole ($RBF = MAP/R_{aff\ art}$). For example, an increase in blood pressure or a decrease in the resistance of the afferent

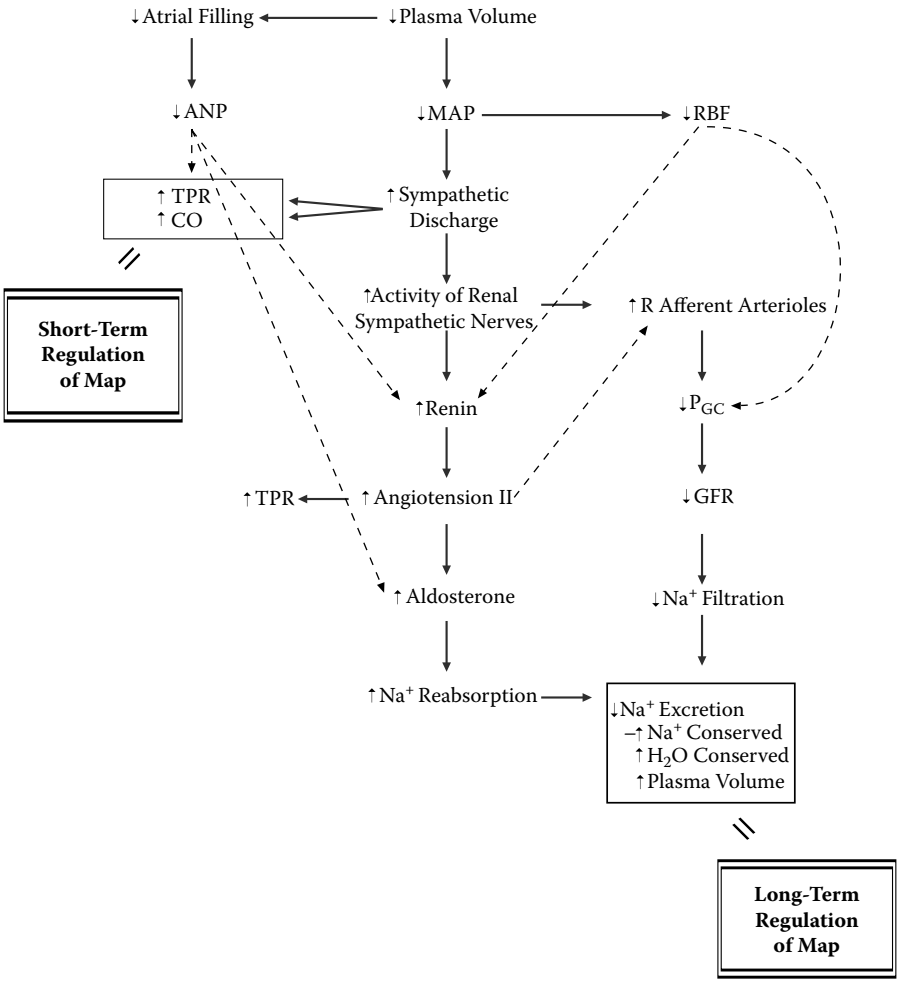


Figure 21.6 The renal handling of sodium.

arteriole will increase RBF, increase GFR, and, consequently, increase the filtration of sodium.

The *amount of sodium reabsorbed* from the tubules is physiologically regulated primarily by aldosterone and, to a lesser extent, by ANP. Aldosterone promotes reabsorption, and ANP inhibits it.

The alterations in sodium filtration and sodium reabsorption in response to a decrease in plasma volume are illustrated in Figure 21.6.

A decrease in plasma volume leads to a decrease in MAP, which is detected by the baroreceptors in the carotid sinuses and the arch of the aorta. By way of the vasomotor center, the baroreceptor reflex results in an overall increase in sympathetic nervous activity. This includes stimulation of the

heart and of vascular smooth muscle, which causes an increase in cardiac output and in total peripheral resistance. These changes are responsible for the short-term regulation of blood pressure, which temporarily increases MAP back toward normal.

Changes in sodium filtration and sodium reabsorption, which lead to a change in sodium excretion, are responsible for the long-term regulation of blood pressure. These changes are brought about by an increase in the activity of the renal sympathetic nerves. Sympathetic stimulation of α_1-adrenergic receptors on renal vascular smooth muscle leads to an increase in the resistance of the afferent arteriole. This causes a decrease in RBF, a decrease in glomerular filtration pressure, a decrease in GFR, and a *decrease in sodium filtration*. If less sodium is filtered, then less sodium is lost in the urine.

Sympathetic stimulation of β_1-adrenergic receptors on the granular cells of the juxtaglomerular apparatus promotes the secretion of renin and, consequently, the formation of angiotensin II. Angiotensin II then causes the following:

- Widespread vasoconstriction
- Increased resistance of the afferent arteriole
- Increased secretion of aldosterone

The widespread vasoconstriction supplements the increase in TPR induced by the sympathetic nervous system. Angiotensin II also causes vasoconstriction of the afferent arteriole, in particular, which enhances the decrease in RBF and sodium filtration. Finally, angiotensin II promotes the secretion of aldosterone from the adrenal cortex. Aldosterone then acts on the distal tubule and the collecting duct to increase *sodium reabsorption*.

Sodium reabsorption is also influenced by ANP. The original decrease in plasma volume leads to a decrease in atrial filling and a decrease in the release of ANP from the myocardium. Atrial natriuretic peptide, which acts on vascular smooth muscle, granular cells of the kidney, and the adrenal cortex, normally causes the following:

- Vasodilation
- Decreased renin release
- Decreased aldosterone secretion

Therefore, inhibition of ANP release leads to vasoconstriction and an increase in MAP. Furthermore, less ANP promotes the release of renin and the secretion of aldosterone, which further enhance sodium reabsorption.

Taken together, the homeostatic responses elicited by the initial decrease in plasma volume serve to decrease sodium filtration, increase sodium reabsorption, and, consequently, decrease sodium excretion in the urine. This conservation of sodium leads to the conservation of water and an expansion of plasma volume back toward normal.

21.11 Control of water excretion: regulation of plasma osmolarity

The regulation of the osmolarity of the extracellular fluid, including that of the plasma, is necessary in order to avoid osmotically induced changes in intracellular fluid volume. If the extracellular fluid were to become hypertonic (too concentrated), water would be pulled out of the cells. If the extracellular fluid were to become hypotonic (too dilute), water would enter the cells.

The osmolarity of the extracellular fluid is maintained at 290 mOsm/L by way of the physiological regulation of water excretion. As with sodium, *water balance* in the body is achieved when water intake is equal to water output. Sources of water input include the following:

- Fluid intake
- Water in food
- Metabolically produced water

Sources of water output include the following:

- Loss from the lungs and nonsweating skin
- Sweating
- Feces
- Urine

The two factors that are controlled physiologically in order to maintain water balance include *fluid intake* and *urine output*. *Fluid intake* is largely influenced by the subjective feeling of *thirst*, which compels an individual to ingest water or other fluids. Urine output is largely influenced by the action of ADH, which promotes the reabsorption of water from the distal tubule and the collecting duct. Thirst and ADH secretion are regulated by the hypothalamus. In fact, three functional regions of the hypothalamus are involved:

1. Osmoreceptors
2. Thirst center
3. ADH-secreting cells

The *osmoreceptors* of the hypothalamus monitor the osmolarity of the extracellular fluid. These receptors are stimulated primarily by an increase in *plasma osmolarity*. The osmoreceptors then provide excitatory inputs to the thirst center and the ADH-secreting cells in the hypothalamus. The threshold for ADH secretion is approximately 285 mOsm/L, and the threshold for the stimulation of thirst is approximately 295 mOsm/L. The stimulation of the thirst center leads to an increase of fluid intake. The stimulation of the ADH-secreting cells leads to the release of ADH from the neurohypophysis

and, ultimately, an increase in the reabsorption of water from the kidneys and a decrease in urine output. These effects increase the water content of the body and dilute the plasma back toward normal. Plasma osmolarity is the major stimulus for thirst and ADH secretion. Two other stimuli for thirst and ADH secretion include *decreased extracellular volume* and *angiotensin II*.

A more moderate stimulus for thirst and ADH secretion is a *decrease in extracellular fluid, or plasma volume*. This stimulus involves the low-pressure receptors in the atria of the heart as well as the baroreceptors in the large arteries. A decrease in plasma volume leads to a decrease in atrial filling, which is detected by the low-pressure receptors, and a decrease in MAP, which is detected by the baroreceptors. Each of these receptors then provides excitatory inputs to the thirst center and to the ADH-secreting cells.

Angiotensin II also stimulates the thirst center, to increase the urge to ingest fluids, and ADH secretion, to promote the reabsorption of water from the kidneys. Other factors that influence ADH-secreting cells (but not the thirst center) include pain, fear, and trauma, which increase ADH secretion, and alcohol, which decreases ADH secretion.

PHARMACY APPLICATION: DRUG-RELATED NEPHROPATHIES

Drug-related nephropathies involve functional or structural changes in the kidneys following the administration of certain drugs. The nephrons are subject to a high rate of exposure to substances in the blood due to the high rate of renal blood flow and the substantial glomerular filtration rate. Furthermore, the kidneys may be involved in the metabolic transformation of some drugs and are, therefore, exposed to their potentially toxic end-products.

Elderly patients are particularly susceptible to kidney damage due to their age-related decrease in renal function. Furthermore, the potential for nephrotoxicity is increased when two or more drugs that are capable of causing renal damage are administered at the same time.

Drugs and toxic end-products of drug metabolism may damage the kidneys by way of the following mechanisms:

- Decrease in renal blood flow
- Direct damage to the tubulointerstitial structures
- Hypersensitivity reactions

Nonsteroidal antiinflammatory drugs (NSAIDs) may damage renal structures, in particular, the interstitial cells of the medulla. Prostaglandins E_2 and I_2 are vasodilators that help to regulate renal blood flow under normal physiological conditions. Because NSAIDs inhibit the synthesis of prostaglandins, renal damage likely results from an inappropriate decrease in renal blood flow.

Chronic analgesic nephritis (inflammation of the nephrons) is associated with analgesic abuse. Ingredients such as aspirin and acetaminophen have been implicated in this disorder.

Acute drug-related hypersensitivity reactions (allergic responses) may cause tubulointerstitial nephritis, which will damage the tubules and the interstitium. These reactions are most commonly observed with the administration of methicillin and other synthetic antibiotics as well as furosemide and the thiazide diuretics. The onset of symptoms occurs in about 15 days. Symptoms include fever, eosinophilia, hematuria (blood in the urine), and proteinuria (proteins in the urine). Signs and symptoms of acute renal failure develop in about 50% of the cases. Discontinued use of the drug usually results in complete recovery. However, some patients, especially the elderly, may experience permanent renal damage.

Medical terminology

Acellular (ā-sĕl′ū-lăr): Not containing cells.
Aquaporin (ă-kwă-pŏr′ĭn): Water channel.
Basolateral (bā-sō-lăt′ĕr-ăl): Referring to the base and side of a cell.
Bowman's capsule (bō′măns căp′sūl): Initial portion of the renal tubule that receives the filtrate.
Cirrhosis (sĭ-rō′sĭs): Irreversible disease of the liver characterized by the loss of normal hepatic tissue and by scarring.
Diuretic (dī″ū-rĕt′ĭk): Substance that increases urine production.
Eosinophilia (ē″ō-sĭn-ō-fĭl′ē-ă): Abnormally high number of eosinophils.
Fenestrated (fĕn′ĕ-strāt-ĕd): Having openings.
Glomerulus (glō-mĕr′ū-lŭs): Capillary network in the kidney where filtration takes place.
Hemoturia (hē″mă-tū′rē-ă): Blood in the urine.
Hypertonic (hī″pĕr-tŏn′ĭk): Having a higher osmotic pressure.
Hypotonic (hī″pō-tŏn′ĭk): Having a lower osmotic pressure.
Juxtaglomerular (jŭks″tă-glō-mĕr′ū-lăr): Near the glomerulus.
Macula densa (măk′ū-lă dĕn′ză): Specialized cells of the distal tubule forming a portion of the juxtaglomerular apparatus.
Natriuretic (nā″trē-ūr-ĕt′ĭk): Substance that increases the excretion of sodium in the urine.
Nephritis (nĕf-rī′tĭs): Inflammation of the kidneys.
Nephron (nĕf′rŏn): Functional unit of the kidney.
Nephropathy (nĕ-frŏp′ă-thē): Disease of the kidney.
Nephrotoxin (nĕf″rō-tŏk′sĭn): Substance that damages the kidney.
Osmolarity (ŏs″mō-lăr′ĭ-tē): Concentration of osmotic particles in a solution.

Osmoreceptor (ŏz″mō-rē-sĕp′tŏr): Receptor sensitive to changes in osmotic pressure.
Podocyte (pŏd′ō-sīt): Specialized epithelial cell in Bowman's capsule.
Proteinuria (prō″tē-ĭn-ūr′ē-ă): Protein in the urine.
Tubuloglomerular (tū″bū-lō-glō-mĕr′ū-lăr): Referring to the tubule and the glomerulus of the nephron.
Ureter (ū′rĕ-tĕr): Tube that transports urine from the kidney to the bladder.
Urethra (ū-rē′thră): Tube that transports urine from the bladder to the outside of the body.
Vasa recta (vā′să rĕk′tă): Capillaries that lie parallel to the loops of Henle.

Bibliography

Costanzo, L., *Physiology*, 3rd ed., W.B. Saunders, Philadelphia, PA, 2006.
Fox, S. I., *Human Physiology,* 9th ed., McGraw-Hill, Boston, MA, 2006.
Guyton, A. C., and Hall, J. E., *Textbook of Medical Physiology,* 11th ed., Elsevier/Saunders, Philadelphia, PA, 2006.
Jackson, E. K., Diuretics, in *Goodman and Gilman's: The Pharmacological Basis of Therapeutics,* 11th ed., Brunton, L. L., Lazo, J. S., and Parker, K. L., Eds., McGraw-Hill, New York, 2006, chap. 28.
Koeppen, B. M., and Stanton, B. A., *Renal Physiology*, 4th ed., Mosby, St. Louis, MO, 2007.
Marieb, E. N., and Hoehn, K., *Anatomy and Physiology,* 3rd ed., Pearson Benjamin Cummings, San Francisco, CA, 2008.
Porth, C. M., *Pathophysiology, Concepts of Altered Health States*, 7th ed., Lippincott, Williams & Wilkins, Philadelphia, PA, 2005.
Sherwood, L., *Human Physiology from Cells to Systems*, 5th ed., Brooks/Cole, Pacific Grove, CA, 2004.
Silverthorn, D., *Human Physiology: An Integrated Approach*, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2007.
Stedman's Medical Dictionary for the Health Professions and Nursing, 5th ed., Lippincott Williams & Wilkins, Philadelphia, PA, 2005.
Taber's Cyclopedic Medical Dictionary, 20th ed., F.A. Davis, Philadelphia, PA, 2005.
Vander, A. J., *Renal Physiology*, 5th ed., McGraw-Hill, New York, 1995.
Widmaier, E. P., Raff, H., and Strang, K. T., *Vander's Human Physiology, The Mechanisms of Body Function,* McGraw-Hill, Boston, MA, 2006.

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Y

Z

Pharmacology and Toxicology

Essentials *of* Human Physiology *for* Pharmacy

A Comprehensive Guide to the Fundamentals of Physiology

A firm grasp of the functions of living organisms is one of the most important prerequisites to pharmacy study. The long-awaited second edition of Essentials of Human Physiology for Pharmacy presents concepts in physiology in a way that prepares students for their subsequent study of pathophysiology, pharmacology, and pharmacotherapeutics.

Thoroughly revised and updated, this text provides important new information to give students a solid grounding in some of the most important aspects of pharmacy. The book begins with an overview of the fundamental aspects of cell membrane physiology with particular emphasis on nerve cell function. This is followed by a detailed discussion of the two major regulatory systems in the body: the nervous system and the endocrine system.

The book then continues with in-depth presentations of the reproductive, muscular, cardiovascular, immune, respiratory, digestive, and renal systems. An important focus throughout the text is how tissue and organ function are regulated in order to maintain homeostasis.

New in This Edition:

- **All 19 chapters updated and expanded**
- **Revised and additional subsections of pharmacy applications**
- **New sections on medical terms with pronunciations and definitions**
- **New chapters on the reproductive and immune systems**
- **Additional figures and tables to clarify information**

Using simple, straightforward language, the text presents the science of physiology clearly so that undergraduate students will understand the relevance of these concepts as they relate to the practice of pharmacy. By studying and mastering the principles in this text, students will develop a secure foundation in human physiology so that they are better equipped to meet the challenges in their pharmacy career.

43900

CRC Press
Taylor & Francis Group
an Informa business
www.crcpress.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK

ISBN 978-1-4200-4390-7
90000
9 781420 043907

www.crcpress.com