Our Marvelous Bodies

Our Marvelous Bodies

An Introduction to the Physiology of Human Health

GARY F. MERRILL

RUTGERS UNIVERSITY PRESS
NEW BRUNSWICK, NEW JERSEY, AND LONDON

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Merrill, Gary F.

Our marvelous bodies : an introduction to the physiology of human health / Gary F. Merrill.

p. cm.

Includes bibliographical references and index.

ISBN-13: 978-0-8135-4281-2 (hardcover : alk. paper)

1. Human physiology.

[DNLM: 1. Physiological Processes. QT 104 M5710 2008] I. Title.

QP34.5.M487 2008

612—dc22 2007029660

A British Cataloging-in-Publication record for this book is available from the British Library.

Visit our Web site: http://rutgerspress.rutgers.edu

Manufactured in the United States of America

CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

I attended an international conference for biomedical scientists in San Diego, California, in the spring of 2005. The deaths of Pope John Paul II and Terry Schiavo made the headlines during that week. Life support had been terminated for Terry Schiavo two weeks before her death. She had suffered a major heart attack and slipped into a coma in 1990. Terry's husband Michael and her parents were bitterly divided over the prolongation of her life support. The Supreme Court was asked to intervene and revealed its division by voting 5 to 4 in favor of not reinstating life support.

Many years before the heart attack that eventually took the life of Terry Schiavo, I knew a family who faced a similar crisis in New Jersey. The mother in this family had gone to a local hospital for a cardiac procedure, lapsed into a coma, and was being sustained on life support. Her three children were adults and had families of their own. She had been married twice, and her current husband was not the father of her children. Two of the children and their families lived in distant states. I empathized as this group of distraught adults kept twenty-four-hour vigils wondering if their loved one would regain consciousness. They were receiving no guidance from attending physicians, did not know what questions to ask, and seemed completely ignored by the medical establishment. It was a lamentable scene of helplessness and confusion.

In private I mentioned to the husband that as a physiologist I had the medical background required to read and interpret medical records. I explained what physiology is and told him I could read his wife's records and apprise him of the status of her kidneys, GI tract, liver, heart, and lungs. He thought about my offer for a couple days and discussed it with her children. As a group they accepted my invitation, and he was given access to the records that he handed over to me. It was clear that my friend had been in multiple-organ failure from the outset of her hospitalization. She was on a mechanical ventilator, had a balloon-pump device implanted in her aorta to assist cardiac output, was on hyperalimentation (nutritional support), and was showing no signs of response to her family or the medical staff. After discussing her records, the family asked what I would do if this was my wife, my mother, my daughter, or my sister. My response was, "I would make

the same decision in each case. I would terminate life support now and would have done so earlier had I seen the records then." The family thought about my input for about twenty-four hours and made a united decision to terminate life support. I sensed immediately the relief they felt. There was no bitterness or regret then or during the months I kept in touch with the family after. I was able to help that family when physicians and others would not—and could not—because of my empathy for them and my background in physiology, the basis of human medicine.

When addressing the American Physiological Society in 1975 as its newest president, Arthur C. Guyton (1919–2003), a medical doctor, said, "What other person whether he be a theologian, a jurist, a doctor of medicine, a physicist, or whatever, knows more than you, a physiologist, about life? For physiology is indeed an explanation of life. What other subject matter is more fascinating, more exciting, more beautiful than the subject of life?" In 1951 Guyton was named one of the ten outstanding men in the nation. In his commemoration to Dr. Guyton, John E. Hall, PhD, wrote, "Arthur Guyton's research contributions, which include more than 600 papers and 40 books, are legendary and place him among the greatest physiologists in history."[1]

As a physiologist, I write this book to help, among others, students in advanced high school biology courses and in introductory college courses such as fundamentals of physiology as well as those in the allied health sciences, for example, respiratory therapists, x-ray technicians, occupational/physical/recreational health therapists, nurses-to-be, and physician's assistants. I write also for those who have recently entered professional fields of health and medicine and are beginning to care for patients. What I have to say about physiology as the basis of medicine and as a prerequisite to understanding one's health should help the student and practitioner gain a clearer understanding of our marvelous bodies and how they work. If someone knows how a thing works, it makes it easier for them to care for that thing. I believe this is true of the heart, the kidneys, and the entire human body. Perhaps with encouragement and information such as that offered here, students (and their patients) will take better care of their bodies, will require less health care intervention, and will live qualitatively more productive lives. Moreover, with this understanding the reader should be better informed when difficult decisions like those raised above have to be made.

Our Marvelous Bodies

1

The Foundation

For students receiving their initial exposure to the life sciences, physiology is the study of how living things work. It is the bedrock of the biomedical sciences. As the American Physiological Society expresses it, physiology is the science of life. Physiology is an analytical, experimental, investigative, and quantitative science. For the medical student completing an MD degree, or any student in the life sciences preparing to see patients, physiology is the basis of human medicine, historically and in the present. Each year a Nobel Prize is awarded in physiology or medicine. No other life science, past or present, has such a distinction.

The physiological approach to problem solving is the mechanistic approach. Physiologists use the words *mechanism* and *mechanistic* when they discuss the functions of living things. Mechanisms of function are studied by physiologists at the molecular, cellular, organ system, and whole animal levels. In the twenty-first century, the challenge for the physiologist is to study life integratively, for example, from the molecular to the organ systems levels. The modern physiologist is also encouraged to work translationally. In other words, if their research has relevance to modern human medicine, what happens in the laboratory must be quickly transferable to the clinic. Although the idea of translational physiology is relatively recent, one of the best examples occurred in the first two decades of the twentieth century when insulin was discovered. In only a matter of weeks between its isolation and purification, insulin was used in a diabetic human subject. From that first trial in a young man in Toronto, use of insulin had an immediate and global impact on human suffering from diabetes.

Structure and Function

Once the student understands what physiology is, it becomes easier, in many cases, to grasp the mechanistic approach by studying the relation of function to

structure. Examples abound of how structure and function are interrelated in the human body (see table 1.1).

Understanding the functions of the body is enhanced by first understanding corresponding structures. In addition, learning medically relevant scientific prefixes, suffixes, and some simple definitions will help both the student and the clinician with physiology. Anatomy, simply stated, is the structure or morphology of a tissue, organ, organ system, or whole animal. The morphology of a living thing can and does change with time and other conditions. Consider, for example, the morphology of a tadpole and contrast it with that of the mature frog. Alternatively, the primitive spermatogonium looks nothing like the mature spermatozoan, that is, the reproductively capable sperm cell. The spermatogonium lacks a tail and well-defined head with acrosome. Metamorphosis, such as that experienced by the tadpole or spermatogonium, happens because of development, differentiation, growth, and proliferation of cells, tissues, organs, and organ systems. Stated simply, anatomy is what we have and physiology is how we use it. Or, anatomy is the cut and chiseled appearance of the conditioned athlete's back and shoulders, biceps, and abdomen. Physiology is the mechanisms by which those same muscles develop force, shorten, and lift a load.

An ideal example of the relationship between structure and function is found in striated muscle. Muscle in human bodies can be subdivided into two broad categories, striated and nonstriated. Striated muscle is so named because when viewed under powerful transmission and scanning electron microscopes,

TABLE 1.1

Examples in human physiology where structure (anatomy, morphology) is related to function

Structure	*Function*
Striated muscle (skeletal, cardiac)	shortening and contraction, development of force, movement of a load
Nephron (kidney)	filters plasma, reabsorbs ultrafiltrate, secretes molecules, excretes urine
Alveolar airway	exchanges the respiratory gases oxygen and carbon dioxide with blood
Circulatory systems	manage/distribute flow (arteries, arterioles), exchange (capillaries), collect (veins)
Epiglottis	controlled by the swallowing reflex, minimizes/prevents choking

these cells have alternating light and dark bands called striations. Examples of striated muscles, among many others, include the biceps and the "six pack" muscles of the abdomen, as well as the ventricles and atria of the heart. Striated muscle is further divided into cardiac and noncardiac or skeletal. All other muscle types are considered nonstriated because they lack the pattern of alternating light and dark bands. One example of a nonstriated muscle would be the vascular smooth muscle that is found in the walls of most blood vessels. Figure 1.1 is an example of striated muscle as seen through an electron microscope at a magnification of 10,000 times.

Striated muscle, whether skeletal or cardiac, is characterized by the well-defined striations shown in figure 1.1. Many such striations occur inside a single muscle cell, also called a myocyte or muscle fiber. The myocyte's cell membrane is called the sarcolemma. The sarcolemma separates the interior contents of the cell from the extracellular fluids. Each muscle cell also has a dense concentration of mitochondria, the metabolic motors of the cell. The light and dark bands are composed of smaller units called thick and thin filaments. Thick filaments (dark bands) are polymers of the protein myosin and thin filaments (light bands) are made from actin chains. There are many other functional proteins associated with thick and thin filaments, but their description is beyond the purposes of this book.

The structural and functional unit of the striated muscle is the sarcomere, which is composed of any two adjacent dark lines running through corresponding light bands. Greater detail of the striations can be studied by further magnifying

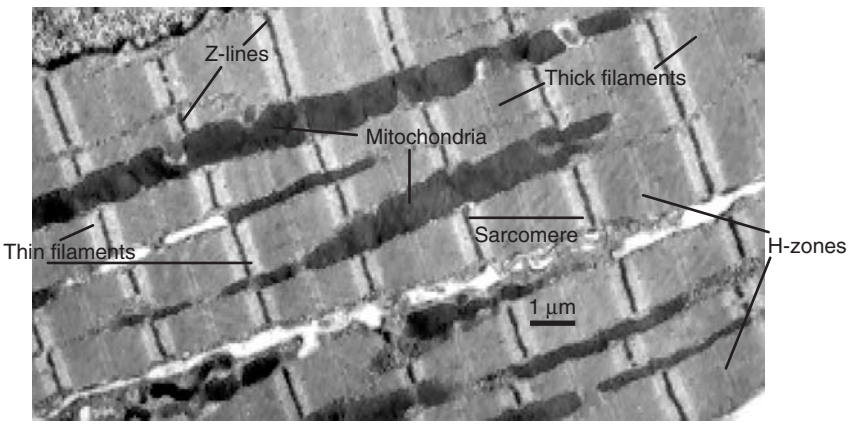

FIGURE 1.1 Electron micrograph (magnified about 10,000 times) showing cardiac striated muscle in the left ventricle of the guinea pig heart. Note the pattern of alternating light (with dark Z-lines running through the center) and dark bands (gray areas labeled thick filaments). Similar patterns are found in the hearts of all mammals, including humans. Note the sarcomeres.

(Merrill, unpublished data)

the structures shown in figure 1.1. We hypothesize that striated muscles contract and create force when the filaments slide over each other (we have not seen it occur, even though we can see the fibers in a microscope). Thin filaments, attached at adjacent Z-lines, slide towards each other and over interposed thick filaments. Thick and thin filaments do not change length, but the sarcomere shortens. Because of the structure and proximity of apposing thick and thin filaments, coupled with the physiology and biochemistry of the sarcomere, as filaments pass each other, energy is released, muscles shorten, force is generated, and a load is lifted (see figure 1.2).

A simple analogy of the sliding filament theory is interlacing the fingers of both hands. With fingers extended and the tips of index fingers touching each other, spread the fingers and then slide the apposing sets of fingers past one another about an inch, keeping the two index fingers fixed. With fingers in this position, note (1) the distance between the two palms, (2) the extent of overlap of the fingers, and (3) the thickness of the region of overlap versus the areas of the fingers that are not overlapped. The thick region corresponds to the area of a sarcomere where the thick and thin filaments overlap. The region of the fingers that do not overlap represent the non-overlapping thin filaments of the sarcomere, and where they are connected to the palms represents their points of attachment at the Z-lines. Now if you slide the fingers closer together, you can visualize how the sliding filament hypothesis of muscle contraction works.

Skeletal muscles are activated to contract by motor nerves that innervate them. A muscle fiber and its associated motor nerve is called a motor unit. The greater the number of motor units firing, the more force a muscle can generate. For example, a person lifting weights with 20-pound dumbbells in each hand will activate fewer motor units than they will using 40-pound dumbbells. The

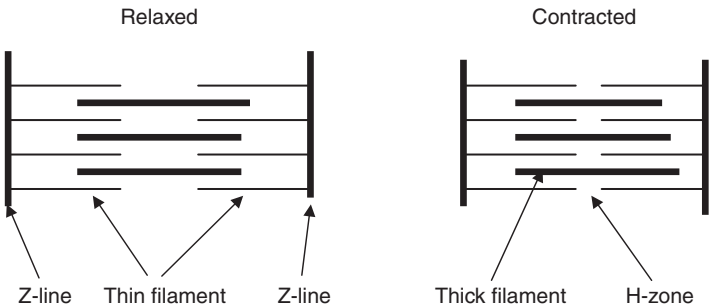

FIGURE 1.2 Relaxed (left) and contracted (right) sarcomeres. These illustrate the sliding-filament theory for muscle contraction (both cardiac and skeletal). Note the extent of overlap of thick and thin filaments under the two conditions. Note also that the lengths of the thick and thin filaments do not change during contraction (only sarcomere length changes).

heavier the weight, the greater the number of motor units recruited. Repeating such motion with progressively increasing weights during an eight- to twelve-week period will cause an increased size of each muscle cell in the affected region (*hypertrophy*). Some weight training (stretching) activities can also increase the number of cells (*hyperplasia*).

A second example of how understanding structure can enhance the understanding of function is the kidney. While the kidneys perform several important physiological functions, one of their main purposes is to filter the blood. In filtering the blood, potentially toxic waste products of metabolism are eliminated—for example, urea—and the balance of body water and electrolytes is maintained. The structural and functional unit of the mammalian kidney is the *nephron*. This word gives rise to the name of a clinician who studies kidney disease, a nephrologist. Nephrology is the study of renal disease. Figure 1.3 illustrates the mammalian nephron. Each of the two kidneys in the human body is constructed of a million plus nephrons. Each nephron is composed of vascular and tubular structures that are important to kidney function.

The general filtering function of the nephron is crudely analogous to the kitchen sink and its plumbing and purposes. The main function of a kitchen sink filled with hot water and detergent is to eliminate waste from glassware, utensils, and pots and pans. The dirty dishwater and its contents eventually get flushed into the kitchen plumbing. However, before they enter the plumbing,

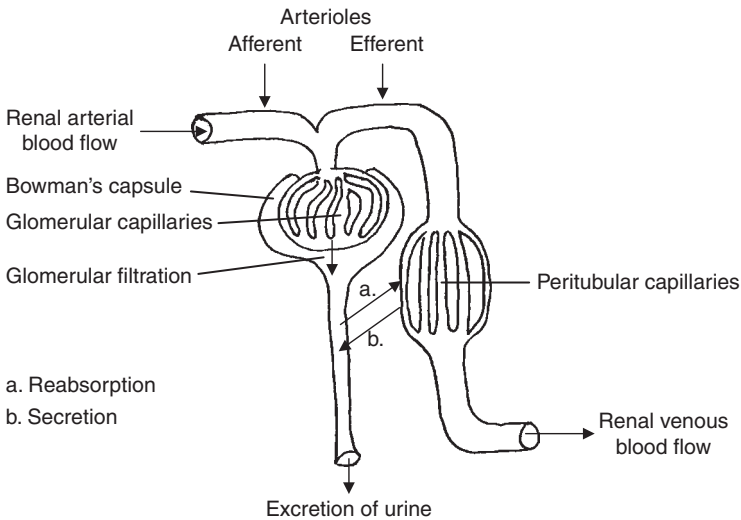

FIGURE 1.3 The mammalian nephron, the basic functional unit of the kidney. Note the basic renal functions of glomerular filtration, reabsorption, secretion, and excretion. Note also the proximities of afferent and efferent arterioles and glomerular and peritubular capillaries.

the waste products and dirty water must pass through a removable sink drain or filter. This drain is designed to allow water to flow across freely but to trap larger waste particles, which can be disposed of in the garbage.

Nephron function is similar. Each nephron has an arterial blood supply that is analogous to the faucet that supplies water to the sink. The nephron has a drain or filter called the glomerular filtration apparatus. It allows water and other small molecules to pass freely into the Bowman's capsule and the tubular nephron (sink gooseneck and house plumbing). The filtration barrier is composed of the wall of small blood vessels called glomerular capillaries. The capillaries have pores in their walls and the walls are overlaid by cells called podocytes. Together, the porous glomerular capillaries and overlaying podocytes form the glomerular filtration barrier. This barrier prevents large molecules from getting out of the renal blood and into the nephron tubules.

As renal arterial blood flows into the glomerular capillaries, water and its smallest solutes are forced from the blood vessels into the tubules of the nephron. Blood pressure and other factors such as the molecular radius of a particular solute and electrical charges on solutes and the filtration barrier determine which solutes get filtered and which ones remain inside the blood vessels. The total process is called glomerular filtration, and its rate can be determined experimentally and clinically. The glomerular filtration rate, abbreviated GFR, is one of the most important indicators of kidney function.

The filtered fluid and its nonaqueous component are referred to as ultrafiltrate, and once it is inside the renal tubules it will either be excreted as urine (renal excretion), or it will be reabsorbed into the capillaries surrounding the nephrons, a process called tubular reabsorption. On occasion, solutes get secreted out of the peritubular capillaries and into the renal tubules. For any given solute in the renal arterial blood, the net effects of filtration, reabsorption, and secretion ultimately determine the composition and volume of urine. As a rule, if a solute gets filtered and is not reabsorbed, it will be excreted as urine. Also, if a solute gets filtered but is completely reabsorbed, it will not appear in the urine. Finally, a solute can appear in the urine without having been filtered at the glomerular capillaries. In this case, that solute had to be secreted by the peritubular capillaries in order to appear in the urine.

One can check the physiological efficiency of these several processes by rapidly consuming a large quantity of water, say a pint or two in ten to fifteen minutes. If you are observant, you will note two things. First, the yellow hue of the urine before consuming the water was darker than it will be two or three hours after drinking. Secondly, the need to urinate after consuming the water will most likely be greater than before. This results as the kidneys reabsorb less water and try to restore body water volume and osmolarity to their physiological states. I will have much more to say about this in chapter 6 on renal function.

Homeostasis, Equilibrium, and the Steady State

In addition to understanding the differences between structure and function and how they interrelate, there are other fundamental physiological concepts that are key to body function. Among the most important of these are the concepts of homeostasis, equilibrium, and the physiological steady state. The term *homeostasis* was coined by the American physiologist Walter B. Cannon (1871–1945). In one of his books, *The Wisdom of the Body*, Cannon described the body's ability to maintain the status quo, and he applied that ability to all known functions of the tissues and organ systems. For contemporary physiologists, homeostasis is the maintenance of static or constant conditions in the internal environment.

Consider body temperature and its regulation. At any point in time, body temperature is the composite of the rates of heat production and heat loss combined with heat storage. When members of polar bear clubs in the northern hemisphere don their swimsuits and plunge into frigid waters on January 1 to ring in the New Year, homeostasis of body temperature is perturbed. While in the cold water, heat loss will outpace heat production, heat storage will decline, and body temperature will fall, even if only transiently. During the next several hours as swimmers dry off, replace their clothing, and move into a warm environment, physiological mechanisms will restore the balance among heat production, heat loss, and heat storage. When body temperature returns to preplunge levels, thermoregulatory homeostasis will have been restored.

The physiological concept of homeostasis suggests a basic mechanism for maintaining the stability of the internal milieu in the face of irregular nutrient, mineral, and water fluxes, as well as physical alterations in the environment. Homeostasis is the control of a vital parameter. The body carefully controls a seemingly endless list of vital parameters. Examples of tightly controlled parameters that affect nearly the whole body are arterial blood pressure and circulating blood volume. At the level of the internal fluids, tightly regulated parameters include body core temperature and plasma levels of oxygen, glucose, potassium ions (K^+), calcium ions (Ca^{2+}), and hydrogen ions (H^+). Homeostasis also occurs at the level of the single cell. Thus, cells regulate many of the same parameters that the body as a whole regulates: volume, the concentrations of many small inorganic ions, and energy levels such as adenosine triphosphate (ATP).

Another example of the concept of homeostasis is energy balance or maintenance of body weight. Simply stated, if the student or patient consumes more calories in a day than are expended, they are going to gain weight. Alternatively, if they expend more calories in a day than they consume, they will lose weight. If they weigh 175 pounds on day one, then consume and expend 2,000 calories per day through day thirty, they will still weigh 175 pounds on day thirty. The homeostasis of energy balance is the formula, over a lifetime, for maintaining body weight at a constant level.

Integral components of the concept of homeostasis are the ideas of equilibrium and steady states. To help explain equilibrium, imagine a washing machine, a rinse basin, and the outlet from the rinse basin into the plumbing system of a house. As the washing machine goes through a rinse cycle, it releases a large volume of rinse water into the basin. The water entering the basin exits into the house's plumbing system. If the plumbing system is clogged or restricted, and depending on the volume of the basin, the rate of entry of water into the basin exceeds the rate of exit from the basin. Under those conditions, the rinse water could overflow onto the laundry room floor. Conversely, if the plumbing system is not plugged, the rates of rinse water entering and exiting the basin are near equal and the level of water in the basin is near constant. At that precise moment in time, the two rates of water flow produce a volume of rinse water in the basin that does not change. At that moment, the entire system—washing machine, entry and exit of rinse water into the basin, and volume of water in the basin—is in a state of equilibrium. Happily, under such conditions there is no threat of the basin overfilling and the water overflowing onto the laundry room floor, that is, there is no arduous task to add to one's workload for the day.

This is the case for many of the functions of the human body. Consider the heart to be the physiological equivalent of the rinse basin. Venous blood flowing into the heart from various regions of the body, such as the head, trunk, internal organs, and limbs, is the counterpart of rinse water flowing into the basin. Arterial blood flowing away from the heart, to nourish tissues and organs of the head, trunk, and limbs, is analogous to rinse water draining the basin. If the volumes of blood flowing into and out of the heart are not equal, then either the heart or the organs will be depleted of or engorged with blood. Either condition is bad, and both cause an imbalance or disequilibrium in the cardiovascular system. Unless corrected promptly, either state can be lethal.

The heart and the vasculature, or blood vessels, acting in concert with other tissues and organ systems, have mechanisms to prevent such disequilibrium from either occurring or from lasting. This, however, can fail in disease states. Consider congestive heart failure as an example. The weakened and reduced capacity of the heart to pump blood is often caused by a heart attack, or myocardial infarction. As the heart expels less blood during each contraction, the volume of blood in the heart at the end of the following relaxation phase increases, that is, the heart becomes congested with blood. As congestion progresses, the pressures inside the heart and in the veins leading to the heart also change. This, in turn, affects arterial blood pressure. Thus, a state of disequilibrium in general circulatory hemodynamics occurs. In other words, cardiac and circulatory homeostasis is lost, the person's health deteriorates, and death can be the end result.

Like the person who consumed and expended 2,000 calories per day for a month, and whose body weight remained constant, the person's heart that ejects all the blood it receives per cycle will produce a circulation that is also in a state

of stability, the physiological steady state. Imagine yourself kicking back and reading the paper about 11 A.M. on a Sunday morning. Your homework is done, there are no midterm exams for a couple of weeks, and you are enjoying some leisure time. As you are comfortably relaxing on the couch, your respiratory rate, heart rate, body temperature, blood pressure (vital signs), and digestive functions are all in their baseline, resting conditions—their steady states. If we were to measure some of these variables, we might find your heart rate to be about seventy cycles per minute (beats per minute), your breathing rate to be about twelve cycles per minute, your body temperature to be about 38°C (98°F), and your mean arterial blood pressure to be about 100 mmHg. Now imagine an unexpected, loud knock at the front door and a frantic neighbor screaming, "My house is on fire, please help me!" The events of the next several minutes will arouse many of your physiological functions. If at the moment of greatest alarm we were to measure your vital signs, they would be elevated and perhaps even continuing to rise—they would not be in the steady state. After the fire and police departments arrive and have things under control, some time will have to pass before these physiological variables return to or near the stable conditions that prevailed before your neighbor's intrusion. During the time of alarm and increased physical activity, these physiological variables will operate under non-steady-state conditions.

Many events can disrupt the physiological steady state. These include but are not limited to periods of dynamic and static physical activity such as dancing, water skiing, and other forms of exercise; marked and/or sudden changes in environmental temperatures, such as would occur if you were marathoning in Death Valley or summiting Mount Everest; different states of consciousness such as during general anesthesia, sleep, and wakefulness; and fasting and feasting, periods of emotional distress and anger, anxiety over a misunderstood date and location of a final exam, and so forth. As students become familiar with physiology in general, they will be more cognizant of events that can influence the steady state.

Physiological Gradients

Another physiological concept of fundamental importance is that of gradients. The term is related to the words *grade* and *slope*. Any traveler driving from the east to the west coasts of the United States has crossed the Continental Divide and Rocky Mountains. Whether in Montana, Idaho, Wyoming, Utah, or Colorado, at some point the traveler will ascend and descend one or more mountain passes. Frequently the steepness of such passes is posted as a grade, for example, 6 percent or 7 percent. This means that for each 100 feet of linear distance forward, the traveler will rise or descend six or seven feet. Although plains and valley floors are already at about 5,000 feet, such as Denver, Colorado, the passes

may be at elevations in excess of 10,000 or 12,000 feet, for example Eisenhower Pass. Drivers of large vehicles such as buses and semis are advised to check their brakes before descending these mountainous passes, and the novice mountain traveler might notice the runaway truck and bus ramps. A gradient is a difference. In the case of Denver and Pike's Peak the difference is about 9,000 feet.

Blood pressure is an example of a physiologically important gradient. Blood flows from regions of high pressure to regions of low pressure. That is to say, blood flow occurs, in part, as the result of pressure gradients. In our reclining newspaper reader above, the average pressure in the left ventricle at the peak of its contraction is about 100 mmHg under resting conditions. Millimeters of mercury are the units of measure used to express blood pressure. Blood pressure levels are similar in the major arteries carrying blood away from the left ventricle during systole. At the peak of excitement during the neighbor's house fire, left ventricular peak blood pressure will be elevated considerably. Blood leaving the left side of the heart eventually ends up in the right side of the heart at the right atrium. From there it goes to the right ventricle en route to the lungs. The average pressure in the right atrium of the reader at rest is less than 5 mmHg; blood flows because of an intravascular pressure gradient of $100 - 5 = 95$ mmHg. Conditions that increase this gradient (house fire) will increase blood flow, and those that decrease the gradient will decrease it.

Another physiologically important gradient is the osmotic gradient. Osmosis is the process by which charged particles such as ions influence the distribution of water between different body compartments. Ions, electrolytes, salts, and other osmotically active chemicals are those that have the ability to attract water across permeable membranes such as cell walls. Osmotically active chemicals exist in three important locations in the body: (1) inside cells, that is, the intracellular space or intracellular compartment; (2) inside blood vessels, for example, the intravascular space or intravascular compartment, an extracellular compartment; and (3) the interstitial space, also an extracellular compartment. *Interstitial* means "between the tissues" and refers to space within the body and its organ systems that is both between the cells (but outside of cells) and outside of the vascular compartment. Under homeostatic conditions, there is a balance in the amount or concentrations of osmotically active chemicals in these three spaces. Because of their abilities to attract water (osmosis), the distribution of these chemicals causes an equally important distribution of water among the three compartments. This mechanism is crucial since the balance of water and salts among the body's various compartments is critically important to the health and well-being of the individual. When either the water or osmotic gradient is upset—for example, when there is a higher concentration of osmotically active chemicals in the interstitial space than in the intracellular space—the other space is disrupted. Unless the imbalance is promptly restored by

homeostatic mechanisms, net losses or gains in water and electrolytes will occur in the various compartments leading to injury and death of the tissues.

Anyone whose job requires them to stand most of the day has experienced swelling and discomfort in the feet. This is especially true in a hot environment, for example, standing on a concrete slab or asphalt during hot summer months. The swelling results from imbalances in osmolality and volume in the water compartments of the tissues of the feet. Periodically elevating the feet above the heart, even during the workday, can do wonders for the swelling and discomfort. Stress such as standing for long hours multiple days per week over years can do irreversible damage to the walls of blood vessels not only in the feet but in vessels of the lower leg as well, leading to conditions such as varicose veins and phlebitis. I will have more to say about this when I discuss kidney function later in the book (see chapter 6, "Kidneys and Renal Physiology").

We acquire oxygen from the atmosphere and eliminate carbon dioxide into it also as a result of gradients. The amount of oxygen in any particular space can be expressed by its partial pressure (the amount of energy or force caused by oxygen molecules as they collide with one another and with the walls of containers confining them). Physiologists express the partial pressure of oxygen with the symbol PO_2. The upper-case P stands for partial pressure and O_2 is the chemist's symbol for molecular oxygen. The greatest partial pressure of oxygen is in the atmosphere we breathe. At sea level it is about 150 mmHg. The next greatest PO_2 is in the lungs (about 105 mmHg), then in the venous blood leaving the lungs (blood that will become arterialized—the addition of oxygen to venous blood—once it reaches the left ventricle and major arteries and is pressurized to about 100 mmHg), then inside the cell (about 40 mmHg), finally in subcellular organelles (structures such as nuclei and mitochondria that are located inside cells), where the pressure is less than 40 mmHg). Thus, there is an oxygen gradient from atmosphere to subcellular organelle of more than 110 mmHg that ensures constant delivery of oxygen to where it is needed. An opposite gradient exists for carbon dioxide, a by-product of cellular metabolism that must not accumulate in the tissues. If carbon dioxide accumulates, especially in brain and other sensitive tissues, it can lead to acidosis, comatosis, and even death. Partial pressure of carbon dioxide in the cells is greater than 45 mmHg and in the lungs is less than 40 mmHg. So, for both these physiologically important, life-sustaining gases, gradients exist that ensure continuous uptake of oxygen and removal of carbon dioxide.

Physiological Reflexes

In their earliest clinical training, students become familiar with the concept of body reflexes. They are taught some very simple reflexes such as the knee-jerk reflex, but, unless enrolled in advanced physiology courses, they might not be

taught the more complex reflexes such as the contralateral flexor/extensor reflexes and reflex movements of the gut wall during and after ingestion of a meal. Structurally, all reflex arcs consist of five components. They are (1) a sensory receptor, (2) a sensory or afferent nerve, (3) a central integrator, (4) a motor or efferent nerve, and (5) a motor effector or activator (see figure 1.4). Unless all five components are involved, it cannot correctly be said that a function has occurred "reflexively."

A sensory receptor is generally thought of as a cell (or group of specialized cells) that is sensitive to changes in its environment. The environmental stimuli that activate these receptors can be mechanical, electrical, thermal, or chemical, and most receptors respond to specific stimuli. For example, receptors that respond to heat and cold (thermoreceptors) do not respond to mechanical changes in their environments. Environmental stimuli refer to events outside the body, at the body surface, or inside the body. In the case of the knee-jerk reflex, the stimulus is mechanical, the deformation or changes in tissue contour caused by a hard object striking the knee. The knee-jerk mechanoreceptors are innervated (attached to sensory nerves) by neurons (nerve cells) that carry the mechanical signal, now transformed into an electrical impulse, centrally or afferently into the spinal cord. Inside the spinal cord, the signal is evaluated and integrated by interneurons (nerve cells interposed between the afferent, sensory nerves and the efferent, motor nerves, that is, the central integrator) that direct the outflow of information to the motor or efferent nerves. Finally, motor nerves carry another electrical impulse to motor effectors or activators. In the case of the knee jerk, the activators are extensor muscles that are stimulated (cause extension of the limb) and flexor muscles that are inhibited (do not oppose extension of the limb). The end result of this reflex is extension and elevation of the leg below the knee.

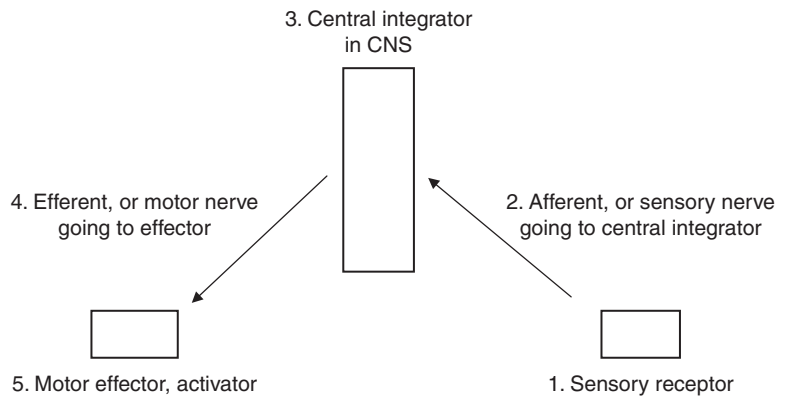

FIGURE 1.4 The five basic components of a mammalian reflex arc. Note that the physiological mechanism of the reflex begins with a sensory receptor (1) and ends with a motor effector (5).

Reflex responses can aid the well-being of the organism, as demonstrated by the nociceptor reflex. Nociceptors are sensory receptors that respond to painful stimuli. Heat can be a painful stimulus. Consider small children who place their hands on hot stoves. Pain receptors in a hand are activated and initiate reflexes that both cause the withdrawal of the hand and the child's removal from the stove. Even though the contact lasted less than one second, the child still suffered a burn. Imagine the damage that would be sustained if thermoreceptors were not present and the child's hand had remained on the stove for several seconds. Examples of other reflexes, such as occur when stepping on a sharp object with a bare foot, should now be more easily understood.

Control Systems Analysis

Physiologists sometimes think and act like engineers. This is especially true when it comes to analyzing functions of the body that are regulated by "control systems." Indeed these systems are similar to and even involve components of the reflex arc. Figure 1.5 illustrates, in simplest terms, the standard physiological control system. The controller is an anatomical structure of the central nervous system, and in the case of the human cardiovascular system, the controller is located in the brain stem and is subdivided into several more discrete units. These are the vasomotor and cardiotonic centers. The vasomotor centers cause adjustments in vascular tone—the degree to which an artery, arteriole, or vein is constricted or dilated. The cardiotonic centers cause acceleration and deceleration of the heart

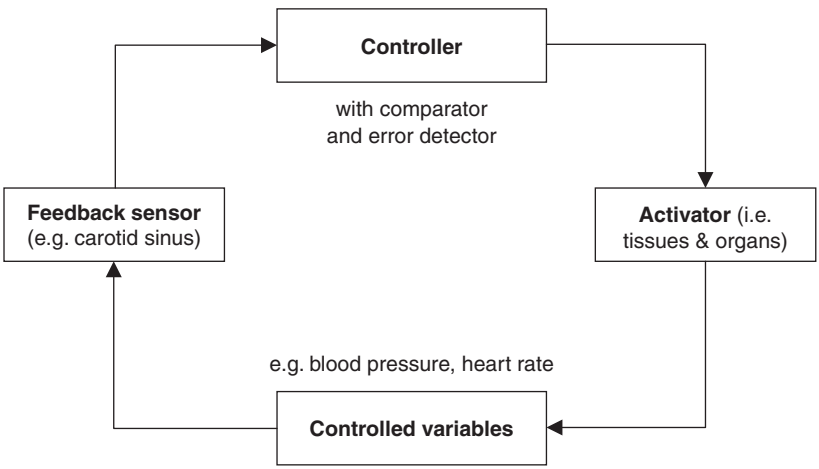

FIGURE 1.5 The physiological mechanism by which closed-loop feedback control systems work. Note that the physiological output of each component becomes the input to a subsequent component. See text for an example of how such a system behaves in controlling blood pressure.

rate, as well as increments and decrements in contractile vigor of the heart. The vasomotor center is further subdivided into vasopressor (causes elevations in blood pressure) and vasodepressor (causes decreases in blood pressure) regions.

The brain stem is a segment of the central nervous system that connects the cerebral hemispheres, spinal cord, and other components of the central nervous system (for example, the cerebellum and midbrain). The brain stem is critically important to the control and regulation of all the major organ systems of the body. Controllers within the brain stem consist of other subcomponents such as *comparators*. Comparators assess feedback signals with preexisting "set points" in the controller. Any differences between feedback signals and set points are seen as error signals and are adjusted by the entire control system. This is done to eliminate the error signal and to reestablish the set point. Output from the controller goes to "activators" (organs and tissues) and become the activator's inputs. The output of an activator is a controlled variable. Examples of controlled variables using a controller mechanism are blood pressure, heart rate, and body temperature. Controlled variables are detected by feedback sensors that send that information directly to the controller where it is compared (by the comparator) with the set point.

One example of a control system is the human cardiovascular system. Consider the short-term regulation of arterial blood pressure. Imagine that your average arterial blood pressure is 100 mmHg and that it has been at this level for several years. This means the set point in your brainstem controller is near 100 mmHg, and this value is what the comparator has been accustomed to detecting for several years. Now imagine that you are nervously seated in your dentist's chair in preparation for a root canal. Through emotionally evoked physiological mechanisms, the thought of a root canal has elevated your mean arterial blood pressure to 150 mmHg, a 50 percent increase. It has also elevated your heart rate. These changes are not good for the body, and if sustained could become pathological. For example, sustained elevated heart rates lead to excess heart work as evidenced by increased oxygen consumption. Prolonged elevation of blood pressure (hypertension) at 150 mmHg or more damages the endothelial lining of blood vessels making it easier for atherosclerosis (hardening of the arteries) to develop. That will further increase blood pressure. Moreover, the higher the blood pressure, the harder the heart has to work to circulate blood. This places undue strain on the heart and chronically predisposes to hypertrophy and failure. While transient elevations of blood pressure (dentist's chair, root canal) do little lasting damage, progressive and sustained hypertension is a leading cause of disability and death in industrialized nations. Transient changes in blood pressure and heart rate can also contribute to dangerous arrhythmias, myocardial infarction, and even sudden death.

A feedback sensor called the *carotid baroreceptor* is located in the wall of the internal carotid artery. This is located just downstream from the point of

branching of the common (main) carotid artery into the internal and external carotid arteries. This critical region of vasculature is located near the bottom of the ear and curvature of the jaw. It is just above the region of the neck where we see EMT specialists, doctors, and nurses place their fingers to detect carotid pulses in injured persons. Once the feedback information from the carotid baroreceptor arrives at the cardiovascular control center, the comparator evaluates the difference between the set point of 100 mmHg and the new feedback of 150 mmHg, and physiological adjustments are made within the activators to bring arterial pressure back to 100 mmHg. Because of physiological variability among dental patients, the above adjustments in the control system might be made while they are still in the dentist's chair. In others, the corrective actions will take several hours to complete.

Feedback and Feedforward

Control systems in humans and other mammals act primarily by negative feedback. Negative feedback is one mechanism that contributes to homeostatic maintenance of physiological variables. Consider the regulation of the concentrations of carbon dioxide in body fluids. If the concentration of carbon dioxide in the blood increases, then pulmonary ventilation will be stimulated. The increase in ventilation will cause a greater release of carbon dioxide to the atmosphere. This will reduce the concentration of carbon dioxide in the blood. In other words, an elevated concentration of carbon dioxide in the blood ultimately leads to a lower concentration of carbon dioxide, which is negative to the initiating stimulus. Conversely, if the concentration of carbon dioxide in the blood falls too low, it will impede respiration causing a feedback increase in the blood concentration of the gas. This response also is negative to the initiating event.

In the regulation of heart rate and blood pressure, elevations in either variable cause a series of events via the control systems described above that lead to reductions in heart rate and blood pressure. In both cases, the final outcome is negative to the initiating stimuli. Therefore, if some physiological variable becomes excessive or deficient, a control system initiates negative feedback, which consists of a series of changes that restores the variable to its original levels.

Positive feedback also occurs but commonly is deleterious and even lethal. Whereas negative feedback typically leads to physiological stability and maintenance of the status quo (protection of homeostasis), positive feedback leads to instability through what some call vicious cycles. Consider again the feedback regulation of arterial blood pressure. If an increase in pressure occurs, a positive feedback system would produce events that lead to a further increase in arterial pressure: an errant change in blood pressure of 100 to 150 mmHg could through positive feedback lead to an arterial blood pressure of 175 or 200 mmHg or greater. Such changes would be dangerous and, if uncorrected, fatal.

TABLE 1.2

Basic concepts of the physiology of human health

Concept	*Principle*
1. Structures and functions	Understanding structure can enhance learning functions
2. Homeostasis	Tendency to stability; the principle that systems are designed to operate at constant levels
3. Equilibrium	The principle that input and output balance one another at a level that is consistent with life and good health; maintaining the *status quo*
4. Steady state	Dynamic equilibrium; conditions when a system operates at a constant level
5. Gradients	Differences, e.g., in concentrations and pressures as a function of time and space
6. Reflexes	Interrelated functions of the five basic components of an involuntary physiological system designed to achieve specific actions
7. Feedback/feedforward	Basic principles of control systems (closed-loop) designed to maintain homeostasis of controlled variables (e.g., blood pressure, body temperature)

A third kind of control called feedforward also occurs in the human body. For example, some movements of the body occur so rapidly that there is insufficient time for peripheral (distantly located tissues relative to the central nervous system) sensory stimuli to travel all the way to the central nervous system, become integrated, and relay motor responses to the affected muscles in time to control a particular movement. In such cases, the brain uses feedforward control to cause the required muscle movements. Sensory signals from the moving parts apprise the brain in retrospect whether the appropriate movement as envisaged by the brain has been performed correctly. If not, the brain corrects the feedforward signals it sends to the muscles the next time the movement is required. If still further corrections must be made, they will be done for subsequent movements. This kind of feedforward activity by the body has also been referred to as adaptive control.

THIS CHAPTER HAS SUMMARIZED some of the fundamental principles of physiology. The first task is to understand the interrelationships between structure and

function of living tissue. Then the concepts of homeostasis, equilibrium, and steady states were discussed. The importance of physiological gradients and reflexes and control systems and feedback were also discussed. While this is by no means an exhaustive list of fundamental physiological principles, it is nonetheless an important list. These principles will help the student and practitioner understand our marvelous bodies and the physiological basis of health and wholeness. For the reader who is interested in a more in-depth understanding of physiology, please see the suggested readings at the end of this and all chapters. The textbooks I have suggested are among the best and are used for teaching the physiology of medicine in medical schools in the United States and elsewhere.

Table 1.2 summarizes some of the important physiological concepts presented in this chapter.

2

Understanding the Mammalian Nervous System

What Are Neurons?

Neurons, or brain cells, come in multiple shapes and sizes. Their common purpose is to communicate. Neurons consist of three basic components that enable them to communicate. These include the cell body or *soma*, an *axon* or cable that extends to an adjacent neuron or other effector, and *dendrites*. Dendrites are shaft-like projections that arise from the cell body and make contact with multiple other neurons in the vicinity of the original cell. Physiologically, dendrites receive incoming information. That information is integrated and analyzed in the cell body. A response is then conducted down the axon to another neuron or some other effector such as a skeletal muscle cell. Neurons are the functional units of all the body's nervous systems.

No one knows how many neurons are present in the human body. In late 2006, I heard a mammalian neuroscientist say that in one cubic millimeter of tissue from the rat cerebral cortex, there are more neurons than there are stars in the universe. I don't know how he arrived at this figure. However, as the universe is composed mainly of cold, dark, matter—as much as 70 percent as estimated by some astrophysicists—so is the human nervous system composed of cells and tissues other than neurons. One of the most abundant of these is the glial cell. Glial cells are thought to be supportive to neurons. They influence blood flow, provide structural support, influence remodeling after disease and damage, and are thought to contribute to the neuron's communicative/metabolic functions.

Neurons communicate via two processes. The first is called *conduction* and the second *transmission*. I will say more about each later, but for now conduction refers to the physiological processes involved in getting a message from the cell body, down the axon, and to the effector.

Nerve axons can be myelinated or nonmyelinated. Myelin is an insulation-like fatty sheath that encases the axon but is disrupted at regular intervals along

its length. Myelinated nerves conduct action potentials with greater velocity than do nonmyelinated axons. This is because the physiological mechanisms that account for conduction in a myelinated neuron occur only at the internodes (the space between adjacent myelinated regions of the axon, which are called nodes). This kind of action-potential-mediated communication is called saltatory conduction. Nonmyelinated axons conduct more slowly because the events of the action potential occur continuously over the entire axon and not just at the internodes.

The process of transmission is confined to the synapse. The synapse is the gap between the terminus of the axon and the activator that it innervates. Transmission takes place either electrically or chemically. Chemical transmission relies on neurotransmitters. These are molecules such as serotonin and norepinephrine that are stored in vesicles in the terminus of the axon, a location referred to as the presynaptic junction. When a signal arrives from the cell body, it causes the release of neurotransmitters into the synapse (also called the synaptic cleft or synaptic junction). By diffusion, these molecules cross the synaptic cleft and bind to protein receptors on the activator or postsynaptic junction. The specific site of binding is called the postsynaptic receptor or postsynaptic docking site.

Functionally, there are many kinds of neurons. Some have axons of only a few micrometers in length and thus need relay information only over short distances. Others have axons that are centimeters or meters long and must convey signals over much greater distances. Quite commonly, but not always, the longer axons are myelinated and thus conduct signals at much greater velocities. Neurons are designed to detect specific kinds of environmental stimuli. For example, a thermosensitive neuron, one that detects changes in temperature, usually is not activated by mechanical signals such as changes in tissue deformation. Classically, this quality has been called the law of specific energies. It reveals that neurons were designed to be specialists.

What Is the Nervous System?

Technically, there are several different nervous systems in the human body. These include the central nervous system (CNS), the peripheral nervous system (PNS), and the autonomic nervous system (ANS). The autonomic nervous system is subdivided in two, into the sympathetic nervous system (SNS) and the parasympathetic nervous system (PSNS). A fourth nervous system is the enteric nervous system (ENS) or gastrointestinal nervous system. Table 2.1 summarizes the various nervous systems, where they are located, and what they do.

Like individual neurons, the different nervous systems were designed to perform specific functions. The CNS consists of neurons contained within the cranium, such as those of the cerebral cortex. It also includes the spinal cord and the brain stem, which connects the cord with the cranial components. The

TABLE 2.1

Summary of the various nervous systems, where they are located, and what they do

Name of system	*Where located*	*Functions*
Autonomic nervous system (ANS)	in/near cord and cranium	involuntary action (e.g., temperature regulation; heart rate)
Central nervous system (CNS)	cranium, spinal cord	voluntary actions, reflexes (e.g., doing homework, knee jerk)
Enteric nervous system (ENS)	gastrointestinal tract	food processing
Parasympathetic nervous system (PSNS)	cranium and sacral cord	involuntary actions
Sympathetic nervous system (SNS)	thoracic and lumbar cord	involuntary actions

Note: there are many body functions that are involuntary, that are too numerous to list here.

central nervous system may be thought of as part of the central axis of the body. It communicates with the periphery of the body via the peripheral nervous system. The periphery of the body includes extremities such as toes, fingers, nose, and ears. The PNS is composed of neurons that extend from the CNS to the periphery (motor neurons) and other neurons that extend from the periphery into the CNS (sensory neurons). An example of the latter would be thermoreceptors in the fingertips. Each neuron in both classes (motor and sensory) is composed of at least a cell body, an axon, and a synapse.

The autonomic nervous system is also a specialized nervous system (see figure 2.1). It is primarily a motor nervous system. That is, the same kinds of sensory environmental stimuli that activate the CNS and PNS also activate the ANS. The cell bodies in ANS neurons are located either in the CNS (for example, somewhere in the spinal cord), or in one of the twelve cranial nerves, or in or very near the activators that these motor neurons innervate. The heart is a good example of the latter. Cell bodies of neurons of the sympathetic ANS are located either in the thoracic or lumbar segments of the spinal cord. This subdivision of the ANS is therefore also correctly referred to as the thoracolumbar ANS. However, these neurons make connections with other SNS neurons that lie, anatomically, outside of but

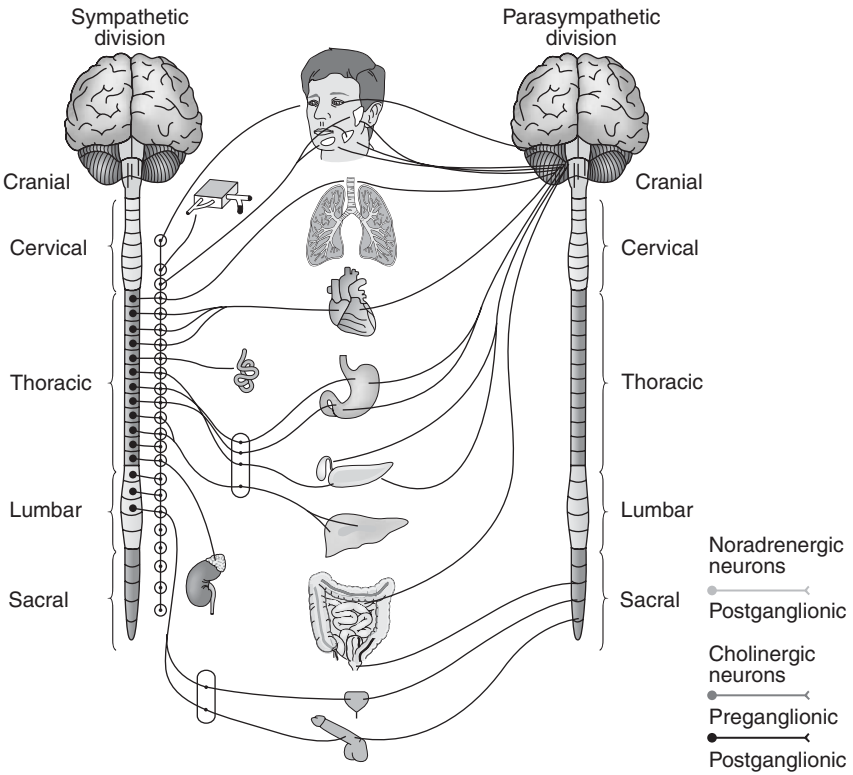

FIGURE 2.1 The mammalian autonomic nervous system, including both sympathetic (left) and parasympathetic (right) branches. Note the presence and proximity of the paravertebral sympathetic chain to the sympathetic system. Successive arrows in each system denote preganglionic and postganglionic axons. Note the anatomic origins of the preganglionic axons in the sympathetic system (thoracic and lumbar spinal columns) versus the parasympathetic system (cranium, sacral spinal column). Note also the relative lengths of pre- and postganglionic axons in the two systems.

(From http://www.arts.uwaterloo.ca/~bfleming/psych 261// image 29.gif via images. google.com)

parallel to the spinal cord. These other neurons are called postganglionic neurons. This makes the first set of neurons preganglionic neurons. The ganglia are interconnected from thoracic to lumbar segments. These ganglia, and their accompanying neurons and axons, form a bilateral chain, that is, two chains arranged parallel to the spinal cord. The chain is sometimes referred to as the paravertebral sympathetic chain. So, in the SNS there are two sets of neurons connected in series with each other. This is also true of neurons in the second branch of the ANS, the parasympathetic ANS or PSNS. Here, the cell bodies of the preganglionic neurons are located either in the cranium (that is, in one of the twelve cranial nerves), or in the sacral segment of the spinal cord. Cell bodies of the postganglionic neurons are located in or very near the effectors or activators

they innervate. Because of these anatomical arrangements, the SNS has pre- and postganglionic axons that are long. The PSNS, on the other hand, has long pre- but short postganglionic axons. (A ganglion is a group or cluster of nerve cell bodies.) In the case of the SNS postganglionic neurons, their cell bodies are found in ganglia (plural for ganglion) that lie a centimeter or so outside the spinal cord.

The ENS is also called the little brain. It is a specialized system of neurons and reflex arcs confined to the muscles and other tissues of the gut wall. It extends from the lower third of the esophagus to the anus. Even though it is anatomically separate from the other nervous systems, it communicates with and is influenced by them. There are three plexuses of nerves in the gut wall. From the inside out, the first is called the submucosal plexus. It is located just beneath the mucosa and functions to stimulate secretions and movements of the mucosa. These help in the digestion, movement, and absorption of food. Auerbach's and Meisner's plexuses are located in the midwall of the gut and serve to cause contractions of the longitudinal and circular muscle layers. Such contractions aid in the mixing and forward propulsion of meals. Ultimately, they assist in the defecation reflex and movement of the bowels.

Figure 2.1 also shows selected activators (organs and organ systems) that are innervated by the autonomic nervous system. For greater detail and a more complete list of the organs innervated by the ANS, consult any of the textbooks of medical physiology I have suggested for additional reading. Note that in the case of the ANS, the activators are not skeletal muscle cells (activators for the PNS), but rather are thoracic and abdominal visceral organs such as kidneys and gut, heart and lungs.

Neuron Communication, Electrical and Chemical Transmission

The ability of the different nervous systems to communicate with each other and with their various effectors is importantly dependent on the structural arrangements described above. How they communicate with each other and with effectors is wholly a matter of physiology, specifically neurophysiology and the action potential.

On December 28, 2006, I watched the inaugural Texas Bowl college football game in which the Scarlet Knights of Rutgers University played the Wildcats of Kansas State University (Rutgers won the contest, 37 to 10). Perhaps the most exciting play of the game was a fifty-yard touchdown pass thrown by Rutgers quarterback Mike Teel to wide receiver Tim Brown. The play was like a fine-tuned Broadway performance. Teel's pass could not have been thrown with greater accuracy. Brown's nearly sixty-five-yard sprint could not have had better timing. Without breaking pace and while obliquely peering over his right shoulder, Brown and the ball met about ten yards shy of the goal line. Brown's momentum carried them the remaining distance to score the touchdown, the first in Brown's collegiate football career.

This finely tuned performance of sensory and motor physiological functions, displayed by two college football athletes, can hardly be explained in a book of this kind. However, it was Teel's sensory perception (physiology of vision and spatial orientation of his body in place and time) of Brown's location and speed (partly motor, partly sensory functions) that determined Teel's ability to deliver the perfectly thrown ball to such a rapidly moving target. Teel's sensory physiology coupled with the strength and coordination of his throwing shoulder, arm, and wrist (physiological motor functions) were both essential to the delivery of his well-thrown ball. This, complimented by Brown's speed and downfield location (partly motor, partly sensory functions), enabled him and the ball to meet in the same place at the same point in time. Nowhere is the physiology of sensory perception and motor function better illustrated than in the precision of competitive sports and the performing arts. An agitated male lion, in hot pursuit of a hyena that has been taunting the hunting lionesses of his pride, comes in a close second.

Both sensory perception and motor function require physiological communication. Physiological communication involves many processes, including flow of ionic current, nerve conduction, and synaptic transmission. The conduction and transmission necessary to convey information take place by electrical and chemical processes. To understand their physiology, the student might be able to identify with an analogy to a subway system. The subway track may be thought of as the nerve axon, and subway cars are the action potential, that is, the electrical signal moving down the track (axon). Individual subway stations are analogous to the internodes of myelinated axons. Commuters enter the cars at station number one. As the cars move at a relatively constant velocity down the track to station number two, some riders stand in place. Others move from one car to another. At station number two, some of those who entered at station number one get off and new riders get on. The train then moves to station number three and a similar exchange of riders, from car to car and between cars and platforms, occurs. This process is repeated locally, that is, at each station, until the subway reaches its final destination.

The commuters moving between cars and platforms are equivalent to electrical charges diffusing in and out of the axons of nerve cells. Some ions stay in place once they enter the axon, while others diffuse longitudinally within the axon. Like charges repel one another and dissimilar charges are attracted to each other. Exchange of ions, or electrical current, between the interior and exterior of an axon occurs only at the internodes of myelinated axons, just like exchange of commuters occurs only at designated stations. That is, commuters are not free to enter or exit cars laterally at will. The space between the cars and the platforms—potentially dangerous to careless passengers—is analogous to the synaptic cleft, that is, the space between the axon terminus and the postsynaptic effector. The station platform is analogous to the postsynaptic protein receptor.

When a single neuron receives an incoming signal via a dendrite or its cell body, the stimulus must be converted into an electrical event that can be

conducted down the corresponding axon (subway track) and to the effector (station platform). *Conduction*, in biological tissues, means the movement of *electrical* current from one region of a cell to another. Electrical current in biological tissues is carried by charged particles called *ions* (subway commuters). The ions of greatest importance in neurons and in other excitable cells are sodium (Na^+), potassium (K^+), and calcium (Ca^{2+}).

Ionic charges are distributed, homeostatically, between the interior and exterior of the cells (like commuters are distributed between platforms and subway cars). The locations and concentrations of ionic charges, like subway commuters, transiently change. When one segment of an axon is stimulated (that is, when ionic current is flowing between outside and inside of the axon), the local region of the axon is said to be depolarized. When the subway cars are stopped at a station, the commuters flow into and out of the cars locally. In the axon, adjacent regions that are not depolarizing are said either to be quiescent or repolarizing. As subway cars exchange commuters spatially and temporally, neurons conduct action potentials by a series of temporal and spatial, but local, exchanges of ions referred to as depolarizations and repolarizations (also known as local responses).

Because ions are charged, they are indeed able to conduct electrical current in living tissues. Thus, part of the process of communication in excitable cells like neurons involves conduction of electrical currents carried by ions. The second part of the process of communication involves transmission across synaptic junctions (from subway cars to platforms). This process occurs via the release of synaptic stores of chemicals. The word *transmission*, in the case of neuron-to-effector communication in physiology, refers to a signal crossing the synapse. Although charged ions are involved in the total process, the actual crossing of the gap depends on *chemicals* called neurotransmitters. The process of neurotransmission has been under investigation by physiologists for more than one hundred years. During the last forty years, physiologists have identified dozens of neurotransmitters that play roles in cell-to-cell communication. I will focus on three or four of the most well-known of these.

In the case of the PNS and its primary activator—the striated skeletal muscle cell—the neurotransmitter of interest is acetylcholine (ACh). This chemical compound is synthesized and stored on the presynaptic side of the gap in subcellular storage organelles called vesicles. The vesicles are part of the presynaptic neuron. When the action potential that was conducted electrically down the axon reaches the presynaptic membrane, it depolarizes. One of the end results of depolarization is fusion of the vesicle's membrane with the presynaptic membrane and release of ACh into the synaptic cleft. As ACh from many vesicles is released into the cleft, its concentration builds up and the molecules diffuse across the gap and come in contact with the postsynaptic membrane. Integral components of the postsynaptic membrane are tiny docking stations called receptors. These are proteins with structural properties that match those of the neurotransmitter.

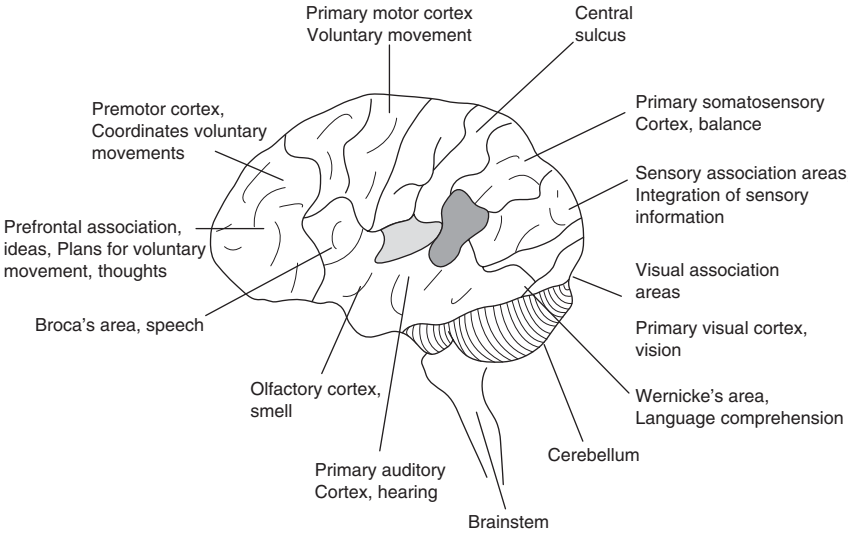

FIGURE 2.2 Lateral view of a cerebral hemisphere. Note the various regions and the associated functions that are regulated there (e.g., occipital lobe and vision; premotor cortex and voluntary movements).

(Author's illustration)

Acetylcholine binds to the receptors and causes activation or depolarization of the postsynaptic membrane. Depolarization is conducted over the entire surface of the muscle cell as well as deep into the interior of the cell. Through a cascade of subsequent ionic/chemical interactions, that is, via a signal-transduction pathway, the depolarization or excitation of the muscle cell is converted into a mechanical response called *muscle contraction*. The overall process involving a nerve action potential and muscle contraction is called excitation-contraction coupling (E-C coupling), and the end result is the shortening of a single muscle cell with the accompanying generation of force (see figures 2.1 and 2.2). When applied to an entire muscle system, E-C coupling enables the arms to move an object or to lift a load. When applied to the entire body, excitation of multiple nerves and the responses generated by their effectors keep systems such as the production and release of hormones and their many physiological regulatory effects intact. This often means the difference between health and disease.

Different Neurotransmitters and Their Locations and Functions

Chemical neurotransmission involves several steps. It begins with an action potential arriving at the presynaptic membrane and depolarizing it. The depolarization causes the opening of potential-dependent (voltage-gated) calcium channels in the presynaptic membrane, which allows ionized calcium (Ca^{2+}) to

enter the presynaptic terminal. The increase in intracellular calcium causes the membranes of presynaptic vesicles (containing neurotransmitters) to fuse with the presynaptic membrane, which releases neurotransmitter into the synaptic cleft. The neurotransmitter diffuses across the cleft and binds to receptors on the postsynaptic membrane. This activates the postsynaptic membrane, which, in turn, depolarizes the postsynaptic cell. The entire process of neurotransmission is terminated when enzymes in the cleft or on the postsynaptic membrane hydrolyze the neurotransmitter (that is, alter its chemical nature thus terminating its effectiveness), when neurotransmitter is taken up by the presynaptic membrane and repackaged into vesicles, and/or when the neurotransmitter diffuses away from the synaptic cleft and can no longer reach the postsynaptic membrane.

This entire process has physiological relevance to health and medicine on a continuous basis. Consider the simple act of eating a meal. Each time one chooses to swallow, reflexes are activated. These briefly terminate respiration thereby minimizing the chances of a bolus of food passing accidentally into the trachea rather than the esophagus. Collectively these nerve/effector actions are called the swallowing reflex. By this reflex, as the bolus of food is forced to the back of the mouth in preparation for swallowing, the epiglottis is lowered over the glottis thus closing entry to the trachea. Simultaneously, inspiratory neurons are inhibited and respiration comes to a temporary standstill in expiration. The upper esophageal sphincter is relaxed thus permitting easy entry of the bolus of food into the esophagus. The entire sequence of reflex events is a highly coordinated, marvelous display involving sensory and motor nerves, neurotransmitters, and their effectors. Under normal conditions, the end result is swallowing of a bolus of food and absence of choking. The process can be frustrated by the careless eating of food and concurrent participation in uncontrolled conversation, exercise activities, and so forth. Careless interference with the swallowing reflex can result in accidental choking and even death.

In the nervous systems defined above, the primary neurotransmitters are acetylcholine (ACh) and norepinephrine (NE). ACh was the first neurotransmitter identified, isolated, and purified for investigation. Norepinephrine was discovered not long after, and both are among the most frequently studied neurotransmitters of the past and present. Many others are listed in table 2.2.

Most neurotransmitters are released from the presynaptic nerve terminal and act on the postsynaptic nerve terminal. However, the case of the autonomic nervous system—where the motor (efferent) action potential travels over two consecutive neurons in series with each other—is special. In both the PSNS and the SNS, there are pre- and postganglionic axons, therefore there are also pre- and postganglionic chemical neurotransmitters. For students of physiology, neurosciences, and medicine, it is essential to associate the proper neurotransmitter with the proper synapse. Their sites of release, reuptake, and receptor recognition

TABLE 2.2

Selected classes of chemical neurotransmitters and examples of each

Class	*Examples*
Amino acids	aspartate, gamma amino butyric acid, glutamate, glycine
Monoamines	acetylcholine, histamine, serotonin
Catecholamines	dopamine, epinephrine, norepinephrine neuroactive peptides
Peptides	angiotensin II, arginine vasopressin, B-endorphin, vasoactive intestinal peptide
Purine derivatives	adenosine, adenosine triphosphate (ATP)

are not intuitive. The only thing that might be intuitive is the coupling of the neurotransmitter with its receptors, but even this can vary. Neither is the location of postsynaptic receptors intuitive. For the neurotransmitters ACh and NE, the student must memorize their pathways within the mammalian autonomic nervous system.

In the PSNS, the neurotransmitter at both the pre- and postganglionic axon is acetylcholine. In the SNS, the neurotransmitter at the preganglionic axon is acetylcholine. At the postganglionic axon, the neurotransmitter is norepinephrine. There are specialized cases in the ANS, CNS, and PNS as well where neurotransmitters and related chemicals are colocalized and coreleased in response to action potentials.

The mammalian nervous systems use dozens of neurotransmitters that act on more than 100 receptors that have been identified to date. The receptors, such as those found in the postsynaptic membranes, operate through second messenger systems that in turn regulate dozens of enzymes and ion channels (pores in cell membranes). The second messenger system is so named because it is activated by a primary or first messenger, that is, the original neurotransmitter. For example, catecholamine neurotransmitters such as norepinephrine (the primary or first messenger) bind to their postsynaptic receptor and activate an enzyme called adenylate cyclase (or adenylyl cyclase). This activated enzyme converts adenosine triphosphate (ATP) to a compound called cyclic adenosine monophosphate, or cyclic AMP (or cAMP). Cyclic AMP becomes the second messenger and activates a cascade of downstream chemical reactions that lead to useful physiological functions. For example, when either norepinephrine or epinephrine binds to its corresponding adrenergic receptor in the liver, cAMP, acting as a second messenger, activates protein kinase A (PKA).

This enzyme in turn activates a glycogen phosphorylase. The entire sequence and other components of this signal-transduction pathway stimulates glycogen breakdown and inhibits glycogen synthesis.

Understanding the above also suggests how errors in these systems can affect the function of a single cell, a tissue, or an organ system. Dysfunction of the behavior of cells, tissues, and organs means disruption of the health of the person in whom these systems have gone awry. Consider Parkinson's disease, for example. This disease is characterized by muscle rigidity, involuntary tremor, and akinesia (difficulty initiating movement). The akinesia is often of greater concern to the patient than either the muscle rigidity or involuntary tremor. This is because, in advanced stages of the disease, performing even the simplest of motor movements requires tremendous concentration, mental effort that deepens to anguish as the disease worsens. Parkinsonism is associated with the neurotransmitter dopamine, and dopamine-containing neurons are found throughout the CNS. They are especially abundant in deep nuclei of the cerebrum (the basal ganglia) and in the structures of the midbrain such as the substantia nigra. Collectively, these structures among others regulate motor functions of the body. The substantia nigra projects axons to the basal ganglia, particularly the caudate nucleus and putamen that influence voluntary movement. Degeneration of the dopamine-containing cells of the substantia nigra produces the progressively worsening and finally fatal motor disorders of Parkinsonism.

Treatment for Parkinson's disease is therefore based on an understanding of dopamine as a neurotransmitter and the structures it affects. Administration of a drug called L-dopa to patients suffering from Parkinsonism usually attenuates the severity of the symptoms, especially rigidity and akinesia. L-dopa is converted to dopamine in the brain. Another treatment for the disease is the drug L-deprenyl. This drug inhibits the enzyme monoamine oxidase (MAO), which destroys dopamine after it has been released into the synaptic cleft. By inhibiting MAO, L-deprenyl enables dopamine to remain in the tissues for a longer period of time. Combination of L-dopa and L-deprenyl generally provides more effective treatment for Parkinsonism than either drug alone. Transplantation of dopamine-secreting cells (obtained from the tissues of aborted fetuses) into the brains of patients has met with some success. However, this treatment is short-lived because the fetal cells survive for only a few months. One of the great hopes the new field of regenerative medicine offers is the prospect of using stem cells to treat diseases such as Parkinsonism. Researchers hypothesize that by administering undifferentiated stem cells in combination with the appropriate nerve growth factors, new neurons will grow in the substantia nigra and basal ganglia, thus replacing dead tissue and restoring motor function.

Huntington's disease also involves neurotransmitters and the central/peripheral nervous systems. Huntington's disease is a heritable disorder that appears at age thirty to forty years. It begins with flicking movements in individual

muscles. These progress to severe distortional movements of the entire body. Severe dementia is associated with the impaired motor functions. Huntington's disease is thought to be caused by loss of the cell bodies of GABA-secreting neurons in the caudate nucleus and putamen and of ACh-secreting neurons in many parts of the brain. GABA-secreting neurons are thought to inhibit excitatory neurons. The loss of inhibition allows spontaneous outbursts of excitatory activity that lead to involuntary muscle contraction and body distortions. The dementia associated with Huntington's disease is said to result more from loss of ACh-secreting neurons than from those that release GABA. This is especially true of areas of the cerebral cortex associated with thinking and reasoning.

Discomfort, Pain, and the Nervous System

Hyperexplexia, or startle disease, is a rare genetic disorder of humans that often begins in infancy. It is characterized by continuous marked rigidity or hypertonia, except during sleep, and marked, generalized muscle contractions to sudden noises (hyperreflexia). As the child grows and develops, the rigidity disappears, but the marked and excessive startle response to unexpected auditory, tactile, or visual stimuli continues. The gene defect has been identified in the alpha one subunit of the inhibitory glycine receptor. Glycine is a major inhibitory transmitter in the spinal cord where it mediates crossed inhibition and inhibition of sensory inputs.

On the night of March 23, 1994, a natural gas line exploded in Edison, New Jersey, by the Durham Woods Apartments. The enormous explosion rocked an area of several square miles and the atom-bomb-like fireball could be seen as far away as Manhattan. Fourteen buildings in the apartment complex were leveled or badly burned. Fortunately only ten people were injured and no deaths resulted directly from the explosion. However, within an hour or two, thirty-seven-year-old Sandra Snyder, a nearby resident, died of a heart attack because in the chaos she could not get emergency care. Her husband later described Sandra's response to the blast. She sprang out of bed, donned a bathrobe, and began pacing the house nervously before subsequently collapsing and going into cardiac arrest. Her husband performed CPR for nearly thirty minutes, but she was later pronounced dead at nearby JFK Hospital. Mrs. Snyder's response to sudden, unexpected noise and confusion marks the most extreme outcome of this neurologic condition (although it is not known if she was an undiagnosed sufferer of either hyperexplexia or cardiac disease).

Both acute and chronic pain is associated with injury to and/or degeneration of the nervous system. The spinal cord is part of the central nervous system. There are millions of neurons/axons that project into and extend away from the spinal cord. An estimated 80 percent of Americans will have to cope with back pain at some time in their lives. Other than coughs and respiratory infections,

back pain is the leading cause of visits to the physician's office. Those who suffer from back pain cost the United States approximately $90 billion annually in medical expenses, disability payments, and lost productivity at work. With the increasing tendency toward lifestyles of physical inactivity and obesity, these numbers are likely to climb in the future.

Barring congenital and genetic anomalies over which we have little control, the best way to avoid injury to the spinal cord and other nerve tissue is to use good judgment and common sense in everyday life. My seventeen-year-old son told me about the athletes he follows in the arena of extreme sports. At one event young men tried to jump, on mountain bikes and skateboards, the Great Wall of China. One mountain bike competitor attempted his jumps with a lightweight helmet and no other body protection. His hand slipped off the handle bars as he hit the receiving ramp, causing him to careen over the edge and away from the safety net. He died as a result of the accident. An American skateboarder wearing better protective gear made the jump but fell on the receiving half-pipe. Fortunately he escaped serious injury. In these extreme cases of potential injury to the nervous system, the athletes involved placed the value of transient excitement, publicity, and glory above the value of life. We, more than anyone else, should work tenaciously to protect our own lives and their value to us and to others.

Having said that, I must acknowledge my own carelessness in protecting my health during the moment's excitement. I have accidentally injured ankles (playing basketball), knees (hiking the Appalachian Trail), shoulders (mountain biking), and wrists (retro-racing)—all in my adulthood. Even after such simple accidents, recovery can be incomplete. For example, the range of motion of an injured joint might not ever be fully restored. The tissue remodeling that takes place in a broken or sprained ankle could prevent one from ever ice skating again. Heaven forbid a horse-jumping accident of the kind that eventually led to the death of actor Christopher Reeve. I have had orthoscopic surgery on both knees, first to remove bone fragments and second to trim a damaged meniscus. In both cases, the orthopedist said I would be back to 100 percent of function after several weeks. The two surgeries took place about ten years apart, and the earlier one took longer to heal than the latter. In neither case did I feel fully recovered until nine to twelve months postsurgery. I still experienced considerable discomfort and pain, as well as noticeable inflammation in both knees after eight to ten weeks. Moreover, neither orthopedist explained some of the side effects I might expect as a result of the procedures. During the period of postsurgical inactivity, I lost muscle and other soft tissue mass from my legs as well as in my feet. I developed bone spurs in my left heel when I tried running too soon after the surgery and loss of tissue mass. For two to three years after surgery on my right knee, I experienced numbness and absence of sensation in the area of the surgery. This was most likely the effect of severed sensory nerves caused by the surgical procedure.

Had the surgeons been more forthright with their experience and information, I would have resumed my exercise program more cautiously. Instead of running several miles a day, I would have walked a lesser distance and gradually built up to my former running regimen. Still, I can take some of the blame for the problems because as a physiologist I should have understood these things and used better judgment after the surgeries. In the case of the second surgery, however, I could not have known that the surgeon and procedure would disrupt function of sensory nerves in my knee. Even the best common sense could not have compensated for this.

Aside from backaches and pain associated with surgical procedures, a more common cause of nerve-related discomfort is headache. Many of us experience on a regular basis the intermittent headaches associated with employment and the routines of daily life. Fortunately most of us do not experience debilitating and sustained migraine headaches. Travelers to high altitudes may experience symptoms of acute mountain sickness defined by headaches that are described as severe, bitemporal, and throbbing; in the United States alone, more than 30 million people travel to elevations over 2,000 meters every year, and at least 25 percent of them will experience symptoms of mountain sickness severe enough to limit activity. Other symptoms besides headache include nausea, insomnia, anorexia, and general fatigue. Although very common and painful, a primary concern of high-altitude headache is that it may be a harbinger of more rare but potentially fatal forms of high-altitude sickness, including high-altitude pulmonary and cerebral edema.

High-altitude headache was first described in writings from ancient China dating to 32 B.C. Nausea and headache are defining symptoms of acute mountain sickness. Although the definitive mechanisms are not known, they are thought to be related to the time-delayed effects of hypoxia (lack of oxygen at altitude). Potential mechanisms involved in high-altitude headache include vasodilation and an increase in cerebral blood flow. This would increase the volume of cranial tissue thus making the brain fit more tightly inside the cranium. Cerebral compression could cause increased sensitivity and damage to meningeal sensory neurons.

Ibuprofen and other nonsteroidal anti-inflammatory drugs have been established as the standard treatment for high-altitude headache and are the gold standard against which other pain-relieving analgesics are compared. However, therapeutic use of these compounds is often associated with gastrointestinal problems including ulceration and bleeding. Stuart Harris and coworkers found that acetaminophen (Tylenol) was just as effective as ibuprofen in reducing the incidence of high-altitude headache in mountain climbers on Everest and suggested that it might be a good replacement for nonsteroidal anti-inflammatory drugs since it is equally efficacious and lacks the gastrointestinal problems (see Suggested Reading).

If acetaminophen behaves in cerebral tissue as it does in cardiac tissue, there could be a simple explanation for its effectiveness. In heart tissue, acetaminophen effectively attenuates the production of nasty oxidants such as hydroxyl radical, peroxynitrite, and superoxide anion, all of which are known to damage living cells including sensory neurons, vascular endothelium, and vascular smooth muscle. These experiments were conducted in hearts where blood flow was temporarily stopped (made ischemic) then restored (reperfused) (see Suggested Reading). We call this kind of experiment myocardial ischemia/reperfusion injury, and acetaminophen appears to protect both the function and structure of the heart under such conditions. Acetaminophen hasn't been studied under similar conditions in the brain. However, it is conceivable that at high altitude and in the presence of hypoxia, the production of oxidants and other damaging molecules in the CNS could outpace their rates of destruction and removal. This could lead to net accumulation and untoward effects including headache. Such a mechanism might apply similarly to the headaches and discomfort many of us experience by the end of a busy workday even at sea level. Of course final judgment must be withheld until the experimental data are collected and published.

Mental Activity, Cerebral Blood Flow, and Health of the Nervous System

Other than therapeutics for damaged/stressed nervous tissue, there are several steps we can take proactively to understand and care for our central and other nervous systems. The image in figure 2.2 is a left-sided lateral view of the cranial components of the CNS (cerebral hemisphere, cerebellum, and brain stem). Imagine this cut in half to yield right and left sections of the cerebrum. Also imagine that we are able to view the interior portions of either half. In the general regions labeled Broca's area, primary auditory cortex, and Wernicke's area, we would find structures collectively called the *limbic system*. The limbic system includes the neuronal circuitry that controls our emotions and motivational drives. A major part of the limbic system is the hypothalamus. In addition to being an important relay station for higher brain regions, the hypothalamus acts as a governor of the pituitary gland, and therefore of all the neuronal and endocrine functions of the body that the pituitary and its hormones regulate. Some simple examples of physiological variables controlled by the hypothalamus are body temperature, osmolality of body water, thirst drive, hunger drive and satiety, and control of body weight and body composition. The hypothalamus, which is less than one percent of the brain's mass, is one of the most important control pathways of the limbic system. Its activities are importantly affected by blood flow. Structures shown in figure 2.2 are to compliment an understanding of experiments described below.

Figure 2.3 and the text accompanying it illustrate regional cerebral blood flow and how it is increased by various physical and mental activities (shaded circles

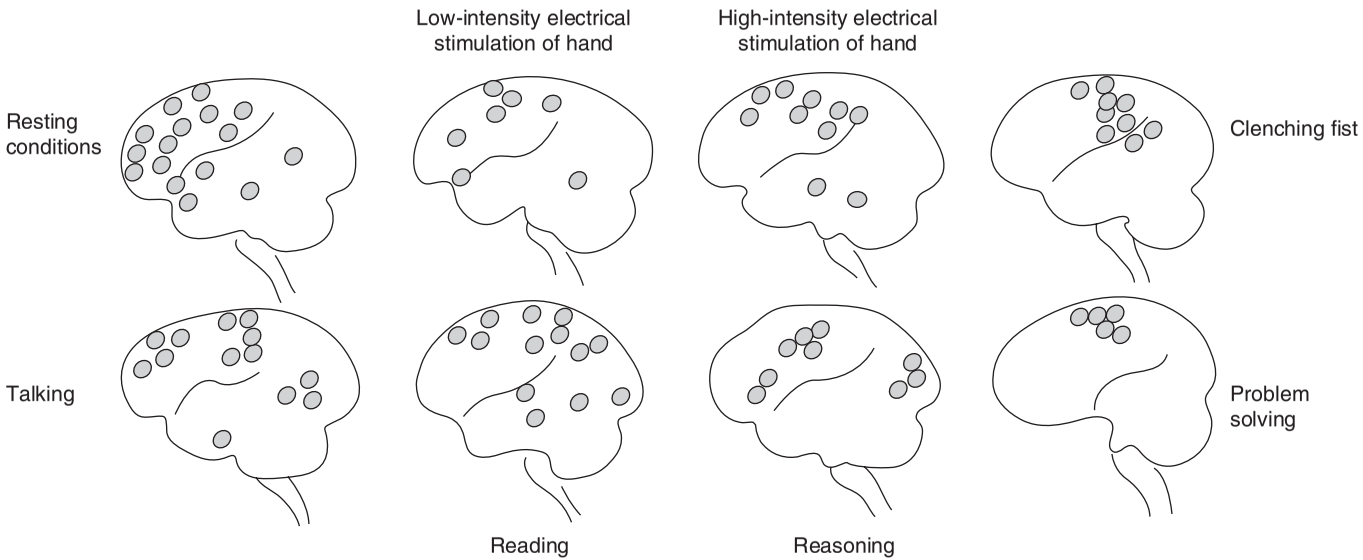

FIGURE 2.3 Regional distribution of cerebral blood flow (circles with shading) during various physical (talking, clenching fist), emotional (low vs. high intensity electrical stimulation of the hand), and intellectual activities (reading, problem solving). Data were collected in patients using the washout of radioactive, but nonharmful, xenon with detectors placed over the sides of the patient's head (see Ingvar DH. *Brain Research* 107: 181–197, 1976).

(Author's illustration)

and adjacent areas). There are several ways to measure the total volume of blood flow to the cerebral hemispheres. Most are experimental, but some are clinical. If the measurement is being made in an experimental animal, the animal might have to be euthanized at the end of the experiment so the brain can be removed and analyzed. Experimental techniques include administration of radiolabeled tracers, dye and thermal dilution methods, and ultrasonic and electromagnetic probes placed on major arteries that deliver blood flow to the brain. Clinical techniques include thermodilution, dye dilution, and detecting radiolabeled, but nonharmful, tracers. Increased blood flow means increased delivery of oxygen, blood sugar (glucose), and other nutrients that are good for nerve tissue. It also means washout and removal of harmful byproducts of metabolism (that is, those that do not get metabolized) that might prove damaging to the brain if allowed to accumulate. Engaging in activities that increase cerebral blood flow is good for us, and we should make a regular practice of doing such things. Among the most important cognitive activities is learning, for example, reading and trying to pronounce unfamiliar words, seeing and learning new words, learning a foreign language, playing Scrabble, even watching a television program like *Jeopardy*.

Even though the two cerebral hemispheres appear to be symmetrical, each half is morphologically, chemically, and functionally specialized. The left cerebral hemisphere deals primarily with the motor and somatosensory (sensory systems of the body) functions of the right side of the body. In addition, in the vast majority of the population the left cerebral hemisphere processes language. Language is a complex activity that includes watching and seeing, hearing and listening, speaking and understanding, and reading and writing. Different regions of the brain seem to be specialized to cope with these various components of language.

The most important take-home message is that engaging one's mind in mental activities is good for the health of the brain. Experiments have been/can be done to investigate the influences of studying crossword puzzles, giving service to others, avoiding rage and anger, and speaking kind and charitable thoughts to others. For the interested undergraduate student, I would recommend experience as a research assistant in the laboratory of a behavioral, cognitive, or biomedical neuroscientist. It is often said that the best way of learning is by doing. Being part of a team that designs, evaluates, and publishes an experiment in neuroscience is probably one of the best ways to drive home the above messages. Such opportunities abound for the average student at well-endowed public and private universities. They are less likely at community colleges and small liberal arts colleges. Even if yours is a shy and timid personality, improving your social confidence and conversational skills (by practicing) will contribute to a healthier central nervous system even as you age.

3

The Endocrine System and Physiological Communication

Advances in Endocrinology

Both the nervous and endocrine systems were designed for communications. The nervous systems rely on the principles of conduction and transmission using electrical and chemical signals associated with individual neurons. Most commonly, these neurons are arranged in series whether they be on the sensory or motor sides of the central nervous system. They are physical structures (biological current-conducting cables) in contact with other physical structures (other neurons and activators such as skeletal muscle cells). Parallel arrangements of neurons and their effectors, however, are not uncommon.

The endocrine system is arranged differently. Historically endocrinology refers to selected organs called glands, for example, the thyroid gland or the adrenal gland, that produce hormones that are released into the circulation and carried to distant targets (cells, tissues, organs) to produce physiologic actions. The details of any single component of the endocrine system have been the topics of dozens of textbooks. In today's world of modern molecular biology and with the continuing revolution in the life sciences, endocrinology is much more than its classic definition. Today's student must also understand the concepts of paracrines and autocrines in addition to the classic concepts of endocrines.

An endocrine hormone is one that gets released from its mother cell into the blood; blood becomes the vehicle of transport for that hormone. Those hormones acting as autocrines and paracrines get released into the interstitial spaces and must rely on mechanisms other than the circulation for transport to their target tissues (for example, diffusion).

An exocrine gland is one that releases its products into tubules and ducts (for example, the acinar cells of the salivary and pancreatic glands). Consider the pancreas gland. It has both endocrine and exocrine functions. It synthesizes and releases the endocrine hormones insulin and glucagon. These are transported by

the blood to distant tissues like fat cells, muscles, and the liver. Among pancreatic exocrine functions are the synthesis, storage, and secretion of digestive enzymes and bicarbonate ions. These are released into pancreatic ducts (tubules), then pass through the common bile duct, and are finally delivered into the duodenum of the small intestine. In the duodenum, the enzymes participate in the breakdown and uptake of fats and proteins and the bicarbonate ions help buffer acids emptied from the stomach into the small intestine.

We have learned in recent decades that a product released from one cell can have an effect on the behavior of an adjacent cell (a paracrine function, see figure 3.1). Moreover, that same product might evoke a physiological or biochemical response on the same cell in which it was synthesized and released (an autocrine function). These concepts were not known in the late nineteenth and early twentieth centuries when the fields of endocrinology and neuroendocrinology were beginning to develop. In fact, though I began to study endocrinology in the late 1960s and early 1970s, I did not become familiar with these terms until the mid 1980s, when I was already into my academic career as a professor of physiology. Paracrines and autocrines often reach their target cells by the principles of simple diffusion. This means the differences in their concentrations in two locations, the distance between the two locations, and the random movements of the molecules (thermodynamics), among other considerations, determine their movements between the two points.

The average student probably thinks of endocrinology as the functions of the pituitary, adrenal, thyroid, and pancreas glands. However in recent decades,

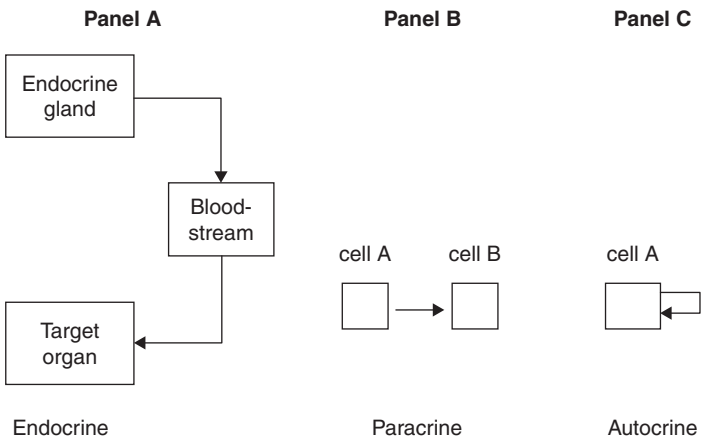

FIGURE 3.1 Contrasting mechanisms by which a hormone can be released, transported, and can act on a target cell. In panel A, the target cell is distant to the endocrine gland (endocrine function). In panel B, both cells A and B reside in the same gland (paracrine function). In panel C, the cell that releases the hormone might itself be the target tissue (autocrine function).

the field has developed and shown that the endocrine system is much broader, more complex, and more dynamic than this. We now know of organs, tissues, and cells that have endocrine-like properties but were formerly not thought of as dedicated endocrine organs. Prior to the second half of the twentieth century, physiologists used the word *neurohumoral* to refer to chemicals that are released either as neurotransmitters or from unidentified organs and tissues. That word fell out of vogue in the second half of the twentieth century but might be appropriate to reintroduce as we think of the more recently defined autocrine, endocrine, and paracrine functions of tissues and organs.

As an example of this new knowledge, let's look at the heart. When I was a graduate student, scientists did not know that the mammalian heart had endocrine properties. However, atrial natriuretic peptide (ANP) is a hormone produced in atrial tissue when the volumes of blood or pressures inside the atria are elevated. Stretch of the atria is a signal for them to release ANP. Under such conditions, ANP is released into the general systemic circulation and causes the kidneys to excrete excess sodium. The loss of sodium coupled with the water it takes with it is referred to as natriuresis. When the kidneys excrete excess sodium and water, the circulating volume of blood can decrease. As blood volume declines, so also does arterial blood pressure and pressures in other parts of the circulatory system, for example, in the atrial chambers. As atrial volume and pressure (stretch) decrease, the signal to release ANP is removed and the heart and circulatory systems return to their homeostatic steady states. This is a feedback system that works as described in chapter 1. Again, stretch of the atrial chambers is a trigger that signals the release of ANP. ANP acts on multiple effectors including cells of the kidneys. These cells are changed in a way that inhibits reabsorption of water and sodium, thus promoting their excretion. As they are excreted, blood volume and blood pressure return to physiological levels. Their return serves as a negative feedback signal terminating further release of ANP, and homeostasis is reestablished.

The kidney is another organ that traditionally has not been thought of as having endocrine functions. However, when there is inadequate oxygen in the circulation, the kidneys sense the hypoxia and release a product called erythropoietin. This agent acts on bone marrow, causing it to release more red blood cells into the systemic circulation. As more red cells are added, the capacity of the blood to take up oxygen from the lungs and deliver it to the hypoxic tissues is increased. As this cascade of events proceeds, the initial signal (hypoxia) that evoked release of erythropoietin from the kidneys is removed and homeostasis is once again restored.

The kidneys also produce an endocrine-like product called renin. Renin acts on a circulating macromolecule, angiotensinogen, to cleave it to a smaller molecule called angiotensin I (AI). In the lungs, angiotensin I is reduced to angiotensin II (AII) by an enzyme called angiotensin-converting enzyme (ACE).

Angiotensin II is a potent vasoconstrictor. This signal-transduction pathway is importantly involved in the regulation of systemic arterial blood pressure and therefore all the physiological phenomena that are affected by it. Whenever blood pressure in the renal artery decreases, a condition known as renal arterial hypotension, renin is released, angiotensin II is produced, and systemic arterial blood pressure gets elevated. I will say more about this process in the concluding chapter.

Hypertension is a disease that affects millions of Americans, as well as millions of residents of other industrialized nations. The pharmaceutical industry has profited considerably in the past few decades by developing ACE inhibitors to help treat or otherwise control hypertension. The physiological rationale suggests that by inhibiting the conversion of AI to AII, the vasoconstrictor contributions of AII to the preexisting hypertension are reduced or eliminated. Therapeutically, this approach has been more or less effective in treating hypertension.

Table 3.1 presents a selected list of endocrine glands, the hormones they produce, and selected functions of those hormones. Hormones are classified according to their molecular structures, their synthetic pathways, and the tissues from which they are derived and upon which they act. For example, the thyroxines or T_4 and T_3 are clearly produced in the thyroid gland. Similarly, follicle

TABLE 3.1

Selected endocrine glands, the hormones they produce and release, and some of their physiological functions

Gland	*Hormones*	*Functions*
Hypothalamus	GHRH	causes release of pituitary growth hormones
Posterior pituitary	oxytocin	stimulates milk ejection and uterine contraction
Thyroid	T_4 and T_3	increase rates of chemical reactions and metabolism
Testes	androgens	promote growth and development of male reproductive tract
Ovaries	estrogens	promote growth and development of female reproductive tract
Kidneys	erythropoietin	increases production and activities of red blood cells
Heart	ANP	reduces blood pressure by promoting renal excretion of sodium ions

stimulating hormone, or FSH, acts on the ovarian follicles but is produced in the anterior pituitary gland. Many steroidal hormones have the same basic molecular structure as cholesterol and vitamin D. Examples of these hormones are the mineralocorticoids produced by the adrenal cortex such as aldosterone. Other hormone-like compounds are derived from proteins, peptides, and amino acids such as tyrosine and glutamate. Still others have the amino group NH_2 in common. Examples in this class are the biogenic amines, including dopamine, epinephrine, histamine, and adenosine.

These classifications are important because the structures of hormones determine their functions, a phenomenon pharmacologists call the structure-activity relationship. Oxytocin, for example, is important to the process of milk production and release in lactating women. There are oxytocin receptors in the alveoli and ductal systems of the female breast that recognize oxytocin but do not recognize ANP. Hence administration of oxytocin to a lactating mammal will stimulate greater production and release of milk whereas treatment with ANP will not. This simple example illustrates the importance of molecular structure to the function of a particular tissue.

Table 3.2 presents a selected list of the chemical classifications of hormones and examples within each classification.

Hormones, Proteins, and Peptides

Most of the hormones in the body are either peptides, proteins, or protein derivatives. The term *peptide* refers to the kind of chemical bond between adjacent amino acids; therefore a polypeptide is an amino acid chain with many peptide bonds. A decapeptide has ten peptide bonds and di- and tri-peptides have two or three bonds, respectively. Hormones made from peptides and proteins have several things in common. Both are composed of chains of amino acids. Both have primary, secondary, and tertiary molecular structures. However, the size of the molecule is an important characteristic of hormones. Peptides are smaller than proteins. One of the smallest peptides is thyrotropin-releasing factor

TABLE 3.2

Chemical classification of selected hormones

Peptide hormones	*Steroid hormones*	*Amino acid hormones*
antidiuretic hormone (ADH)	aldosterone	norepinephrine, epinephrine, tyrosine
secretin	estradiol (E_2)	serotonin (5-hydroxytryptamine)
insulin	androgens	thryoxine (T_4)

(TRF, or thyrotropin-releasing hormone, TRH), which consists of three amino acids. Growth hormone (GH) and prolactin (PRO), on the other hand, have as many as 200 amino acids each. Polypeptides with fewer than 100 amino acids are called peptides and those with greater than 100 amino acids are considered proteins. Another way of classifying hormones is by their molecular weights. Proteins have molecular weights greater than 5,000 daltons and peptides have molecular weights less than 5,000 daltons.

Functionally proteins and peptide hormones are synthesized inside cells on structures called the endoplasmic reticula. They are originally constructed as macromolecules that are biologically inactive. That means they cannot be used by the body to transmit signals between cells or tissues. Being bound within vesicles or to the endoplasmic reticulum prevents them from doing unwanted damage to the interior machinery of the cell such as the nucleus and mitochondrion. These macromolecules are called preprohormones when they are initially synthesized, but they ultimately get broken down to smaller products called prohormones, which are packaged inside secretory vesicles. Inside the vesicles, enzymes cut down the prohormones to biologically active hormones. Like neurotransmission, when the appropriate signal activates the endocrine cell, the vesicle fuses with the cell membrane, both membranes are momentarily and locally disrupted, and the active molecule gets released to the outside of the cell. The neurohumor (hormone) is free then to act on the original secretory cell (autocrine), to diffuse through the interstitial spaces to act on a neighboring cell (paracrine), or to be carried away in the bloodstream to act on some distant organ or cell (endocrine).

In the case of the endocrine function, many hormones circulate freely in the plasma until they arrive at their target organ. Others are bound to proteins called carriers. The bound hormone and its carrier are called a circulating or binding complex. Good examples of this kind of hormone transport system are the thyroid hormones T_4 and T_3, steroid hormones such as estrogen and androgen, and insulin-like growth factors.

At least two physiologically important functions are served by the binding of hormones to protein carriers. First, it creates a pool of readily available hormone in the vascular compartment and minimizes the hour-to-hour or day-to-day fluctuations in the hormone concentrations and actions that would occur otherwise. Second, it extends the half-life of the hormone. When in the unbound state, these hormones might have half-lives of only a few minutes, whereas they might have half-lives of several days when bound to protein carriers. The half-life of a hormone (abbreviated $t\ ^1/_2$) is the time it takes an initial concentration of the active hormone to be reduced to one half the original quantity. As an example, more than 99 percent of T_4 is bound to carrier protein in the circulation. The bound form has a $t\ ^1/_2$ of about seven or eight days. The $t\ ^1/_2$ of free T_4 is only a few minutes. If the physiological half-life of the unbound form of

T_4 was several days, general metabolic processes would be so accelerated that cells, tissues, and organs would burn out well before their intended life expectancy. Having a half-life of only a few minutes ensures a graded physiological response from the target tissue without excess damage to it.

The Hypothalamic-Pituitary Axis

There are several physiologically important endocrine glands. Among the most thoroughly studied are the pituitary gland, the gonads (testes, ovaries), the pancreas, the adrenal glands, and the thyroid gland. The latter four operate under the regulatory control of the pituitary gland. The pituitary gland is found at the base of the brain and consists of several types of endocrine tissues: the anterior pituitary (adenohypophysis), the posterior pituitary (neurohypophysis), and the median eminence, also known as the pituitary stalk. In humans, the pituitary gland is attached by the stalk to the floor of the third ventricle of the brain. The walls of the third ventricle are composed of specialized CNS nuclei including those that comprise the hypothalamus, a structure key to the physiology of the pituitary and other glands.

There is continuous communication between the hypothalamus and the pituitary. The neurohypophysis consists of neurons whose cell bodies are located above it in the hypothalamus. This means the axons of hypothalamic neurons descend through the pituitary stalk and into the neurohypophysis. A couple of the more important nuclei are the supraoptic and paraventricular nuclei. Respectively, these nuclei synthesize, among other products, ADH (antidiuretic hormone; also known as AVP, arginine vasopressin, or simply vasopressin) and oxytocin. These chemicals move from the neuronal cell bodies down the axons of the stalk and are stored in the presynaptic terminals of the posterior pituitary. ADH is involved in the physiological regulation of body water and osmolality. Oxytocin helps initiate and sustain lactation in postpartum women.

The hypothalamus also produces other chemicals that affect function of the adenohypophysis. These compounds are called releasing factors, releasing hormones, and inhibiting factors or inhibiting hormones. Such compounds are synthesized in the hypothalamus but are released under the appropriate stimuli into a special circulatory system in the median eminence of the pituitary stalk and are carried by these blood vessels to the anterior pituitary. There they activate or inhibit the release of several hormones that are synthesized in the anterior pituitary. One of the factors is gonadotropin-releasing factor (GnRF, also called gonadotropin-releasing hormone, GnRH). Its name implies, correctly, that it has something to do with the physiology of gonadal products such as the androgens (predominantly male hormones produced in the testis) and estrogens (predominantly female hormones produced in the ovaries). GnRH causes

select groups of cells in the adenohypophysis to release follicle-stimulating hormone (FSH) and leutinizing hormone (LH). These are carried by the circulatory system to the gonads where they augment the growth and reproductive functions of those glands. Other hormones of the anterior pituitary include thyroid-stimulating hormone (TSH), adrenocorticotropic hormone (ACTH), and growth hormone (GH).

The hypothalamic-pituitary axis refers to the interconnections and interactions between these tissues. While all hormones and endocrine glands are designed to promote the health of the body, I will highlight only a couple that affect growth. These are growth hormone (GH) and the insulin-like growth factors (IGF-I, IGF-II). In addition to its general effect in promoting growth, GH has several specific effects that include the increased rate of synthesis of proteins, the reduced use of glucose, and the increased breakdown and consumption of fat deposits. This effectively means that GH conserves body carbohydrates, mobilizes fat (thus minimizing its storage), and enhances uptake and use of proteins by the body. GH promotes use of protein by increased uptake of amino acids through cell membranes, increased translation of RNA (ribonucleic acid) to promote the synthesis of proteins by the ribosomes, increased transcription of DNA (deoxyribonucleic acid) to form new RNA, and decreased breakdown of structural proteins and amino acids (for example, those that comprise muscle tissue).

While many of the effects of GH are acute (they take minutes to hours to occur), they are distinct from those that are important in the long-term regulation of growth. We are not yet aware of all the effects of growth hormone nor of all the ways its physiological actions can be influenced. However a general takehome message for younger students is this. Most of us do not attain our full adult height and function until our late teens and early twenties. This means the actions of GH on height and size are still taking place up to that point. One should avoid doing anything prior to these ages—and at any age—that could impair the physiological actions of growth hormone. For example, use of toxic substances, including but not limited to illicit drugs, tobacco, and alcohol, should be avoided regardless of how socially popular they are.

We have also heard stories of the effects of weight-bearing exercise (strength training and conditioning) on height and size of youth. Use of steroids often finds its way into such stories. Serious weight training at too early an age should be avoided, or one should at least consult with the family physician before beginning a routine. I am always amused after giving endocrine-related lectures at the number of young men who approach me in private to seek help for a friend who is on steroids and working out. They are concerned for the friend's welfare and want to know what I think about athletic use of steroids. I tell them to not use these drugs and to encourage their friends to stop using as well. High school and college locker rooms are generally not great sources of

information about steroid hormones, body size and conditioning, and the appropriate physiology.

About half a century ago, it was reported that GH itself does not have growth-promoting action on epiphyseal cartilage (the site of longitudinal growth of bones). Using in vitro experiments, investigators found that plasma from physiologically normal animals stimulated the growth of cartilage. Plasma taken from animals whose pituitary glands had been removed (hypophysectomized) failed to promote growth of cartilage. The addition of GH to plasma deficient in GH did not restore the plasma's ability to promote growth of cartilage. However, if the animals that were deficient in GH were administered GH several weeks later when the plasma of these treated animals was withdrawn and placed in a petri dish with cartilage from a second, younger animal, the treated animal's plasma promoted the growth of the younger animal's cartilage. This experiment led to the hypothesis that GH stimulates the secretion of another circulating factor that mediates the action of GH. This intermediate substance was called somatomedin because it mediates the somatic (body) effects of GH. The responsible proteins were isolated, purified, and subsequently found to be identical to peptides with structures similar to proinsulin. Hence they were given the name of insulin-like growth factors.

Rats are a commonly used experimental animal in both biomedical and social sciences. They rarely live to be more than twenty-four to thirty-six months of age. Investigators use litter mates (rat pups born to the same mother) to study the effects of GH and IGFs on body weight and the rate of growth. At the end of such an experiment, rats die either of natural causes (old age) or they are euthanized. Investigators then compare, for example, the masses of organs, the dimensions of long bones, or even more physiologic variables such as hematocrit, circulating blood volume, and oxygen carrying capacity of the blood. Over the course of their life span, the GH-treated litter mates gain weight at a more rapid pace. They also achieve final adult weights that are 40 to 60 percent greater than vehicle-treated litter mates. Even though the experiments have not been done, it is reasonable to conclude that similar results would be obtained in higher mammals, including humans. This is one more reason to avoid lifestyle behaviors that can in any way impair the physiological functions of growth hormone.

To estimate the growth potential and final adult height of a child today, science and medicine commonly use the average height of the child's parents as the most reliable indicator. For example, I am approximately five feet eleven inches tall. My wife, the biological mother of our nine children, is approximately five feet one inch tall. We have six sons and three daughters. At the time of this writing, all except our last child (an eighteen-year-old son) have reached their adult heights. My best estimate of their average height is five feet six inches. Our daughters, on the average, are shorter than our sons. Some of our sons are over

six feet and some are under six feet (so parental heights are not an exact predictor).

Many physiological variables change rhythmically. The frequency of change can be measured on a time scale of seconds to minutes (for example, heart rate and respiratory rhythm), by the hour (ultradian, for example, release of GH), by the day (diurnal or circadian, for example, sleep and wakefulness), or by the season/year (circannual, for example, hibernation, march of the lemmings, migrations of orcas, migration of arctic terns). Doubtless there are other physiologically important temporal rhythms waiting to be defined. Chronobiology, or the study of the timing of life's events, is a fascinating and important field. The interested student can go to any search engine, for example, pubmed.com, enter key words that relate to chronobiology, and find studies on a wide variety of physiological factors that are influenced by it. For example, in the field of general medicine, the response of female patients to general anesthesia is influenced by the timing of their menstrual cycles. The time of day a patient takes a prescription medication has an effect on physiological responses to the medication. In febrile patients, body temperatures fluctuate diurnally and are often highest in the early evening hours.

Growth hormone is one of the endocrine products that is secreted on a diurnal basis. This means that if an investigator collected venous blood samples from human subjects every hour on a twenty-four-hour cycle, then measured the concentrations of GH in those samples, the concentrations would vary throughout the twenty-four-hour period. The greatest circulating concentrations would be found during the first few hours of deep sleep (around midnight). The lowest concentrations would be found in the early morning hours (between 4 and 7 A.M.).

Excesses and Deficiencies of GH and Health

According to the *Guinness Book of Records*, Robert Pershing Wadlow (1918–1940) of Alton, Illinois, was the tallest person in history. At the time of his death, Robert was eight feet eleven inches tall and weighed 490 pounds. The Alton Giant or Gentle Giant, as he was known, was six feet two inches tall and weighed 195 pounds as an eight-year-old boy. At age thirteen he became the tallest Boy Scout in history at seven feet four inches tall.

An overactive pituitary gland caused by a tumor in or near the GH-producing cells of the adenohypophysis can lead to two related diseases, gigantism and acromegaly. Gigantism is associated with disease that begins before puberty while acromegaly develops after puberty. The epiphyses are the growth-producing regions of cartilage at the ends of the shafts of long bones such as the femur in the upper leg and the humerus in the upper arm. When the epiphyses fuse with the shafts of the long bones, the long bones no longer increase in length and

one's height is fixed. In an individual with a disturbance in pituitary function, growth and development either before or after adolescence can be markedly altered. Robert Wadlow and others like him probably suffered from gigantism originating from a congenital pituitary disorder.

If the disease is initiated after adolescence (after fusion of the epiphyses and long bones), the affected individual will not increase in height but will increase in size (acromegaly). In particular, bone structures of the face, hands, and feet enlarge. Internal organs are also affected—the heart, liver, and kidneys are enlarged (cardiomegaly, hepatomegaly, nephromegaly). Enlargement and dysfunction of internal organs can continue until the death of the patient, which occurs earlier than in the healthy adult. In the extremities, the fingers and toes get thicker, growing to perhaps twice the size of the fingers and toes of a normal person of similar age. The forehead is enlarged and protruding, particularly around the orbital ridges of the eyes. The nose, lips, and lower jaw can also grow to twice the normal size. In addition to these facial features, other signs of the disease are present such as a deepened raspy voice.

Many years ago my wife and I befriended another couple. As our association developed I could not help notice the signs of acromegaly in this man. Though I am not a small man, at five feet eleven inches in height and 175 pounds neither would I be considered large. My friend was about the same height, but his face, hands, and feet were disproportionately larger than mine. This became most evident when we shook hands. His voice was deep and husky, but he was among the gentlest, kindest men I have known. As our friendship grew, I queried him about his health. Our conversations lead me to conclude that he was not aware of the signs and symptoms of pituitary disease. I asked if he or his wife had ever heard the terms *acromegaly* or *gigantism*. Both said no. I also asked my friend if he had ever had a physician's examination. He replied yes and that he saw his physician once every few years. However, none of the physicians he had seen had bothered to raise the question of acromegaly.

After I had described these two diseases, I gave my friend a list of terms and questions he should take with him the next time he saw his doctor. My friend was referred by his primary care physician to several specialists. They ran tests including CAT scans and MRIs and concluded that he did indeed have acromegaly: the enlarged and tumorous pituitary was revealed in CAT/MRI images and blood samples revealed elevated concentrations of GH, somatomedin, and insulin-like growth factor. My friend and his physicians were then faced with the decision of how to treat the disease. They decided that because of his age, the longevity of the disease, and other factors, he should have the pituitary tumor removed surgically. They chose a hospital where surgeons were expert in performing hypophysectomies (removal of pituitary tumors/glands), and he was treated successfully. After the operation, there were marked changes in both his physical body and in his voice. All his facial features—the heavy

protruding jaw, the thick enlarged nose and lips—as well as the size of his hands and fingers were much less pronounced. His voice was not as deep and raspy. My friend seemed less lethargic as well. He thanked me for his new lease on life.

As a health practitioner or student planning to become one, you must be aware of such clear signs of diseases like gigantism and acromegaly. Once you are aware it then becomes your professional obligation to reveal this information to your unsuspecting patient. You must educate them on the causes and symptoms of the disease so they can think about it in an informed way. Finally, with the patient's input, you must present all the options for treatment. Identifying and diagnosing disease should not be left to patients even if they are reasonably informed physiologists.

Other diseases of the anterior pituitary lead to the opposite effects of gigantism and acromegaly. Moreover, such diseases need not be limited to the pituitary gland. In general, there are two classifications, panhypopituitarism and dwarfism. The former means a generalized decrease in secretion of all hormones of the anterior pituitary. It can be congenital or have its onset at any time of life and can arise slowly or rapidly. It results most commonly from a tumor that destroys the entire pituitary gland. Dwarfism, on the other hand, results from the generalized lack of secretions from the anterior pituitary in early childhood. The limbs, organs, and other body parts grow in relative proportion to one another, but the rates of growth are markedly impaired. A child of ten to twelve years of age might look proportionately balanced but will have the body growth of a four- to six-year-old. The same person at ages eighteen to twenty might have the body development of an eight- to ten-year-old child. In contrast with dwarfism, a person with panhypopituitarism does not pass through puberty and never secretes sufficient quantities of gonadotropic hormones to develop adult sexual functions. In about one-third of individuals with panhypopituitarism, however, only GH is impaired, and these people do pass through puberty and develop adult sexual functions because secretion and release of FSH and ACTH are not impaired. In other forms of dwarfism, the release of GH is normal but there is an inherited inability to synthesize somatomedins, which are necessary for normal growth and development.

Another friend of mine was born with dwarfism. As an adult, this man was about four feet five inches tall. His wife had been married once before and had two sons from that marriage. Several years later my friend and his wife had a baby girl, their first child. The newborn was the age of one of our children. She grew up to be a beautiful young woman and appeared physiologically nearly perfect. She got married in her mid-twenties and had children of her own. Her father's dwarfism seemed to have no effect on her or her children. Her father was also a kind soul (despite the fact that he was employed in a county prosecutor's office).

The Adrenal Gland and Growth

The adrenal cortex can also be problematic and can lead to arrested growth and development. This gland is so named because it is located bilaterally atop the kidneys. The adrenal is composed of two glands, an inner medulla and an outer cortex. They have different cell types, look different morphologically, and secrete different hormone products. The adrenal medulla in humans and other mammals is a specialized type of nerve tissue. Cells in the medulla synthesize, store, and secrete the catecholamines tyrosine, dopamine, epinephrine, and norepinephrine (the latter two are also known as adrenaline and noradrenaline). These compounds are coreleased when the gland is stimulated, but epinephrine is the major secretory product. It is this portion of the gland and its secretory products epinephrine and norepinephrine from which the archaic phrase "fight or flight hormones" has arisen. The flight or fight hormones have multiple effects on the body that prepare one for adverse events including taking exams.

In the cardiovascular system, epinephrine and norepinephrine increase heart rate and contractility. They also increase blood flow to many organs. These are some of the main events in the fight or flight preparation. This function of the gland is closely related to function of the autonomic nervous system. In fact, any sensory signal that prepares the body via stimulation of the adrenal medulla for an adverse event also stimulates the sympathetic division of the autonomic nervous system. This allows the two systems, nervous and endocrine, to compliment one another.

The outer portion of the adrenal gland is called the cortex and in humans is divided into three zones. Each zone produces a different class of hormones collectively called adrenocorticoids. The three main classes are glucocorticoids, mineralocorticoids, and sex steroids. Their names indicate their physiological functions. Glucocorticoids directly affect the metabolism of carbohydrates (glucose being one of the main targets). Therefore they indirectly affect the metabolism of the other energy substrates, protein and fat. Two glucocorticoids are cortisol and deoxycorticosterone. These are classically considered anti-inflammatory agents although in recent years the health food industry has tried to link them negatively with abdominal fat and obesity. Mineralocorticoids affect the physiology of minerals, most notably sodium and potassium ions but also chloride and bicarbonate ions. Specifically they exert a major influence on how the kidneys process these so-called electrolytes. One important mineralocorticoid is aldosterone. Aldosterone's main effect is promoting sodium reabsorption. Its release is also linked to the renin-angiotensin system and to systemic arterial hypotension. The third class, the sex steroids, influence reproductive structure and function. They do this via their direct and indirect actions on target tissues such as the testes and ovaries. Examples of the sex steroids are pregnenolone (leads to the synthesis of progesterone) and androstenedione (leads to the synthesis of androgens).

The adrenal hormones can have a profound effect on growth and development. Their effects are best understood by considering the consequences of excess or deficiency of such hormones. Excess androgens and estrogens before puberty can enhance bone growth. This also accelerates the rate at which the skeleton matures and thus can diminish the time between lengthening of long bones and closure of epiphyseal plates. In many cases, this is revealed as reduced longitudinal growth of long bones, therefore in reduced adult height and stature. This outcome is clearly demonstrated in cases where children are exposed to excess sex steroids at an early age such as might occur from tumors of the adrenal glands or taking sex steroids (for purposes of body building or other athletic performance enhancement) that have been prescribed as medications for their parents or older adults.

Like the sex steroids, the adrenal glucocorticoids in excess can also inhibit growth. Glucocorticoids are used therapeutically to treat serious illnesses of childhood such as asthma and autoimmune deficiencies, but the side effect can be impaired growth, which is not restored to normal until the plasma concentrations of glucocorticoids are returned to physiological levels. Use of GH to treat the impaired rate of growth that occurs with glucocorticoid therapeutics is not effective in restoring normal growth rate. This reveals that the underlying cause of the impaired growth is not due to inefficiency of GH. It is most likely that glucocorticoids impair physiological functions at growth plates, since this is where longitudinal growth takes place. Finally, in recent years the public has become concerned about being exposed to growth hormone through drinking cow's milk. On occasion, dairy cows are treated with growth hormone to increase the cow's production of milk. The potential detrimental effects of drinking milk from treated cows has not been investigated.

Table 3.3 presents a list of some of the known factors that are related to growth and development. This list has evolved in recent decades and is likely to continue expanding as research in life sciences and biomedicine continues.

The Thyroid Gland

Growth hormone and products of the adrenal glands are not the only endocrine hormones that affect growth and development. Generally any neurohumoral agent that affects metabolism and use of energy substrates will affect growth either directly or indirectly. Thyroid hormones are another good example of endocrines that influence growth and development. Thyroid hormones have two important functions. In growing animals and humans, they are crucially important in development especially of nervous tissue. In adult animals and humans, the thyroid hormones act to maintain metabolic homeostasis and thus affect the functions of virtually all living tissues. To meet these requirements, the thyroid gland is able to synthesize and store large quantities of thyroxine

TABLE 3.3

Short list of various physiological growth factors

Acronym	*Full name*	*Target tissues*
AGF	angiogenic growth factor	growth of new blood vessels and related tissues
EGF	epidermal growth factor	growth of skin and subcutaneous tissues such as fat
FGF	fibroblast growth factor	growth and development of cells and tissues in general
HGF	hepatocyte growth factor	growth and development of liver cells and liver sinusoids
IGF-I, II	insulin-like growth factors	mediate effects of growth hormones
NGF	nerve growth factor	growth and healing of nerves and supportive tissues
VEGF	vascular endothelial growth factor	growth and development of endothelial cells that line blood vessels

(T_4) and triiodothyronine (T_3). Metabolism of the thyroid hormones takes place primarily in the liver although some metabolism takes place in the local target tissues such as the brain. Circulating plasma concentrations of thyroid hormones are regulated precisely by the anterior pituitary hormone TSH (also known as thyrotropin) in a classic negative-feedback mechanism. The predominant actions of thyroid hormones occur by binding to nuclear receptors and through modulating transcription of specific genes.

Disorders of the thyroid gland are common and consist of two general types: changes in morphology of the gland and changes in the secretory activities of the gland. Thyroid nodules and goiter are the most common pathologies and can be caused both by benign and malignant tumors (morphology). Such tumors might or might not influence secretory behavior of the gland. Conversely, patients appearing with overt hyper- or hypothyroidism often present with marked clinical manifestations. These are not always accompanied by changes in morphology but are associated with marked fluctuations in secretory products. For example, global congenital hypothyroidism due to iodine deficiency remains the major cause of mental retardation and is particularly problematic in many third-world countries. Iodine is an absolute requirement of the thyroid gland's ability to synthesize thyroxines. Screening of newborns coupled with the thyroid-related institution of hormone replacement therapy has markedly decreased the incidence of mental retardation in the United States.

Effective treatment of most thyroid disorders is readily available. Treatment of the hypothyroid patient is straightforward and consists of hormone replacement. However, treatment of hyperthyroidism in adults is more complex and includes blocking the synthesis of thyroid hormones, surgical removal of the gland, and chemical destruction of the gland by administration of radioactive iodine. Still, treatment of thyroid disorders is extremely satisfying as most patients can either be cured or have their disease controlled.

The thyroid gland is located anteriorally at the base of the neck and is attached by connective tissue to the trachea and surrounding tissues. Its main secretory products are thyroxine (T_4) and triiodothyronine (T_3). Their synthesis and release are under the control of TSH from the anterior pituitary, which is regulated by the hypothalamus by thyrotropin-releasing hormone or factor, TRH/TRF. Thyroxine is the major secretory product and, like GH, influences many bodily functions. Most notably it is involved in the homeostasis of metabolism and internal body temperature. An overactive thyroid (hyperthyroidism) is the most common cause of the disease thyrotoxicosis, characterized by excess activity, insomnia, and inability to gain weight/excess loss of weight. A definitive diagnosis is obtained by accelerated excess uptake of radiolabelled iodine during a twenty-four-hour period. Grave's disease or toxic diffuse goiter is the most common cause of excess uptake of iodine and accounts for a majority of the cases. Grave's disease, like Hashimoto's thyroiditis, is caused by an abnormal immune response that includes the production of antithyroid antibodies against the thyroid follicular cells, microsomes, and TSH receptors. Many of the signs and symptoms of thyrotoxicosis stem from the excessive production of heat, from increased motor (physical) activity, and from activation of the sympathetic nervous system. Activation of the sympathetic nervous system can lead to cardiovascular effects such as elevated heart rate, cardiac enlargement, reduced resistance to blood flow, and increased pulse pressure. The increase in basal metabolic activity produces excess appetite and, if food intake is insufficient, to loss of weight. Left undiagnosed and untreated, people suffering from hyperthyroidism may be misunderstood as hyperactive type-A personalities. In older patients and/or as the disease progresses, angina, cardiac arrhythmias, and heart failure may be present.

Hypothyroidism, also known as myxedema when severe, is the most common disorder of thyroid function. An underactive thyroid can result from hormonal insufficiency at one or more levels of the neuroendocrine axis. The neuroendocrine axis involves cranial structures such as the hypothalamus and neurons descending into it, the pituitary stalk and glands, and more peripheral endocrine glands like the thyroid, adrenals, and pancreas. Failure of the thyroid to produce sufficient hormone is called primary hypothyroidism. Central hypothyroidism is a CNS problem and results from failure of the pituitary to release TSH (secondary hypothyroidism) or from failure of the hypothalamus to release TRH

(tertiary hypothyroidism). Central hypothyroidism is much less common than primary hypothyroidism. Associated symptoms in these individuals are usually the opposite of hyperthyroidism. They often display reduced metabolic rates. Their consumption of oxygen and production of carbon dioxide are below the norms for their age and gender, they are intolerant to cold, and have unexplained gain of weight, lethargy, and generalized apathy. The face is often expressionless, puffy, and pallid. The skin is cold and dry and the hair is coarse, brittle, and sparse. The voice is hoarse and low-pitched, speech is slow, and the appetite is suppressed leading to gastrointestinal inactivity and regular constipation.

Because the thyroid gland has such far-reaching effects throughout the body, its surgical removal (to treat hyperthyroidism) or use of replacement therapy (to treat hypothyroidism) can lead to long-standing imbalance in homeostatic regulation. Changes in body temperature can be marked and sustained for days or weeks. The patient's sensations of being warm and cold can fluctuate several times during a single twenty-four-hour period. It might take months for a physician to titrate the dose of hormone to bring this single physiological variable under control. A similar picture is painted for the regulation of appetite, thirst drives, sleep, and physical activity.

Our sixteen-year-old son complained of being tired and having little energy. We initially thought this was due to his relatively poor diet that included too much sugar and too few vegetables and fruits. Because his complaints came early in the track season, we thought his lethargy might be related to the transition from an unfit to a more fit state. But after adjusting to training and conditioning, and changing to a diet of more fruits and vegetables and less sugar, he still complained. My wife made an appointment with a pediatrician, and his blood chemistry was evaluated. His T_3 and T_4 concentrations suggested the possibility

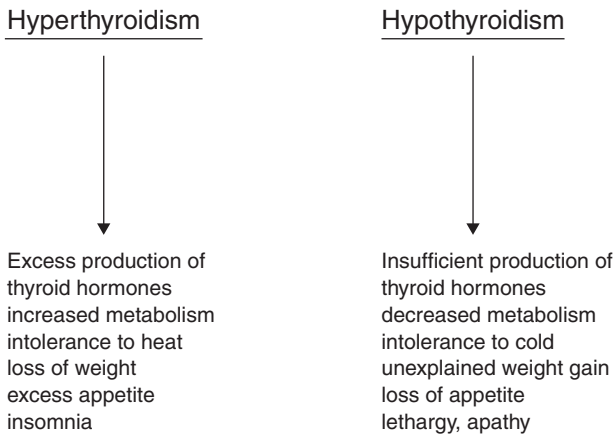

FIGURE 3.2 Classic symptoms associated with the diseases hyperthyroidism (left column) and hypothyroidism (right column).

of hypothyroidism, so he was sent a month later for more blood work. This blood screen included measures of TSH (thyroid-stimulating hormone) and thyroglobulins (proteins that bind thyroid hormones). TSH concentrations were high and thyroglobulins were normal. This was confusing since elevated TSH would normally be expected to augment the production and release of T_3 and T_4. The pediatrician also mentioned decelerated growth rate in our son starting at about ages thirteen or fourteen years. One of his favorite role models was his older brother. Indeed the record of their growth throughout the teenage years was marked on the same scale on trim around the door frame. We could see from these data that our younger son was falling behind his older brother's growth rate. These observations were consistent with the possibility of hypothyroidism.

More tests were prescribed, and both the results and the physician's interpretations were inconsistent and even more confusing. To add to the confusion, our son was by this time getting up at 6 A.M. to go to a summer job and was a dedicated weight lifter. He had started training at age fourteen and was still working out three years later. We began to question whether the early weight lifting had impaired his growth but felt the main problem was our inability to get conclusive clinical blood chemistry. Physicians are not scientists and are not trained in the scientific method. Moreover, insurance companies and HMOs strongly influence a physician's ability to prescribe tests and make referrals. Therefore the blood screens were infrequent, random, and disjointed. They were not completed under the systematic, disciplined conditions that would be expected in an endocrinologist's or physiologist's laboratory and trying to make sense of them was an exercise in futility.

Carefully observing our son for the next few years and making decisions with his input seemed like the best course of action. We knew that during the interim he would gain his adult stature and that this could affect his condition. Moreover, there was some evidence of cardiac disease in our family, and this had shown up in the son with whom the growth of our youngest son was being compared. With his consent, we decided to give no further attention to the matter until he reached young adulthood and his full adult stature. His complaints of lethargy had diminished markedly by the time we reached this decision.

4

The Cardiovascular System and the Blood

Homeostasis in the mammalian cardiovascular system depends importantly on the interactions among blood pressure, blood flow, resistance to blood flow, and other hemodynamic variables. Moreover, there are several important reflexes such as the baroreceptor reflex and the Bainbridge reflex that help maintain an equilibrium in the above hemodynamics and that try to restore homeostasis when it is disturbed.

The mammalian cardiovascular system is best understood by analyzing its component parts. In the simplest terms, these are the heart, the blood vessels, and the blood. The hearts of all mammals have four chambers: a left and a right atrium and left and right ventricles. The atria are separated by a thin wall of muscle called the interatrial septum and the ventricles are separated by the thicker interventricular septum. There are effectively two cardiovascular systems in mammals: the pulmonary and systemic (peripheral) circulatory systems. Each is composed of a pump (right or left ventricle) and a set of blood vessels. The ventricles function as the pump mechanism for generating pressure and volume work, and need to create high enough internal pressures to overcome the arterial resistance to blood flow. The ventricles must also eject a sufficient volume of blood to initiate and sustain circulation. The main differences between the two cardiovascular systems are that (1) the pulmonary circulatory system supplies blood flow only to the lungs whereas the systemic circulatory system provides blood flow to all remaining organs and tissues, and (2) the pulmonary is a low-pressure system while the systemic is a high-pressure system. The pulmonary cardiovascular system begins at the right ventricle and ends at the left atrium. The systemic or peripheral cardiovascular system begins at the left ventricle and ends at the right atrium. All other organs and tissues of the body are interposed between these two chambers. All vessel components between the two end points in either system are an integral part of that particular system.

This means that under normal conditions, maximum ventricular/arterial pressures in the systemic circulatory system are about 100 mmHg whereas corresponding pressures in the pulmonary circulatory system are about 25 mmHg. This is important because the systemic circulation and left ventricle work against a high resistance. The pulmonary circulation and right ventricle do not.

The Blood

Blood is a tissue. A tissue is a group of like-behaving cells that serve a common purpose. In the case of blood, the purpose is delivery of nutrients and gases and removal of by-products of metabolism. Blood is composed of a variety of cell types found in a fluid matrix. The cells are red blood cells (RBCs or erythrocytes), white blood cells (WBCs or leukocytes), and blood platelets (fragments of larger, more primitive blood-forming cells). In adult humans, neither the platelets nor the RBCs have nuclei. All WBCs have nuclei, but the nuclei have a wide range of morphologies and staining characteristics. Some do not stain and are called neutrophils. Others bind basic stains (blue) and are called basophils. Still others bind acidic stains (red) and are called eosinophils. The fluid matrix in which the cells are suspended is called the plasma. By both weight and volume, plasma is mostly water, about 90 percent. Because plasma is mostly water and because water is the universal chemical solvent, plasma is considered a solvent. It contains a wide range of solutes including but not limited to macromolecules such as the immunoglobulins (antibodies) and other proteins, electrolytes, vitamins and minerals, carbohydrates, lipids, and amino acids. One of the important constituents is fibrinogen, a protein that is essential to the blood-clotting mechanism. Under conditions where increased clotting is necessary, fibrinogen is broken down to fibrin monomers that polymerize to help form fibrin clots. These serve as vascular plugs to minimize loss of blood from severed blood vessels. When fibrinogen and its breakdown products are removed from the plasma, it (plasma) will no longer clot and is then known as serum. The two terms *plasma* and *serum* are frequently misused. Clinically speaking, if withdrawn plasma is stored and intended for reinfusion, it is either anticoagulated or has the fibrinogen removed before storage. In this latter case, the plasma becomes serum.

The health of the blood is a good indicator of the overall health of the body, and it can be assessed in a number of ways. One of the most important quantifications of the blood is what we call hematocrit. Hematocrit is the percentage of whole blood (cells plus plasma) that is composed of cellular elements, mostly RBCs. It is determined by withdrawing a sample of peripheral venous blood from, for example, the vein between the forearm and biceps. The sample is then centrifuged to separate cellular elements from the plasma. Hematocrit is demonstrated in figure 4.1. The heavier RBCs settle and compact at the bottom of the centrifuge tube while the lighter-molecular-weight plasma fills the top half.

Between the two volumes is a narrow layer called the buffy coat. This volume is composed of WBCs and platelets.

Note in figure 4.1 that the hematocrit (erythrocytes or red blood cells) is approximately 45 percent. This is physiologically normal. A reasonable range for hematocrits in adult males is 42 to 46 percent and slightly lower in menstruating women. Diseases and other conditions that elevate hematocrit are called polycythemias and those that decrease hematocrit are known as anemias. Many conditions that lower the amount of oxygen in the circulation (hypoxia) stimulate erythropoiesis, or the enhanced production of erythrocytes. Hypoxia-induced erythropoiesis in turn elevates the hematocrit. One of the most well-known erythropoietic stimuli is spending time at high altitudes. Mountain climbers who reside at sea level but periodically venture to altitudes above 15,000 to 25,000 feet, the height of the world's highest mountains, have hematocrits in excess of 50 percent while they are at the high elevations. The increase in red blood cells compensates for the lower partial pressures of atmospheric oxygen and thus helps maintain oxygen homeostasis. When these sojourners return to sea level,

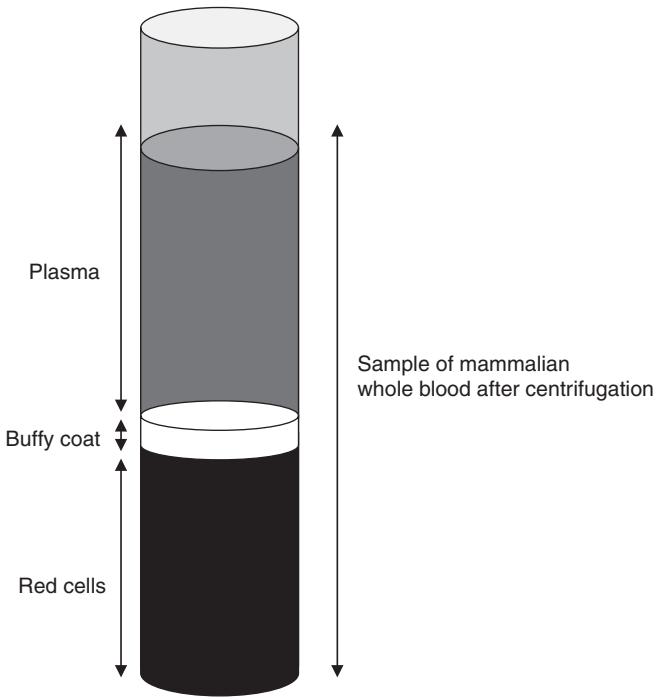

FIGURE 4.1 A cylinder illustrating the blood hematocrit. The height of the column of fluid inside the cylinder (right arrow) illustrates the whole blood sample after centrifugation. Note there are three separate layers, the bottom layer consisting mostly of packed red blood cells, the second of white blood cells and platelets (about 1 percent of the total volume), and the top layer of cell-free plasma.

hematocrit is restored to its sea level value. People who live or spend long periods of time at high altitude typically have hematocrits in excess of 55 to 60 percent.

Mature red blood cells in adults lack nuclei. In terms of their morphology or multidimensional shape, erythrocytes are biconcave disks. Cross-sectionally, they look like donuts and are thickest at the perimeter and thinnest in the center. The physiological explanation for this design lies in the function of the cell. Arguably the most important molecular component of the red blood cell is the protein hemoglobin (Hb). Hemoglobin is the main protein carrier of oxygen in the circulation. Hemoglobin molecules are most densely concentrated around the edges of the cells, giving the perimeter its thicker shape. This is important for at least two reasons. First, it decreases the distance that oxygen must diffuse to get from the red blood cell to the tissue or from the pulmonary alveoli to the red blood cells. This facilitates exchange of gases. A second purpose for the biconcave design of the RBC is flexibility. The outside diameter of RBCs is often equal to or modestly greater than the inside diameter of the smallest blood vessels through which they must pass. Were the cells not flexible, they could not conform to the interior dimensions of the microcirculation and would become trapped, thereby impeding blood flow.

The flexibility of RBCs is analogous to rides at water parks. Many of these rides involve curved and tortuous cylinders of a few feet in diameter. These slides with water flowing through them descend from heights of forty to eighty feet to ground level. Riders sit or lie on inflated rubber tubes. Some tubes have space for only one rider; others can accommodate two or three riders. Because the tubes are rubber and inflated, they are flexible. They deform and reshape according to the weight of the rider and the contour of the cylinders through which they pass. If they were inflexible, they could become jammed in the cylinders, thus preventing their passage as well as that of subsequent riders.

When evaluating the health of one's blood by measuring concentrations of electrolytes such as Na^+, K^+, and Ca^{2+}, plasma is used. Electrolytes are maintained at near-constant levels by homeostatic control mechanisms. These mechanisms involve their dietary intake, renal excretion, and distribution and balance among the various body water compartments. Extracellular potassium is normally about 4–5 milliequivalents per liter in both plasma and interstitial fluids. A doubling of this concentration—elevations to 8–10 milliequivalents per liter—can be life threatening. The heart is extremely sensitive to circulating concentrations of potassium. Imbalances can lead to disturbances in rhythmicity as well as to sudden death. Loss of calcium homeostasis can also be dangerous. Calcium is an important second messenger, and disturbances in its transmembrane distribution can lead to disruption of cellular signal-transduction pathways in all cell types.

The volume of blood in circulation at any moment in time is maintained at relatively constant levels. This takes place through sensitive feedback mechanisms involving all the components described earlier. For example, if blood

volume is suddenly increased appreciably above physiological levels, stretch receptors in the blood vessels and cardiac chambers as well as elsewhere in the body are activated. This leads to increased renal excretion of water, inhibition of thirst mechanisms, and subsequent restoration of blood volume. The converse happens when blood volume is suddenly decreased such as with hemorrhage.

Blood volume is also regulated hemostatically. Hemostasis is the ability of blood to clot. Circulating blood volume in adults ranges from about four to eight liters and can be estimated by multiplying one's body weight by 0.08—blood volume is an estimated 8 percent of body weight in mammals. I weigh 175 pounds or 80 kilograms. Eight percent of 80 kilograms is 6 kilograms. One kilogram of whole blood (1,000 grams) has a density or mass per volume of about one gram per milliliter. Therefore one kilogram of blood is approximately equal to one liter of blood. So my circulating blood volume is about six liters.

Any injury to one or more blood vessels is a potential threat to circulating blood volume. If blood volume decreases, blood pressure also decreases. If blood pressure decreases, blood flow falls. If blood flow diminishes, so also will the delivery of oxygen and nutrients to the tissues as well as the removal of waste products from them. Hemostasis is the body's design to minimize and/or prevent loss of circulating blood volume. When a blood vessel is severed and blood loss occurs, a physiological sequence of events takes place to minimize the loss. The first event is vasoconstriction of the severed vessel and those in the immediate vicinity. Vasoconstriction happens because endothelial cells in the walls of the damaged blood vessels release chemicals that cause smooth muscle cells in the wall to contract. When the smooth muscle cells contract, the inside diameters of the vessels are reduced, that is, the vessels vasoconstrict. The sympathetic neurons in the region of the damage help mediate this process by releasing neurotransmitters that cause vasoconstriction. The second event is the formation of platelet plugs. Other signaling molecules from the region of damaged tissue attract blood platelets to the wound. The platelets adhere to the luminal endothelial cells adjacent to the wound and establish a cargo-net-like meshwork of polymerized fibrin molecules that begin trapping other platelets, RBCs, and WBCs. This is referred to as a platelet plug. The chemistry of the platelet plug is beyond the purposes of this book, but it involves a cascade of events importantly involving conversion of fibrinogen to fibrin and thrombinogen to thrombin. Finally, the third step is the entrapment of many RBCs to form a blood clot. Subsequent extraction of plasma from the meshwork strengthens the clot, and the loss of blood to the outside is reduced, then stopped.

The Heart and Cardiac Cycle

The typical heart illustrated on Valentine's Day cards shows the apex as the pointed end of the heart and a base that is the fatter upper end of the heart. The

mammalian heart bears little resemblance to the idealized Valentine's Day card although one can see the similarities. If you look at your chest and place your hand (fingers close together and thumb extended) on edge and obliquely from the upper right side toward the lower left side and across the sternum, you will have defined the position of the heart inside the thoracic cavity. The base of the heart is near the base of your hand where it joins the wrist. The apex of your heart is near the tip of the small finger where it contacts the chest wall.

Among the mammalian organs, the human heart is special in many ways. The rhythmic contractile activity we call the heart rate is based on the performance of a specialized conduction system that originates in the right atrium and terminates at the base of the ventricles. Structurally and morphologically, the heart is composed primarily of striated muscle cells, though it is cardiac striated muscle as opposed to skeletal striated muscle. Adjacent cells are interconnected at specialized regions called intercalated disks. These do not occur in skeletal muscle, and they give the heart the ability to function as a syncytium, that is, as if it was a single cell. This means that when one cell contracts, all others contract nearly simultaneously. The exterior surface of the heart is called the epicardium and is composed of tough epithelial cells. The internal surface of the cardiac chambers is called the endocardium. These are flat, smooth cells that form the interface between the blood and the wall of the heart. Everything located between these two layers of tissue is called the myocardium—whether in the atria or ventricles. The myocardium can be further subdivided into the subepicardium, the midmyocardium, and the subendocardium. This classification of layers helps cardiologists, thoracic surgeons, and pathologists describe damage caused during heart attacks and other disease states. It helps physiologists define their experimental investigations. For example, a myocardial infarction, or heart attack, might have been confined to the subendocardium at the base of the posterior left ventricle. Myocardial infarcts of the left ventricle are more common than those of the right ventricle. Full-thickness infarcts—those that extend from the subepicardium to the subendocardium—are more serious than those that involve only the subepicardium, midmyocardium, or subendocardium. Infarcts of the left ventricular lateral free wall are more common that those of the left ventricular posterior free wall. This is all importantly dependent on the architecture of the coronary vasculature and the degree of blockage of these blood vessels. Prognosis varies with the individual, the size and location of the infarct, timing of postinfarction treatment, treatment itself, surgical and cardiological teams treating the patient, and so on.

The normal mammalian heart is found inside a membranous sac called the pericardium. The pericardium is composed of several layers of epithelial cells and tough connective tissue. There is generally a small amount of fluid inside the pericardium. This serves to lubricate the interior surface of the sac and the outer surface of the heart where they come in contact. Friction between the two

tissues is thereby reduced. Too much fluid can lead to compression and disease of the heart.

The specialized conduction system of the heart is composed of modified muscle cells (not nerves). It begins in the posterior wall of the right atrium where it joins with the superior vena cava, the primary vein carrying blood from the head, neck, and regions of the upper body to the right atrium. In humans, there is a small specialized region of tissue at that location called the sinoatrial (SA) node. It is also known as the pacemaker of the heart. Here a single cell depolarizes and then near-simultaneously activates all surrounding syncytial cells. An action potential is created that spreads throughout the right and left atria. This leads to electrical activation then contraction of the atria. When the atria contract, they empty blood into the corresponding ventricles. The wave of electrical activity then converges at another specialized group of cells called the atrioventricular (AV) node. The AV node is located at the basal terminus of the interventricular septum. It is approximately fifteen millimeters long and ten millimeters wide and extends into the interventricular septum as the bundle of His (pronounced *hiss*). Because of the cellular construction of the AV node, it is a tissue that is especially sensitive to any obstruction of the action potential and to the influences of disease and drugs. Indeed, disturbances in rhythmicity of the heart can often be traced to the AV node. The bundle of His divides into left and right branches that descend down the left and right sides of the interventricular septum, eventually branching multiple times to become the Purkinje fibers. Purkinje fibers interconnect with muscle fibers of both ventricles and terminate at the base of the heart. This specialized conduction system originates at the top of the right atrium and terminates at the base of the ventricles and is able to conduct an action potential throughout the entire heart in less than one-half of one second. The frequency of depolarizations of the SA node is what determines heart rate or rate of ventricular contraction.

I want to stress that the heart's specialized conduction system is composed of muscle cells, not neurons. That is not to say the conduction system is not influenced by the nervous system; it is. The autonomic nervous system innervates the heart and is especially abundant in the conduction system. Both branches of the autonomic nervous system influence the atrioventricular conduction system. Postganglionic fibers of both the parasympathetic and the sympathetic nervous systems are generously distributed to the SA and AV nodes. However, sympathetic fibers are more abundant in the SA than AV nodes and parasympathetic fibers are more abundant in the AV than SA nodes. The bundle of His and the Purkinje fibers are also influenced by both autonomic branches. Conversely, only the sympathetic nervous system innervates the walls of the ventricles to any appreciable degree. As a general rule, activation of sympathetic fibers increases heart rate and cardiac contractility. Activation of parasympathetic fibers slows heart rate but does not affect contractility.

Electrical activity in the heart causes mechanical activity. That is, generation of action potentials in the SA node leads to contraction of both the atria and ventricles. The coordination of electrical and contractile (mechanical) events in the heart is known as excitation-contraction coupling (E-C coupling). In more recent years, it has also been called electromechanical coupling or electromechanical induction. Coupling of these two events and the physiological actions of calcium ions are intimately interconnected. As graduate students, my peers and I were taught that the involvement of calcium was the final common pathway in E-C coupling. We were told that stimulation of sympathetic nerves led to enhanced heart rate and contractility and that the neurotransmitter norepinephrine (released by the SNS postganglionic axons) acted via calcium in mediating these actions. My instruction as a graduate student took place in the early 1970s prior to the current revolution in life sciences and the era of modern molecular biology. At that point in time, we thought that calcium release and its interaction with contractile proteins was indeed the final step in E-C coupling. As the era of modern molecular biology has expanded our understanding, we can now speak of the roles of stimulatory and inhibitory G proteins, calcium-induced calcium release, ryanodine and dihydropyridine calcium receptors, and many other components of signal-transduction pathways that were unknown in the early 1970s. The take-home message for the student of the early twenty-first century is this: enjoy your formal education as you receive it, but realize that what you learn in physiology and medicine today will be markedly influenced by what those who come after you learn and do in experimental laboratories ten, twenty, and forty years from now.

The coronary circulation provides blood flow to the heart and all its associated tissues. Of the total volume of blood that the heart pumps to the entire body, the coronary circulation under baseline conditions receives only about 5 percent. Considering the tremendous volume of work the heart does over a lifetime, it is certainly one of the most underperfused organs of the body. That means that the ratio of its supply of oxygen and important nutrients to its demand for these is among the lowest of all tissues. Functionally, this is not as important for amphibians and other lower vertebrates and invertebrates as it is for mammals. Some lower forms of animals including amphibians are able to extract oxygen from the blood in the chambers of the heart. As far as we know, mammals are not able to do this. Thus for their oxygen supply, mammals must rely on a well-developed circuit of coronary blood vessels, as well as mechanisms that proportionately increase coronary blood flow as oxygen supply dwindles and as oxygen demand increases.

In humans there are two main coronary arteries. They are called the right and left main coronary arteries. They are the first blood vessels to arise near the root of the aorta. Because the left ventricle contains much more tissue, hence mass, than the right ventricle, the left main coronary artery carries most of the

heart's blood flow. The two main coronary arteries are each only a couple centimeters long. Both branch into smaller arteries that provide blood flow to the anterior and posterior walls of both ventricles. The left ventricle receives most of its blood supply during the resting phase of the cardiac cycle while the right ventricle receives the greatest fraction of its flow during the contracting phase of the cycle. Thus the regulation of the coronary circulation is an important subject of investigation and has far-reaching clinical implications. One of the main avenues of investigation in both the basic physiology laboratory and by the pharmaceutical industry has been the identification and development of coronary vasodilators. This has required physiologists and pharmacologists to identify protein receptors in the membranes of vascular smooth muscle and endothelial cells. Once these proteins are identified, purified, and sequenced, organic chemists are able to synthesize compounds that will bind to them that either relax vascular smooth muscle directly or cause release of other agents that are coronary vasodilators. Then physicians can test these agents under controlled clinical conditions.

The average resting heart rate of adults ages twenty-one and above is about seventy-two beats per minute. There are two subdivisions of each cycle: a resting phase called diastole and a contracting phase called systole (see table 4.1). The majority of the time of a single cycle is spent in diastole. Physiologists believe the reason for this is to allow adequate time for the ventricles to fill with blood during the resting phase of each cycle. The filled volume replenishes the

TABLE 4.1

Phases of the mammalian cardiac cycle

Phases (subphase)	*Abbreviated physiological activities*
Systole	
isovolumetric contraction	valves closed, ventricular volumes constant, pressure/tension increasing rapidly
rapid ejection	arterial valves open, much of stroke volume ejected, arterial pressures rise rapidly
reduced ejection	stroke volume completed, arterial pressures stabilize, then begin to decline
Diastole	
isovolumetric relaxation	valves closed, arterial pressures fall rapidly, pressure/tension decrease rapidly
rapid filling phase	atrial valves open as ventricular pressure drops below atrial pressure, rapid filling begins
reduced filling phase	ventricles finish filling, atrial pressures rise, valves close, and cycle repeats

volume of blood that was ejected during the previous systole and ensures development of sufficient wall tension and ventricular pressure to eject another volume during the subsequent systole. If diastole is shortened appreciably, then the time of filling is shortened correspondingly. This causes a reduction in the volume of blood that can subsequently be ejected during systole. If such a condition lasted for more than a few cardiac cycles, blood flow to the body would be compromised and tissue dysfunction, damage, and death could result. Both diastole and systole can be further subdivided into rapid and reduced filling and emptying phases. That is, most of the blood that will fill the ventricles during the 550 to 600 msec of diastole enters in the early milliseconds (the rapid filling phase, that is, the first 100 to 200 msec). The rest of the blood enters during the latter portion of diastole (the reduced filling phase). A similar phenomenon is seen during systole. In experimental cardiovascular laboratories, these subphases can be characterized and quantified by measuring pressures, volumes, and rates of blood flow into and out of both ventricular chambers. Cardiovascular disease can thus be characterized according to such data.

Stimulating sympathetic nerves that innervate the heart by means of either electrical instruments or by drugs causes both an increase in heart rate, that is, a positive chronotropic response, and an increase in contractility, or a positive inotropic response. The duration of a single cardiac cycle will be reduced proportionally. Stimulating excessively can double or even triple heart rate in some mammals. In humans if heart rate accelerates to 175 to 200 cycles per minute, the amount of time spent in diastole during each cycle is dangerously reduced thus impairing ventricular filling. Conversely, stimulating the parasympathetic nerves can markedly reduce heart rate. The parasympathetic nerves are carried to the heart by the tenth cranial or vagus nerve. Excessive stimulation of the vagus nerve causes momentary cardiac arrest, that is, the heart actually stops beating for several seconds. Syncope or fainting can accompany such an event. A complete description of all the maneuvers, tests, and drugs that can influence heart rate and function via either the sympathetic or parasympathetic nerves is beyond the scope of this book. However, students should become familiar with the common ones and should even query their cardiologists should they be seen for heart-related problems. Moreover, practitioners should be sufficiently concerned for the health of their patients to teach a few of these principles to them.

The Blood Vessels

The third major component of the cardiovascular system is the blood vessels. Structurally and functionally, blood vessels can be subdivided into several types, most notably arteries, capillaries, and veins. They all have properties in common, that is, endothelial cells that form the interface between flowing blood and the vessel walls, but they are also uniquely different: for example, arteries

withstand high pressures, veins have high compliance, and capillaries have only a single layer of cells in their walls. The luminal dimensions or inside diameters of all except true capillaries are under both metabolic and neurogenic control. This gives all blood vessels except true capillaries the capacities to actively and passively contract (vasoconstrict) and relax (vasodilate).

Since blood flows sequentially through the different segments of blood vessels, I will describe them and their functions in sequential order. Arteries can be thought of as conducting, distributing, and resisting. The largest of the conducting arteries are those that arise at the base of the ventricles; the aorta at the base of the left ventricle and the pulmonary artery at the base of the right ventricle (see figure 4.2 for schematic). These two conducting arteries give rise to downstream conducting arteries of progressively narrowing inside diameters,

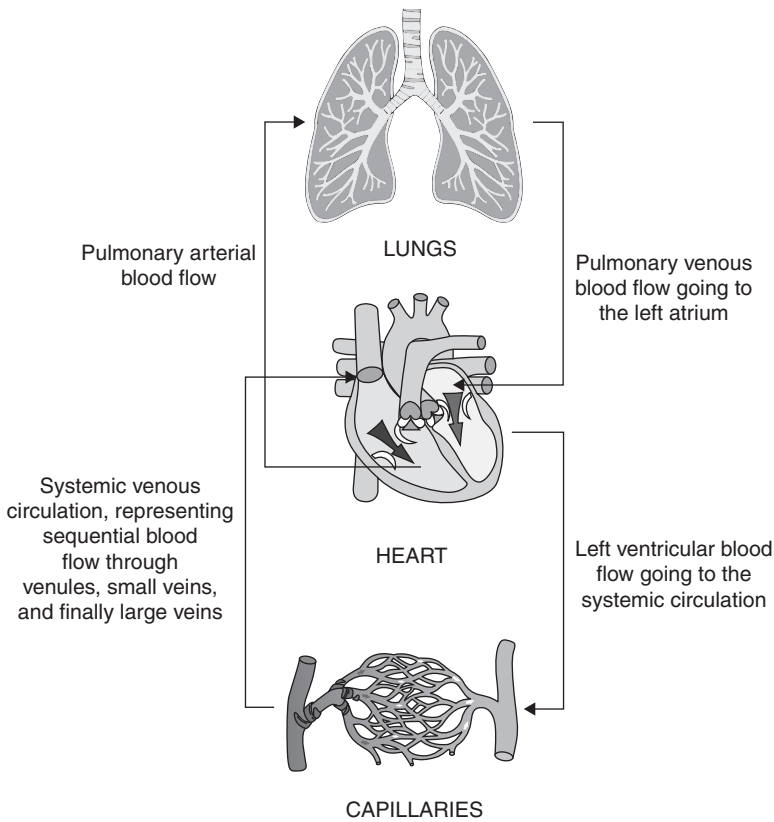

FIGURE 4.2 The mammalian cardiovascular system and the relationships of the two circulatory systems (pulmonary and systemic) to the heart and to each other. Tails and heads of arrows depict the origins and terminations, respectively, of blood flow in the two circulatory systems. Note also that the heart has four chambers, left and right atrium and left and right ventricle.

for example, subclavian arteries, brachiocephalic arteries, and iliac arteries. Conducting arteries convey blood away from the ventricles and towards the upper/lower regions of the body and extremities. They are the most compliant among the arteries. Compliance is computed differently but basically refers to how much volume a segment of vessel can accommodate with little corresponding increase in pressure. Large veins are the most compliant among the vessels and resistance arteries are the least compliant.

Distributing arteries are those that arise from conducting arteries—they are the second segmental arrangement of arteries—and they carry fractions of the total blood supply to the various organs and tissues. They distribute blood flow among and within the organs and tissues. Some of the distributing arteries are cerebral arteries (for example, the middle cerebral artery or MCA), coronary arteries (the left anterior descending coronary artery or LAD), renal arteries (for example the interlobular artery or ILA, radial arteries, RA), and femoral arteries (FA). Distributing arteries carry varying fractions of the total blood supply at any given moment. Total blood supply refers to cardiac output, the volume of blood that either ventricle ejects per minute of time. Quantitatively, cardiac output is expressed in liters per minute and is symbolized as Q = l/min. The average adult male weighs about seventy kilograms and has a Q of about five to six liters per minute (women proportionally less). The distributing arteries carry changing fractions of this volume of blood to their corresponding organs. For example, in a hot arid environment, the body might need to eliminate excess heat. In this condition, blood flow to the skin and sweat glands will increase while that to the kidney might decrease. Likewise, where hemorrhage and blood loss threaten one's health, blood flow will be preferentially distributed to the heart, brain, and a few other essential tissues at the expense of organs like skin, gut, muscle, and kidneys, which receive disproportionately less flow during such crises.

The distribution and redistribution of blood volume and blood flow are among those cardiovascular variables that apply to the negative feedback control systems referred to in chapter 1. Mechanoreceptors, which detect changes in the deformation of a vessel wall or an atrial/ventricular chamber, respond to changes in blood volume. Changes in blood volume are often associated with corresponding changes in blood flow. If a heart chamber or a blood vessel is gorged with blood, the mechanoreceptors will be stretched, thus detecting a change in volume. An afferent sensory signal will be relayed to the central nervous system. An error detector will compare this signal with the set point for that variable and will make appropriate adjustments. This corrected response will be relayed back to the blood vessel by a motor nerve, causing the vessel to constrict. As the vessel constricts, its capacity to retain the added volume will decrease and a new steady state will be achieved. With the new steady state comes a withdrawal or termination of the initial sensory signal that activated the feedback control system.

Resisting arteries are more correctly called resistance arteries or arterioles. Physiologists almost uniformly call them arterioles. Arterioles are the third segment of arteries in the arterial tree. They are so named because of the marked drop in blood pressure that is measured across them. That is, pressure at the upstream end of an arteriole is considerably greater than pressure at the downstream end. This is much less the case for distributing and conducting arteries where there is little change in pressure from one end to the other. For example, pressures through a large conducting or distributing artery might start at 100 and drop to 90 mmHg. For the typical arteriole, corresponding numbers might be 75 and 35 mmHg. The pressure drop in the latter case is four times that in the former case. Since the total volume of blood passing through all arteries in the body is the same, and since resistance is calculated by dividing differences in pressure by blood flow, in this case resistance to flow would be four times greater in the resistance arterioles. The walls of arterioles are composed of much more vascular smooth muscle than are the other classes of arteries. It is the vascular smooth muscle in the wall of a blood vessel that determines its capacity to contract and relax or to cause baseline resting tone or relative degree of contraction.

Segmentally, or in series, the next downstream set of blood vessels are the capillaries (see table 4.2). Capillaries are the smallest set of blood vessels in the

TABLE 4.2

Segmental arrangement and function of mammalian blood vessels

Vessel segment	*Physiological functions*
Arteries	
conducting	aorta, pulmonary, major downstream branches; transiently store blood during systole, recoil and contract during diastole, deliver blood to distributing arteries
distributing	renal, coronary, cerebral; distribute fractions of cardiac output to organs and tissues
arterioles	resistance vessels; determine distribution/partitioning of blood flow to tissues based on their needs for oxygen and nutrients (resting states vs periods of physical activity; e.g., dynamic exercise)
Capillaries	major, but not only, site for exchange between blood and tissues
Veins	
venules	adjacent to capillaries, receive capillary blood, help determine exchange, some resistance
small veins	collect blood flow from venules and microcirculation
large veins	deliver blood flow to atria, store blood, serve as adjustable reservoirs of blood volume

body. The average capillary is less than 25 micrometers in length and has an internal diameter of less than 10 micrometers (recall that it takes 1,000 micrometers to equal one millimeter and 1,000 millimeters to equal a meter, about one yard). The walls of capillaries, unlike all other segments of blood vessel, are a single cell layer thick. These cells are called endothelial cells. Collectively they are referred to as the vascular endothelium. Capillaries are not under regulatory influence by nerves or metabolites. As already mentioned, luminal changes in diameter of blood vessels are caused by the contraction/relaxation of overlying vascular smooth muscle cells. Capillaries cannot contract and relax because they lack vascular smooth muscle cells. Capillaries are the sites of most of the exchange that occurs between the flowing blood and the interstitium. A site of less significant exchange is the downstream venule, or the smallest of the veins. Blood pressure inside capillaries is lower than that in upstream arterioles and higher than that in downstream veins.

The exchange between the circulatory system and the surrounding tissue spaces and cells is critically important, and all regulatory phenomena within the circulatory system seem to be directed at protecting capillary hemodynamics. In homeostasis, there is a steady egress of fluid and nutrients from the capillary spaces into the interstitial spaces. This ensures adequate nutrition for cells and simultaneous removal of the by-products of metabolism. It also ensures the physiological balance between water and electrolytes in the various aqueous body compartments. In short, the homeostasis of microcirculatory hemodynamics ensures the existence of life. The key variables in this process are called Starling forces in honor of the British physiologist Ernest Starling (1866–1927), who first investigated and enunciated them. These variables are capillary hydrostatic pressure, capillary colloid osmotic pressure, interstitial hydrostatic pressure, and interstitial colloid osmotic pressure. I will not quantify these here, but know that in the circulatory steady state there is a harmonious balance among them. Anything that disturbs this balance more than just transiently is potentially life-threatening.

During the course of a twenty-four-hour period, there is a net efflux of water from the arteriolar end of the capillaries that is not compensated for by a corresponding influx at the venular end. This difference of two or three liters of water could be stored in the interstitial spaces and thus not get returned to the circulatory system. If this was all that happened, circulatory collapse, systemic hypotension, and eventually death would ensue. Fortunately, under normal circumstances, this does not happen because of our lymphatic circulatory system. Lymphatic capillaries that are distributed throughout the tissues and organs absorb the excess two or three liters of water and return it to the systemic circulation. If the lymphatics malfunction and/or if there is excess loss of protein (colloid) and water from the capillaries, accumulation of water in the interstitium can occur more rapidly than the lymphatics can remove it, and the steady

state will be impaired. Under these circumstances, the end result is temporary tissue edema and inflammation such as might occur with the sting of an insect.

Venules are the first postcapillary segment of blood vessels. Like capillaries they have low intraluminal pressures, but like arteries they do have vascular smooth muscle cells in their walls. Therefore venules and small veins downstream to them are under both neurogenic and metabolic control. This means that under the right circumstances a venule can contract, impede the outflow of blood from capillaries, and thus increase the volume of blood and pressure inside those capillaries. When this happens, the hemodynamics in and around the capillaries change. Such changes can again lead to increased loss of fluid and proteins from the capillaries and to inflammation and swelling of the tissues, an edemogenic response. Thus, under many conditions, venules are just as important as arterioles in determining the dynamics of fluid and solute exchange in the capillaries. Their contributions to the health and welfare of any tissue and, indeed, the whole body must not be underestimated.

Small and large veins are the counterparts to the distributing and conducting arteries of the upstream side of the capillaries. They are the compliance vessels of the circulatory system. As mentioned above, compliance has a variety of definitions. For our purposes, it can be thought of as stretchability, distensibility, and/or elasticity. Because veins are more compliant than corresponding arteries, they can accommodate larger volumes of blood without experiencing significant increases in blood pressure. Experimentally, if the volume of blood in the systemic circulatory system is measured while the heart is arrested, greater than or equal to 60 percent of the volume of blood would be found in the veins. Much less would be found in the arteries, capillaries, and heart chambers. Such experiments in animals have helped to establish that veins are more compliant than arteries. Also, segments of large and small arteries and veins can be removed from the body and studied in vitro or ex vivo to confirm the above.

Because of their compliance, veins serve as blood reservoirs or blood sources. When arterial pressure falls due to hemorrhage (a blood sink), veins can constrict and in so doing transfer large portions of their blood volume to the arterial side of the circulatory system. This helps to sustain arterial blood pressure and perfusion of the organs, and has been found repeatedly under conditions of shock and trauma and in hemorrhagic hypotension. This natural transfusion of blood can help a person survive until artificial transfusions are available. This system is one of the body's best examples of a negative feedback control. Sensory receptors located near the bifurcation of the common carotid arteries into external and internal branches, known as baroreceptors, are continuously sensing pulsatile and mean arterial blood pressure at these points. That sensory information is relayed to cardiovascular control centers in the medullary regions of the brain stem where error detectors compare it with physiological set points. Any adjustments that are needed to protect the homeostasis of blood

pressure are made via motor nerves that send activating signals to the heart, blood vessels, and other organs and tissues. For example, if blood pressure in the carotid bifurcation decreases, the baroreceptor reflex response will increase heart rate, cardiac contractility, venoconstriction, and generalized arterial vasoconstriction. Collectively these changes will restore carotid blood pressure to its normotensive level.

Lifestyle and Monitoring Cardiovascular Health

There are simple, reliable, and reproducible ways for students and practitioners to monitor the health of their cardiovascular systems and to help patients and clients do the same. Heart rate, when expressed in cycles or beats per minute, is a cardiovascular variable. Changes in heart rate can be influenced by many things (table 4.3). In the resting basal state, slower heart rates are better for one's health than higher heart rates. The explanation is simple. The more times a heart contracts each minute, the more work it does. The fewer times the heart beats per minute, the less work it does. For the student and clinician, the heart does two kinds of work: pressure-related and volume-related work. Both kinds require the expenditure of chemical energy and the consumption of molecular oxygen. Simply stated, the more oxygen and energy an individual or an organ consume per unit of time, the more work they are doing. Working an animal, a person, or an organ more will wear it out sooner than one worked less. Thus, having a slow heart rate under resting conditions is better for the heart and the person than having a fast heart rate.

TABLE 4.3

Selected activities that affect heart rate in mammals

Event	*Directional change in heart rate (increase, decrease)*
Rest and sleep	heart rate decreases
Dynamic exercise (manual labor)	heart rate increases
Emotional distress	heart rate increases
Exposure to extreme cold	heart rate decreases
Exposure to extreme heat	heart rate increases
Resting state (conditioned person)	heart rate slower than in unfit people
Treadmill exercise stress test	heart rate rises incrementally
Intentional breath holding	heart rate should decrease transiently
Straining at stool	heart rate can increase dangerously

Excluding genetic considerations, conditioned athletes are those who have the slowest resting heart rates. They are said to have exercise-induced bradycardia. Bradycardia means slow heart rate and tachycardia means fast heart rate. Some elite athletes have resting heart rates below 45 cycles per minute (cpm, or bpm for beats per minute): when Lance Armstrong won his seventh consecutive Tour de France in 2005, I was told his resting heart rate was 35 cpm. Let's imagine that you are a physically unfit student and that your basal, resting heart rate is 75 cpm. You want to become healthier, and as a practitioner you wish to encourage your patients similarly. By palpating the pressure pulse in either your carotid or radial arteries, you can determine your heart rate under basal conditions. You can do this several times, record the numbers, and calculate the average. Consider this your basal heart rate and make a permanent record of it. Now go climb the stairs or do some push-ups and other form of dynamic exercise for about ten minutes, then stop and record your heart rate again. Perhaps it rose to 125 cpm during that ten-minute bout of exercise. Repeat the same activity for several days, then progressively increase the duration of your exercise sessions over a period of several weeks from ten to twenty to forty minutes. If you do this consistently for a few months, you will discover that your basal resting heart rate has decreased, depending on the duration (minutes per session), intensity (effort exerted), and longevity (weeks or months) of your workout. In other words, the health of your cardiovascular system as assessed by your resting heart rate will have improved.

Once you have convinced yourself of the physiological reality of this exercise and have felt the mental and emotional exhilaration that accompanies it, not only will you want to continue indefinitely but you will want those you care about to have a similar experience. You should then be motivated to encourage your patients, clients, and others over whom you have some influence to do something similar. As you progress, you are becoming a healthier student, practitioner, and person.

No reasonably minded person in the twenty-first century will debate the connections between diet, exercise, and cardiovascular health and disease. The data are too firmly established to deny the beneficial effects of wise dietary choices and sustained physical activity on the heart and circulation. For the past several decades, the American public has been inundated with information about the influence of diet on cardiovascular health. Unfortunately, fad diets come and go, and in their wake millions are confused and some even die. Most physicians who write books about diets and nutrition are neither nutritionists nor dieticians. Therefore, before subscribing to their diets and buying their products you might ask yourself what qualifies this person to write on the topic.

While you are reflecting on fad diets, consider also the bottled water/sports drink fads of the past decade and a half. What is the evidence that bottled water is better for your health than regular tap water? Contact any branch of

government and ask for literature on purification standards for drinking water in your community, state, and the nation. Once you have obtained the literature, contact any manufacturer of bottled water and ask for their literature regarding the quality control standards for their products. Compare the two sets of standards and you will discover that few if any manufacturers of bottled water exceed governmental standards for the potability of drinking water. This means several things. First, commercially available bottled water is no safer for your health than is your community water supply. Second, since someone is already paying a utility service to purify and pipe water into your home, you are throwing away money by purchasing bottled water. Finally, paying for bottled water means that your ignorance of the facts has allowed someone else to exploit you. I will have more to say about sports drinks and the myths associated with their use. By becoming an educated consumer, the student also becomes an educated clinician. Teaching your patients to be similarly wise is part of your clinical obligation. Now back to diet, exercise, and cardiovascular health.

Eating a balanced diet cannot be overemphasized for one's health. Dieticians and nutritionists have taught us that excess fats, carbohydrates, and calories are not good for our health. But how many students have heard and understand the phrase *energy balance*? Energy balance refers to the intake and combustion of food or calories. It is used most commonly when talking about maintenance or loss of weight. Barring genetic errors of metabolism that we have little or no control over (despite the fact that some advances have been made in this era of molecular biology), there is a very simple rule for energy balance. If the amount of energy consumed is balanced by the amount of energy expended, then there will be no gain in body weight. In other words, no matter what your current body weight, if you consume 2,000 calories today and expend 2,000 calories today, then tomorrow you will weigh the same that you did today. Conversely, if you consume 3,000 calories today but only expend 2,000 calories, then tomorrow you will weigh more than you did today. Alternatively, if you consume 2,000 calories and expend 3,000 calories, then tomorrow you will weigh less. Energy balance obeys the same scientific laws and principles as does water balance, sodium or salt balance, or balance of any other physiological variable. Speaking physiologically, balance is analogous to maintaining one's checkbook or savings account. The key to maintaining your weight or to either losing or gaining weight is to understand and apply the physiological principles of energy balance.

Discipline is a key ingredient in maintaining a healthy cardiovascular system through one's lifestyle and diet. There is nothing wrong with having a normally healthy diet today and then enjoying cake and ice cream tonight at a birthday party. If your cake and ice cream constitute the equivalent of 500 to 1,000 extra calories, it is still not bad. The bad comes from doing nothing to counterbalance the energy intake within the next twenty-four hours or so. For the disciplined student, patient, or practitioner, several things can be done about

those extra 500 to 1,000 calories. You can (1) partially fast the following day, for example, miss breakfast or lunch, or eat a smaller evening meal; (2) engage in some dynamic activity during the next twenty-four hours that burns the equivalent calories; for example, you can run, swim, cycle, dance, or do vigorous yard work; (3) do any combination of dieting and exercising that compensates for the 500 to 1,000 calories in the cake and ice cream. The main problem we have relative to diet and cardiovascular health is sustained overindulgence with energy intake and lack of discipline in compensatory or purposeful energy expenditure. If you think you can regularly overeat and underexercise, and not eventually have heart, vascular, and other health-related problems, you don't understand life and health.

It is not enough to simply avoid eating too much junk food as a means to good cardiovascular health. You must also eat the proper kinds of foods in the appropriate daily proportions. The best way to do this is through education, good judgment, and common sense. To help the student and clinician, I encourage getting the *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2005* (new editions are published every five years). This dietary guideline is saturated with good advice about food choices as they relate not only to cardiovascular disease but also to diseases such as cancer and diabetes. Moreover, there is information in the report to help individuals of both genders and several ages select exercise programs to supplement their diets and quests for improved health.

Another piece of good advice is to ask your physician to refer you for a complete screen of blood lipids. Obtain a copy of the report and take special note of your cholesterol and triglyceride numbers. Then work and eat to either lower or maintain them. Body weight is the physiological variable that is most highly correlated with elevated levels of circulating lipids including cholesterol. The heavier you are, the higher your blood cholesterol. The higher your blood cholesterol, the more likely you are to have an unhealthy cardiovascular system and to die an early death from heart-related cardiovascular disease. Dietary fats are one of the main determinants of circulating cholesterol concentrations. Most dietary fats come in the form of triglycerides. Therefore it is important to keep both the cholesterol and the triglycerides in control.

One thing that is inversely related to circulating concentrations of cholesterol is intake of dietary fiber, particularly the insoluble variety. That is, as your daily consumption of insoluble fiber increases your blood cholesterol concentrations come down. On the average, American women consume ten to twelve grams per day of dietary fiber. American men consume proportionately more, about fifteen to seventeen grams per day, because they are larger and tend to eat more. A good target for both genders is twenty-five to thirty-five grams of dietary fiber per day; in other words, most of us should try to double our fiber intake.

To test the influence of diet on your cardiovascular health, here is a simple experiment to try. Go to your doctors and tell them you are planning to change your diet, but before doing so you would like your blood lipids measured. For about three days before the phlebotomist withdraws your blood, make an accurate record of your daily fiber intake. You can go online to find out how much fiber is in the foods you eat. If you have trouble finding information about the fiber content, check the USDA's Diet and Nutrition Service (see also the suggested reading at the end of this chapter). Dietary fiber is found mainly in plant products such as fruits and vegetables but can also be found in breads and cereals, particularly those containing wheat or whole wheat. After your blood has been withdrawn and you have measurements for cholesterol and triglycerides, modify your diet by doubling your daily intake of fiber for about thirty days. On days 28, 29, and 30, reestimate your daily fiber intake at this higher level of fiber. Then go for a second blood withdrawal, get those numbers for cholesterol and triglycerides, and compare the two sets of numbers. If you are a typical American, have been on the typical Western diet, and have done nothing else during the thirty-day period that might affect your blood lipids (for example, you have not changed your exercise routine, started smoking, or started or stopped taking oral contraceptives), then you should see an appreciable decline in your cholesterol and triglyceride numbers.

For most, this experiment can be done discretely and with little or no expense. If doctors understand their patients' concerns about improving health, they can prescribe dual blood screens (that is, before and after samples) and raise little concern for insurance companies because such repetitions are common, everyday occurrences. The insurance company does not need to know that yours is an experimental investigation. That is between you and your physician.

Reactive Hyperemia and Blood Flow to Organs and Tissues

Blood flow is one of several important cardiovascular hemodynamic variables. It is routinely measured in experimental cardiovascular laboratories, during drug development in the pharmaceutical industry, and clinically. Coupled with corresponding measurements of blood pressure, blood viscosity or hematocrit, and calculations of vascular resistance, much can be learned about the circulatory health of an organ by studying its blood flow.

Reactive hyperemia is one measure of an organ's blood flow. It is an indicator of the intrinsic ability of an organ or tissue to locally regulate its own blood supply. This is in contrast to and distinctly different from the remote control of blood flow by the central nervous system. Several forms of local blood flow control exist in humans and other mammals. They are (1) reactive hyperemia, (2) active hyperemia, or exercise/functional hyperemia, and (3) pressure-flow autoregulation. Most organs, including the heart, kidneys, skeletal muscle, gut, and skin, display one or more of these flow-controlling mechanisms. The

mechanisms of these local blood flow phenomena have been under experimental investigation by physiologists for decades.

Reactive hyperemia is achieved by the temporary occlusion of the arterial blood flow to the tissue followed by release of the occlusion. The excess flow that is measured upon release of the occlusion defines the response. *Hyperemia* means excess blood flow. *Reactive* means in response to a preceding action, that is, occlusion of blood flow. As an example, picture the forehead of someone who has rested his head for a few minutes on the palm of his supported hand (elbow resting on a knee, arm supported by a chair). Imagine the blood vessels in the skin of the forehead. During the period of rest, those blood vessels are compressed between two relatively hard surfaces: the skull and the bones of the hand or clenched fist. During the period of vascular compression, blood flow to tissues downstream to the point of occlusion is impaired or completely stopped. After several minutes of occlusion and when the person has removed his head from the resting position, the color of the skin over the point of occlusion is a brighter hue of red than it was before the occlusion. This reveals reactive hyperemia in the blood vessels of the skin (see figure 4.3). The mechanisms that cause reactive hyperemia are beyond the scope of this book but have to do with the balance between oxygen supply to the skin and the removal of by-products of metabolism before, during, and after the period of occlusion of blood flow.

Thus a measure of the health of one's cardiovascular system is whether or not an organ or tissue displays an appropriate reactive hyperemic response.

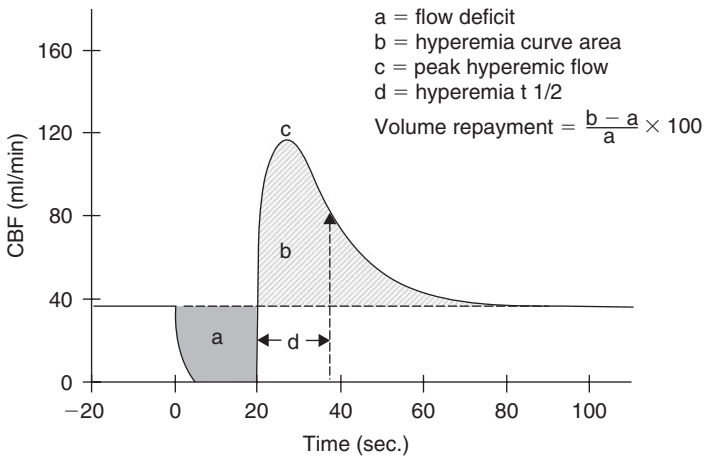

FIGURE 4.3 The blood flow phenomenon of coronary reactive hyperemia. During the period 0–20 seconds (a), blood flow is arrested (falls to zero). During the period 20–80 seconds (b), the occlusion is released and blood flow is restored (reactive hyperemia). The letters a, b, c, and d are indications of the different ways this phenomenon can be quantified.

Merrill, unpublished data

A clinically more relevant way of assessing reactive hyperemia is called the Alan or modified Alan test of collateral perfusion. This is a test of reactive hyperemia of the cutaneous and subcutaneous tissues of the hand. Two people and a clock with a second hand are needed. The first person is the experimental subject. The second person is a technician or assistant. The experimental subject should be seated or lying comfortably in a supine (on the back) position with his arms extended, palms facing up and resting on a solid immobile surface. Both arms should be easily accessible to the technician.

After five or ten minutes in this resting position and ensuring that no physical exertion occurred immediately before the period of rest, both persons should observe the coloration of the palms of the experimental subject's hands. Both should also identify the locations of the radial and ulnar arteries. When the arm is in the supine resting position, palm facing up, the radial artery is located laterally projecting toward the thumb and near the outside edge of the wrist, and the ulnar artery is located medially projecting towards the ring and little fingers and near the inside edge of the wrist. Using curved and apposed index, middle, and ring fingers, the technician should practice locating and palpating the arterial pulses in both arteries. He must be able to correctly identify the location of each artery. During the actual test, the technician will place his thumbs directly over the radial and ulnar arteries of one of the subject's arms. On the appropriate signal, the subject will tightly clench the fist of that arm for a predetermined period of time, for example, ten, fifteen, twenty, or thirty seconds. Simultaneously, the technician will apply thumb pressure to both arteries in an attempt to occlude blood flow through them. If the pressure is properly applied and flow is occluded, the subject might feel modest tingling or numbing of the hand during the period of occlusion. At the end of the predetermined time interval, the technician releases both occlusions and the subject opens his clenched fist. Coloration of the palm and fingers is observed. The yellow/white patches should turn a uniform red hue within about five seconds upon release of the occlusion and clenched fist. The change in color from yellowish/white to red indicates refilling of blood vessels with oxygenated blood, that is, reactive hyperemia and reperfusion of the tissues of the hand. If the time of reperfusion significantly exceeds five seconds, the subject should consider seeing a physician and obtaining a referral to a cardiologist.

Occlusion of radial and ulnar arteries using direct thumb pressure is only one method of disrupting blood flow to the clenched fist. Another method is occlusion of the brachiocephalic arteries in the arm between the shoulder and elbow. This occlusion is routinely performed using inflatable blood pressure cuffs by nurses and others when they monitor systemic arterial blood pressure. In that case, the nurses also use a stethoscope to listen to the sounds made by flowing blood. When pressure in the cuff is elevated above physiological arterial blood pressures, for example, to about 180–200 mmHg, this causes a mechanical

occlusion of the brachiocephalic arteries and the cessation of blood flow downstream from the point of occlusion. When pressure in the cuff falls modestly below systolic pressure, that is, our highest blood pressure value in the artery, blood flow through the now partially occluded arteries will occur during the systolic phase of each cardiac cycle. As these boluses of blood collide with the walls of the previously occluded downstream arteries, they cause the vessels to vibrate, and the vibrations can be detected as sound. Once pressure in the cuff is modestly reduced below diastolic pressure, that is, the lowest pressure in our arterial blood vessels, the previously occluded vessels are now fully opened and the sounds disappear. Thus, the method routinely used to measure a patient's arterial blood pressure in the clinic can be used to determine other aspects of the health of one's cardiovascular system, for example, reactive hyperemia and reperfusion of the cutaneous circulation of the palm of the hand and fingers.

On occasion, results of the Alan test of tissue perfusion suggest the need for more direct procedures to monitor blood pressure in one or both arteries. These direct procedures can include local cannulation of either artery and direct monitoring of arterial blood pressure. Under these conditions, aphysiologically low pressures might indicate occlusive disease such as upstream atherosclerosis. Sometimes patients with Raynaud's disease are diagnosed using procedures such as those described here.

5

Health and the Respiratory System

Breathing in the human includes both ventilation and respiration. The rib cage, diaphragm, and intercostal muscles constitute a bellows-like system in which the lungs are found. Neurogenically controlled movements of the thoracic cavity cause the expansions and contractions of the lung that respiratory physiologists call ventilation. During the inspiratory and expiratory phases of each respiratory cycle, atmospheric air moves into and out of the lungs in a rhythm that is analogous to the flow of ocean tides. In both cases, air and water flow over the same path during each cycle. Because of this analogy, respiratory physiologists, pulmonologists, and respiratory therapists call the cyclic flow of air in and out of the lungs the *tidal volume*. The number of respiratory cycles in a minute multiplied by the tidal volume is called the *minute ventilation*.

Respiration concerns the fate of the gases that are carried in and out of the lungs with each tidal volume. The primary respiratory gases are oxygen and carbon dioxide. The mechanisms by which these gases get from the lungs to the tissues, and from the tissues to the lungs, respiratory physiologists call *external respiration*. It involves mainly the exchange and transport of gases between lungs and blood and between blood and tissues. Biochemists call the mechanisms by which oxygen gets used by the cells *tissue* or *internal respiration*. This involves use of oxygen by mitochondria and the processes of electron transfer. Therefore, to fully understand the human respiratory system, the student must grasp the physiological concepts of ventilation, gaseous exchange and transport, and uptake and release of gases by cells and subcellular organelles such as the mitochondria.

The human respiratory system serves both respiratory and nonrespiratory purposes. Physiologically, the respiratory system delivers fresh ambient air to the blood and releases gaseous metabolic by-products to the atmosphere. The physiologically important ingredient of fresh ambient air is oxygen. The significant gaseous component released to the atmosphere is carbon dioxide. Nitrogen,

water vapor, and inert gases are nonphysiological participants. Nonrespiratory functions of the lungs include the delivery of odorants to the olfactory epithelium. This is important to animals that rely on sniffing to detect their environment without the danger of bringing noxious volatile agents deep into the lungs. The lungs also warm, filter, and moisturize the air. This prevents airborne pathogens from entering alveolar gas spaces, it minimizes dessication of the mucosal linings of the respiratory channels, and it protects exchange of gases by ensuring the homeostasis of temperature in the alveoli. The lungs also play an important role in immunology. Pulmonary macrophages prevent ingested and inhaled foreign substances from getting into the alveoli.

The respiratory system is subject to feedback control, gradients for pressure and airflow, and other physiological regulatory phenomena mentioned in chapter 1. For example, there would be no exchange of blood gases between the tissues, blood, and alveoli without the appropriate partial pressure gradients for gases. Atmospheric oxygen is delivered to cellular mitochondria down a pressure gradient that decreases in each location the gases move through: the partial pressure in the atmosphere (PO_2) is greater than that in the alveoli, which is greater than in pulmonary capillaries, then interstitial spaces, then cytosol, then mitochondrion. For carbon dioxide, produced as a by-product of cellular metabolism, the gradient is reversed. The mitochondrial PCO_2 is greater than the cytosolic PCO_2, which is greater than, in turn, the interstitial PCO_2, venous blood PCO_2, alveolar PCO_2, and finally atmospheric PCO_2.

The human respiratory system is subject to central control mechanisms and negative feedback in the same way that the cardiovascular system is. For example, during inspiration as the alveoli fill with air, causing both them and the thoracic cavity to expand, mechanoreceptors in the walls of the airways and in the chest wall detect the expansion and send sensory signals to the brain stem to terminate inspiration. These signals and others from higher regions of the central nervous system get integrated in the respiratory control centers of the brain stem medulla. Collectively this stops the firing of inspiratory neurons, brings airflow and lung expansion to an end, and allows expiration to begin passively. This is one more example of a well-ordered, cyclic, negative feedback control system. Like all other major mammalian organ systems, respiratory homeostasis is ensured by a system of sensory receptors, afferent—sensory nerve—central comparators and integrators and motor—efferent nerve—and corresponding activators. This is true whether the respiratory system is regulating gas exchange or participating in the control of blood pH.

Components of the Respiratory System

The human respiratory system consists of two lungs (trilobed right lung and bilobed left lung); a conducting system of airways that includes the trachea, left

and right bronchi, and nonrespiratory bronchioles; and an air-exchange system composed of respiratory bronchioles, alveolar sacs, and individual alveoli. The conducting airways are passive structures and simply convey a volume of air to and from the lungs. The bronchioles of the conducting system are active structures. In addition to other cellular and molecular elements, the walls of the nonrespiratory bronchioles are composed of bronchiolar smooth muscle cells that are innervated by the autonomic nervous system. Parasympathetic efferent neurons release acetylcholine onto bronchiolar smooth muscle and cause it to contract (the bronchospasm caused by nicotine in cigarette smoke involves a vagal, efferent limb that innervates the bronchioles). The sympathetic nerves release norepinephrine, and this neurotransmitter has the opposite effect on bronchioles. They are relaxed in response to sympathetic stimulation. Thus, the inside diameter of the respiratory airways is actively regulated. They can relax by a process called bronchodilation and contract by bronchoconstrictor mechanisms. Downstream to the respiratory bronchioles are the alveoli (alveolus is the singular), the main site for exchange of alveolar gases with pulmonary blood. Table 5.1 outlines general details of the main components of the respiratory system in humans.

The trachea divides into a left and right branch as do all the bronchi and bronchioles. Each branch, also called a bifurcation, is considered a new generation. There are twenty-five to thirty generations of bronchi leading ultimately to about ten million bronchioles. These give rise to hundreds of millions of alveoli, maximizing the surface available for exchange of gases. If all the alveoli of the adult human lungs could be opened and spread out on a flat surface, they would cover an area nearly the size of a tennis court. This is how much surface area is available in the lungs for air and blood to come in contact with one another. Of course there are conditions and disease states that impair both the size and physiological effectiveness of this surface area. Such diseases can affect either the respiratory component of the area, the vascular component, or both.

Other physical processes illustrate the effectiveness of a large surface area. Think of drying towels on a clothes line in the wind. If the towels are fully opened and attached at only one edge, they will dry faster than if they are not fully opened, draped over the line so that two halves come in contact with each other, or any other combination. This is because the fully opened towel has a greater surface area to which the wind will have access. A large block of ice will melt more slowly than the same volume/mass of ice broken into smaller pieces, which have more surface area exposed to warmer temperatures (useful knowledge when buying ice for an ice chest). The importance of surface area to daily living and to respiratory physiology can hardly be overstated.

Consider the subject of hypoxia, or lack of oxygen, in either the air we breathe or the blood circulating to the tissues and cells. In most mammalian organ systems, lack of oxygen in the blood causes vasodilation of small arteries and arterioles. This allows blood flow to increase, provided the arterial perfusion

TABLE 5.1

Structures of the conducting and respiratory zones of the mammalian respiratory system

Structure	*Inside diameter (mm)*	*Cilia present (y/n)*	*Cartilage present (y/n)*	*Smooth muscle present (y/n)*
Larynx	35–45	yes	yes	no
Trachea	20–25	yes	yes (C-shaped)	yes (modest amount)
Primary bronchi	12–16	yes	yes (full rings)	yes
Secondary bronchi	10–12	yes	yes (plates)	yes
Tertiary bronchi	8–10	yes	yes (plates)	yes
Smaller bronchi	1–8	yes	yes (modest, plates)	yes
Bronchioles	0.5–1.0	yes (reduced amount)	no	yes (plentiful)
Terminal bronchioles	<0.5	yes (reduced amount)	no	yes (plentiful)
Respiratory bronchioles	<0.5	yes (minimal)	no	yes (modest amount)
Alveolar sacs	<0.3	no	no	no

Data adapted from Germann and Stanfield, *Principles of Human Physiology*, 2nd ed. 2005, Pearson Benjamin Cummings.

pressure is maintained or even elevated. The increase in blood flow can compensate to some degree for the lack of oxygen. Oddly, the small arteries and arterioles of the lungs vasoconstrict in response to hypoxia. When this happens, blood flow is diverted away from those alveoli that have a low oxygen content. Hence, there is either reduced or no gas exchange within such alveoli. This results in the same net effect, as far as gas exchange is concerned, as reducing the alveolar surface area. Such a response was designed to match pulmonary blood flow with alveolar ventilation. Under homeostatic conditions, the volume of airflow to each alveolus is almost perfectly matched by the volume of blood flow to the same alveolus. That is, the ratio of ventilation to perfusion is about one. This is true for individual alveoli and for the entire lung. Matching ventilation and perfusion ensures that blood is delivered only to those alveoli that are filled with air. Alternatively, it means that ventilation occurs only in those alveoli that are perfused with blood.

Within the walls of the alveoli, there are three main kinds of cells. They are designated type I and type II respiratory cells and macrophages. Type I respiratory cells are flat epithelial cells. They overlay a basement membrane that is made of glycoproteins. Adjacent to the basement membrane is the pulmonary capillary whose wall is comprised of a single layer of endothelial cells. The alveolar epithelium, the basement membrane, and the pulmonary capillary endothelium constitute a respiratory membrane, or barrier, of approximately 0.2 microns thick through which gas molecules like oxygen, nitrogen, carbon dioxide, and water vapor must pass. Changes in its thickness can markedly influence exchange of gases. At high altitude, the thickness of the barrier increases. This increases the distance that either oxygen or carbon dioxide have to travel to get to the pulmonary blood or alveolar air. The end result is body hypoxia and hypercapnia: low oxygen and simultaneously high carbon dioxide. Both conditions change gradients for gas exchange, signal danger, and are potentially lethal alone or in combination.

The third type of cell found in the alveoli is the macrophage. Macrophages are white blood cells produced by bone marrow, released into the systemic circulatory system, then extracted by the tissues of organs after a few days. They are an important part of the body's immune system. Macrophages line the walls of the alveoli but are able to move about in order to detect, ingest, and destroy, by phagocytosis, foreign substances that get into the lungs.

Pulmonary Ventilation, or Lung Inflation and Deflation

Lungs inflate and deflate much like balloons. Balloons have a neck and a body. Blowing forcefully into the neck of a balloon creates a downhill pressure gradient that is high in the lungs and low inside the body of the balloon. Because of the downhill pressure gradient, air is able to flow from the lungs into the balloon. The key to this process is the interrelationship among airflow, pressure, and volume. The thoracic cavity and the lungs have been designed to cyclically inflate and deflate. The process is called the respiratory cycle. The cycle consists of three phases: an inspiratory phase during which air moves from the atmosphere into the respiratory system, an expiratory phase when air moves from the respiratory system into the atmosphere, and an end-expiratory phase during which there is no net movement of air. Respiratory physiologists refer to the entire process of inflating and deflating the lungs as pulmonary ventilation or simply ventilation. To correctly understand pulmonary ventilation, the student must have more information about the structure of the thoracic cavity and the lungs, as well as a minimal appreciation for some of the simpler respiratory gas laws.

The lungs and trachea are contained within the thoracic cavity, or thorax. At the top this cavity is composed of the base of the neck and related structures. The lateral, anterior, and posterior portions consist of ribs, the sternum, and the vertebrae, respectively. At the bottom margin is the diaphragm, a skeletal/visceral

muscular structure. The ribs attach at the sternum and vertebrae and are connected to one another by intercostal skeletal muscles. For any particular rib, its point of attachment at the sternum is slightly below the corresponding point of attachment at the vertebra. The apex or superior aspect of both lungs is narrower than the corresponding base or inferior margin. This is to accommodate the heart and large blood vessels. The region of the thorax where the heart, great vessels, and apices of the lungs come in contact is called the mediastinum. Motor nerves innervate the diaphragm, the intercostal muscles, and the conducting/respiratory airways. Sensory nerves are found in the walls of the conducting airways and are located in other intrathoracic structures as well. As mentioned above, this means that the respiratory system is subject to neurogenic control.

The outermost layer of cells on the surface of the lungs is called the visceral pleura and the innermost layer of epithelial cells on the inside surface of the thorax is called the parietal pleura. Both layers of pleura are composed of flat epithelial cells that have physiological functions like all other epithelial cells, including secretion of aqueous solutions. The potential space between the two pleural membranes is called the intrapleural space. Under physiological conditions, these two membranes are not quite in contact with each other over most of their surface areas. This is because both membranes are covered by a thin film of aqueous material. This fluid-like substance lubricates the two membranes, reduces friction between them, and has other physical and chemical properties that prevent the two membranes from pulling apart. This contact is maintained throughout the respiratory cycle. To help explain ventilation, respiratory physiologists refer to an intrapleural pressure and to intrapleural fluid.

There are many gas laws, both chemical and physical, that apply to the respiratory system. The serious student and practitioner should explore all of these, including Boyle's law, Charles law, Dalton's law, Gay-Lussac's law, and Henry's law. A textbook of respiratory physiology will explain some of the more important gas equations. For our discussion, Boyle's law is most important. Boyle's law relates pressures and volumes: in a closed system such as the lungs and thoracic cavity of the human, whenever volume increases there must be a corresponding decrease in pressure. Conversely, if pressure increases there will be a corresponding decrease in volume. In other words, according to this law, the product of pressure and volume remains constant. Consider your own respiratory system and imagine that you are at the end-expiratory phase of the current respiratory cycle. At that moment in time, there is no net movement of air into or out of the lungs and the product of pressure and volume equals some number, but neither pressure nor volume are changing. Now the inspiratory phase of the next respiratory cycle begins. As your thoracic cavity and lungs expand and increase in volume, Boyle's law indicates that pressure inside the lungs called intrapulmonary pressure must decrease. As the intrapulmonary pressure decreases, it

becomes modestly lower than atmospheric pressure. Atmospheric pressure means barometric pressure. The decrease in intrapulmonary pressure establishes a downhill pressure gradient from atmosphere to lungs. As a result of this gradient, due in part to expansion of the lungs during inspiration, fresh ambient air flows into the lungs and they are filled by the end of inspiration.

Another pressure gradient that is important to ventilation of the lungs is that between the intra-alveolar (intrapulmonary) and intrapleural spaces. Again imagine yourself to be at the end-expiratory phase of a normal respiratory cycle. At that point in time, there is no net airflow and intra-alveolar pressure is said to be zero. This means that ambient and intrapulmonary pressures are the same. Intrapleural pressure under physiological conditions is always subatmospheric, which means it is less than ambient pressure. In textbooks of respiratory physiology and in the current scientific and medical literature, it is not unusual to see intrapleural pressure referred to as a negative pressure. The word negative and/or the minus sign (−), in this case, simply means that intrapleural pressure is lower or less than atmospheric pressure. In the mammalian respiratory system, there is really no such thing as a negative pressure, and use of the word *subatmospheric* in place of the word *negative* is preferred. A subatmospheric intrapleural pressure ensures that the alveoli specifically and lungs in general are always partially inflated and inflatable. Should the thorax ever be punctured, the intrapleural and ambient pressures would become equal; they would equilibrate, the alveoli and lungs would collapse, and the patient's respiratory condition would be compromised. This condition is known as pneumothorax, and it can be caused by events such as a knife or gunshot wound to the chest or, in the case of an automobile accident, the penetration of a steering column. A first responder, such as an emergency medical technician at the scene, should locate the site of the wound and close it to the atmosphere. Should you find a victim conscious, able to speak, but having difficulty breathing, you must always consider pneumothorax.

Chronic obstructive pulmonary diseases (COPD) cause difficulty in ventilating the lungs. These conditions occur more commonly in the elderly, especially those who abused their respiratory systems with years or decades of tobacco use. These diseases can also occur in the young for no cause of their own, for example, small children who are unable to escape the pathological effects of secondary smoke. COPDs include asthma, chronic bronchitis, and emphysema. The first two conditions increase resistance to airflow and thus make inspiration difficult. The third causes eventual breakdown of alveolar walls. This results in difficulty deflating the lungs. All three conditions influence the ratio of air flow to blood flow in the alveoli and usually require medical attention.

Pulmonary function tests are used to diagnose respiratory diseases. These tests frequently involve analyzing one's ability to eliminate large volumes of air in short periods of time. The patient is brought to the clinic or laboratory. Using

the equipment and principles of spirometry, the technician determines the patient's maximal lung capacity, that is, how many liters of air he or she can inhale after eliminating as much air as possible during a prior prolonged and forced expiration. Once the maximal lung capacity is determined, the technician asks the patient to repeat the exercise. This time the technician determines how rapidly the patient can eliminate air and the exact volume that is expired. Both the volume of air released, measured in liters, and the time it takes to release it, measured in seconds, are evaluated. These numbers are compared with the population of healthy subjects of the same age, weight, and gender. One pulmonary function test of this kind is called the forced expiratory volume test or FEV. It measures how much air subjects can forcefully expel from their lungs. A variation of the FEV is the volume of air one can forcefully expel in the first one second of the test, also called the FEV_1. Most healthy individuals are able to eliminate about 80 percent of the maximal volume in one second. If your FEV_1 is significantly less than this, there is reason to be concerned.

Respiration, Exchange, and Transport of Blood Gases

In the above few pages, we have reviewed pulmonary ventilation or the mechanical movements of the chest wall and the corresponding inflation and deflation of the lungs. The main purpose of those activities is to cause the exchange of respiratory gases between the atmosphere, flowing blood, and the tissues and cells of the body. We will now turn our attention to the topic of external respiration. As already mentioned, there are several ways that scientists describe respiration. The student or clinician can speak of internal (mitochondrial, cellular) and external (exchange and transport of respiratory gases). Here we will focus only on external respiration. The physiologically important respiratory gases are oxygen and carbon dioxide. These are the gases that are most important to our health. Nitrogen, however, which composes about 79 percent of the ambient air, is also important in determining barometric pressures in the alveoli, and thus in determining alveolar inflation/deflation. It does not participate, however, in the important metabolic, regulatory, and respiratory processes in the body that oxygen and carbon dioxide do. Thus it is of little concern to us here. For the current discussion, respiration can be thought of as the exchange of oxygen and carbon dioxide between tissues, circulating blood, and the alveoli. That definition can be expanded to include how these gases move from one place to the other.

Let's begin our discussion at the alveoli. Imagine two events taking place simultaneously. The first is the inspiration of an amount of fresh ambient air. This amount, taken in a single breath, is called the tidal volume. The second event is the arrival at the pulmonary capillaries of a volume of blood just returning from the tissues where it has given up oxygen and gained carbon dioxide.

The volume of air that a person moves in and out of both lungs each minute is called the minute ventilation, V_E. It is determined by calculating the product of respiratory rate, measured in cycles per minute, and tidal volume, expressed in liters. An average V_E = 12 cycles per minute × 0.5 liters = 6.0 liters/min. The volume of fresh ambient air that is delivered to the alveoli and respiratory bronchioles each minute is called minute alveolar ventilation and is designated V_A. It is different from minute ventilation. The difference in the two volumes is called the dead space volume. Dead space volume is any and all portions of the respiratory system where air exists but does not get exchanged with pulmonary blood during the respiratory cycle. Specifically it includes all tubular structures of the respiratory system upstream to the respiratory bronchioles as well as the larynx, pharynx, and oral and nasal passages (the conducting zone). In each tidal volume, approximately 150 ml of air fills the dead space and therefore is not available for exchange with the blood. Of course, this volume depends on the gender and size of the person as well as on the state of their health. In order to estimate the alveolar ventilation, minute ventilation must be corrected by subtracting the volume of dead space air (or dead space ventilation). In this example, V_A = (500 ml – 150 ml) × 12 cpm, or 4.2 liters/min.

Each inspiration of new air gets divided into alveolar air and dead space air. Only the alveolar air is available for exchange with the arriving volume of pulmonary capillary blood. This is because the pulmonary capillaries lie adjacent to the alveoli and respiratory bronchioles and not to the bronchi, trachea, and nasal passageways. As the two volumes, blood and air, come in close proximity to each other, separated only by the respiratory membrane, they simultaneously pick up oxygen and release carbon dioxide. The gases get from one compartment to the other, that is, from the alveolar space to the pulmonary capillary lumen, by the principles of diffusion. Moreover, they have only a few milliseconds to make the transfer.

Key to the process of transport and exchange of gases are the red blood cells and the hemoglobin molecules they carry. Hemoglobin arriving at the alveoli from the tissues carries both carbon dioxide and hydrogen ions. Both are potential forms of acid. Acidity of blood is expressed by pH (an index of hydrogen ion activity, or hydrogen ion content, H^+). Venous blood is more acidic (for example, about 7.32 pH units) than arterial blood (about 7.40 pH units). Acidity is represented on a spectrum of 1.0 (most acidic, for example, stomach acid after ingestion of a meal containing a high content of protein) to 14 (least acidic or most basic). A neutral pH is 7.0. In human physiology, a neutral blood pH is considered to be in the range of 7.38 to 7.42 or thereabouts. Hemoglobin molecules that are carrying these potential acids are called reduced hemoglobin. They can be abbreviated HbH to designate hemoglobin bound to hydrogen ions and $HbCO_2$ to designate hemoglobin bound to carbon dioxide. In the few milliseconds that it takes the hemoglobin molecule to pass from the arterial to the venous ends of

the pulmonary capillary, chemical reactions take place inside the red blood cells that convert hydrogen ions to carbon dioxide and release it to the alveolar space. Now freed of acid, these same hemoglobin molecules are available to bind to new oxygen. Hemoglobin's molecular structure is modestly changed as it binds oxygen so it can accommodate more oxygen. Hemoglobin that is bound to oxygen is called oxyhemoglobin and is abbreviated HbO_2.

Pulmonary capillary blood that milliseconds earlier arrived at the alveoli now leaves the pulmonary capillary en route, via the left atrium and ventricle, to the tissues that moments ago released their carbon dioxide and hydrogen ions to the same or similar hemoglobin molecules. For moving products from one location to another, the hemoglobin gas-transport system is arguably the most efficient known. Can you imagine how the cost of living could be reduced if every bus, car, plane, ship, train, and truck that departed a location full of goods reached its destination, unloaded, and then refilled entirely before returning to the same location instead of returning empty or only partially filled? For years, I have watched campus buses traveling in one direction filled with students that return in the opposite direction carrying one or two students. I have traveled across the country in large airplanes when I was one of only a handful of passengers on the plane. Semi trucks depart full of cargo and carry empty trailers on return. In the case of physiology, a hemoglobin molecule arrives at the tissues and deposits its load of oxygen (see table 5.2). It then retrieves an equally important load of carbon dioxide and returns to the lungs to start over. What a marvelous picture of efficiency.

TABLE 5.2

Exchange and transport of respiratory gases (oxygen, carbon dioxide) between lungs, blood, and tissues

Site	*Oxygen*	*Cabon dioxide*	*Mode of transportation*
Alveolar pulmonary capillaries			
arterial end	low	high	bound to proteins (Hb), dissolved in solution,
venous end	high	low	bound to proteins (Hb), dissolved in solution,
Peripheral tissue capillaries			
arterial end	high	low	bound to proteins (Hb), dissolved in solution,
venous end	low	high	bound to proteins (Hb), dissolved in solution,

This respiratory system of exchange and transport is effective for healthy individuals living at or near sea level. However, the scene can change lethally in extreme conditions or the unhealthy person. Consider the novice adventure traveler. There is a lucrative business for those willing to fill the public's desire for thrill. For $75,000 to $100,000, clients can pay travel adventure companies to take them to the summit of Mount Everest, the world's highest peak at 29,028 feet. At those altitudes, barometric pressure is only one-third that at sea level. The partial pressure of oxygen that drives diffusion of the gas from one space to another inside the body is equally low relative to sea level. The partial pressure of oxygen in the alveoli at such altitudes might be below 35 mmHg, compared to about 100 mmHg in the alveoli at sea level. Under such extreme conditions, the exchange and transport of oxygen are markedly impaired. Persons often become hypoxic and disconnected, and their abilities to think and perform are seriously impaired. Between 26,000 feet and the summit is where most deaths on Everest occur; oddly, the majority of those deaths take place during the descent from the summit. Add to hypoxia the effects of malnourishment, dehydration, sleep deprivation, fatigue, hypothermia, and severe weather, and survival becomes very difficult, as novice to experienced climbers discovered on Everest in May 1996. Travel adventure expeditions to Everest, one from the United States led by Scott Fischer and another from New Zealand led by Rob Hall, were trapped by an unexpected blizzard in the death zone above 26,000 feet. Both Fischer and Hall were expert mountaineers and had summited Everest before. Nine people lost their lives in the May blizzard while seeking adventure on Mount Everest. The deaths, including those of Fischer and Hall, were all preventable and were caused, in large part, by poor judgment and the pathophysiology of oxygen and carbon dioxide.

One of the main variables that determines the exchange and transport of oxygen is the partial pressure gradient of the gas between two compartments. Under normal circumstances, and at sea level, the gradient for oxygen between alveolar air and pulmonary capillary blood is 60 mmHg or more. In the death zone on Mount Everest, and for those not using supplemental oxygen, this gradient can be reduced to less than 10 mmHg. John West and colleagues collected these data during their 1981 American Medical Research Expedition to Everest.

Central and Peripheral Regulation of the Respiratory System

As early as the second century A.D., investigators of experimental physiology suspected that control of respiration resided, at least in part, above the midcervical spinal cord. Galen of Pergamon (circa 130–210 A.D.) studied medicine in Greece, Asia Minor, and Alexandria. When he returned to Pergamon, he became chief physician and surgeon to the gladiators. Later, in Rome, he was appointed court physician to Marcus Aurelius. During Galen's tenure caring for traumatized gladiators, he observed that those wounded in the high cervical spinal cord

were not breathing when brought to him. Those receiving battle injuries in the lower cervical cord were breathing but were paralyzed in the arms and legs. He later confirmed these observations in experimental animals and reasonably concluded that the brain sends messages via the midcervical spinal cord to respiratory structures including the phrenic nerve, diaphragm, and intercostal muscles. However, Thomas Lumsden and others working more than eighteen centuries later conducted definitive physiological experiments that revealed that, in fact, control of respiration resides in the brain stem, or above the spinal cord. Galen was so revered and his theories so unchallenged that advances in respiratory medicine were stalled for nearly eighteen centuries.

Regulation of the respiratory system is under the kind of negative feedback control described in chapter 1. It includes vascular, respiratory, and neurogenic components. In these reflex arcs, the sensors are located in the systemic circulatory system (peripheral sensors) and in the brain stem (central sensors). Collectively these sensors are called chemoreceptors, and they are sensitive to changes in the partial pressures of oxygen and carbon dioxide as well as to changes in the concentrations of hydrogen ions. Sensors located in the systemic circulatory system are called peripheral chemoreceptors, and those located in the central nervous system are called central chemoreceptors. When activated, their outputs cause changes in the depth and frequency of respiration as well as changes in heart rate, blood pressure, and other cardiovascular variables. The main purpose of the respiratory control system is to maintain the partial pressures of carbon dioxide and oxygen. The system is also intimately involved in maintaining acid/base homeostasis, or pH.

The peripheral chemoreceptors are located primarily at the bifurcation of the left and right common carotid arteries where these arteries form internal and external branches. Peripheral chemoreceptors are also found near the ventral surface of the aortic arch and, depending on the mammalian species, near the bases of other large conducting arteries. Collectively they are called carotid bodies and aortic bodies. The ascending limbs of their sensory nerves project into the brain stem and other structures of the central nervous system. The central chemoreceptors are found in the pons and medulla of the brain stem, are located bilaterally, and can be divided into a dorsal respiratory group, a ventral respiratory group, and a pneumotaxic center (see table 5.3). They receive sensory input from the peripheral chemoreceptors, from respiratory structures such as the conducting airways, and from the blood perfusing them. Output from the respiratory control centers flows to the abdominal muscles, intercostal muscles, and diaphragm, as well as to the heart, vasculature, and other organs.

The dorsal respiratory group is located dorsally (toward the back) in the medulla and extends over much of the length of the medulla. This respiratory control center controls inspiration and so plays the most important role in respiration. This region of respiratory neurons displays a ramp-like discharge of

TABLE 5.3

Central respiratory control centers and their functions

Name of center	*Location*	*Function*
Dorsal respiratory group (DRG)	bilateral dorsal medulla (brainstem near nucleus tractus solitarius)	controls inspiration
Ventral respiratory group (VRG)	bilateral medulla (brainstem ventrolateral to DRG)	active during expiration
Pneumotaxic center	rostral pons (top of brainstem)	stops inspiration
Apneustic center	caudal pons (bottom of brainstem)	historic significance

action potentials that steadily increases for a couple seconds and then is terminated for about three seconds. This causes progressive inflation of the lungs and allows adequate time for the lungs and chest wall to recoil during expiration before the next ramp begins. The basal rhythmicity of inspiration is intrinsic to the neurons of the dorsal respiratory group just as basal heart rate is intrinsic to the SA node. These neurons are often referred to as the central pattern generator. The underlying cause of this intrinsic rhythm, like so many other rhythms of physiology, is unknown.

The ventral respiratory group is anterior (toward the front) and lateral (toward the side) to the corresponding dorsal respiratory group. There is little or no evidence of involvement of these respiratory neurons in quiet passive breathing, when they are mostly inactive. So we can conclude that under resting conditions, the dorsal, and not ventral, respiratory group controls respiration. However, under more dynamic conditions such as strenuous exercise, neurons of the ventral respiratory group contribute to both inspiration and expiration. During conditions such as extreme exertion, they are more involved in expiration and seem to control the powerful respiratory movements of the abdominal muscles.

The pneumotaxic center is located dorsally in the pons region of the brain stem. Its main function is to limit the duration of inspiration. When strong signals descend from this center to the dorsal respiratory group, the duration of inspiration might last only half a second. When the signals are weak, inspiration can last as long as five seconds. The pneumotaxic center plays a variety of roles in autonomic function but is not essential to normal respiration or eupnic breathing.

Not all respiratory physiologists agree on the existence of a fourth set of respiratory neurons, the apneustic center. When this term appears in modern textbooks

of physiology and medicine, it is usually to pay tribute to early investigators and their contributions. This center is more historical than physiological.

Peripheral regulation of respiration begins with changes in blood gases and blood pH. The critical changes are those that occur at the sites of the aortic and carotid bodies. The specific changes that stimulate an increase in respiration are (1) a decrease in the partial pressure of oxygen in arterial blood (PaO_2), (2) an increase in the partial pressure of carbon dioxide in the arterial blood ($PaCO_2$), and (3) a decrease in the pH of the arterial blood, that is, an increase in hydrogen ion concentration. Under physiological conditions, the PaO_2 of arterial blood perfusing the aortic and carotid bodies is about 100 mmHg. Thus, these chemically sensitive neurons are conditioned to this level of dissolved oxygen in the blood. They generate action potentials that are carried centrally to act on the respiratory control centers. These in turn cause the respiratory system to operate at a frequency of about twelve to fifteen cycles per minute and at a tidal volume of about 400 to 600 milliliters. This establishes basal minute and alveolar ventilations. These are closely matched to rates of blood flow in the pulmonary and systemic circulatory systems. Such basal physiological rates of respiration sustain life.

Imagine yourself as a mountaineer. Your PaO_2 has been reduced to 50 to 60 mmHg as you near the summit of a 20,000 foot peak without breathing supplemental oxygen. Your aortic and carotid bodies will detect this aphysiological change and will fire action potentials with increased frequency to the medullary central pattern generator. The output of respiratory neurons to the muscles of inspiration will increase, and both your tidal volume and respiratory frequency will increase. This will stimulate ventilation of the lungs and will deliver more oxygen to the alveoli and pulmonary capillary blood. Subsequently the arterial PaO_2 will increase, compensating for the reduced PaO_2 you had experienced. Soon a new steady state for PaO_2 will be established, and PaO_2 will be within a range that is somewhere between 95–100 and 50–60 mmHg. The new steady state for oxygen delivery will sustain you for the several hours you remain at 20,000 feet. This is one more example of how negative feedback systems work. The initial stimulus was an aphysiologically low partial pressure of oxygen at the chemoreceptors. The response was a compensatory increase in respiration that delivered more oxygen to the blood. The final response was a return of the partial pressure of oxygen to more physiological levels.

Lungs and Balance of Acids and Bases

Under physiological conditions, the partial pressures of oxygen and carbon dioxide in human arterial blood are, respectively, 95 ± 5 mmHg and 40 ± 2 mmHg. The pH of arterial blood under these conditions is about 7.40 ± 0.02 units. Corresponding values for venous blood under physiological conditions are 40 ± 2 mmHg, 45 ± 2 mmHg, and about 7.35 ± 0.02 units. All three physiological

variables are under homeostatic control and must remain relatively constant to sustain life. There are physiological mechanisms in place to compensate for simultaneous changes in one or more of these variables. Some of these mechanisms are performed by the blood, such as the presence of proteins that can buffer changes in blood pH, others are found in the interstitial fluids, such as the hydration of carbon dioxide to produce a weaker acid, and still others reside in the respiratory and renal systems.

Acidosis refers to any condition that results in an excess of acid in any compartment of the body, including intracellular, intravascular, and interstitial spaces. *Alkalosis* refers to any condition that results in an excess of base or the opposite of acid. If acidosis or alkalosis is caused by a problem in the respiratory system, the condition is called respiratory acidosis or alkalosis. If disruption of kidney function causes the problem, the ailment is called renal acidosis or alkalosis. From these simple definitions, the student can tell that the lungs and kidneys are the organs charged with resolving acid/base problems that might compromise one's health. The lungs are a rapidly responding system, and corrective changes occur within seconds or minutes. The kidneys respond more slowly, and corrective changes take hours to days. From this perspective, it is prudent to view the two systems as complimentary to one another. However, only the acid/base responses of the lungs will be discussed here.

Table 5.4 lists some examples of the pH of various fluids and body compartments, while table 5.5 defines the four kinds of clinically significant acidosis and alkalosis. Both will be useful references as you read this section, and each should become familiar to the student and well-known to the practitioner.

TABLE 5.4

Acidity/alkalinity of fluids and body compartments in mammals

Location	*Conditions*	*pH range*
Stomach antrum	histamine-induced secretion of gastric acid	0.6–0.8
Secretory granules	basal, resting state	5.45–5.65
Tap water	at body temperature	6.75–6.95
Cytoplasm (inside cells)	sampled in any cell; skeletal muscle, nerve	7.15–7.25
Arterial plasma	sampled in any systemic artery at body temperature	7.38–7.42
Venous plasma	sampled in any systemic vein at body temperature (at rest)	7.28–7.38
Pancreatic fluid	after ingestion of a meal	7.95–8.15

From table 5.5, one can see the conditions that decrease alveolar ventilation or that impair diffusion of gases across the respiratory membrane and elevate the partial pressure of carbon dioxide in blood plasma and other extracellular spaces. This causes a secondary increase in the concentrations of hydrogen ions and carbonic acid in those same spaces. Carbonic acid (H_2CO_3) is produced when carbon dioxide (CO_2) combines with water (H_2O). This chemical reaction is called the hydration of carbon dioxide. Carbonic acid is a weak acid but can still contribute to respiratory acidosis. Respiratory acidosis can occur whenever damage to the respiratory membrane takes place, such as with disease, accidents, or smoking. It can also be caused by injury or disease to the respiratory control centers. Obstruction of the respiratory passageways caused either acutely, such as temporarily choking on food, or chronically, as with emphysema and asthma, can lead to respiratory acidosis. It is compensated for by increased buffer capacity of body fluids and/or by renal excretion of excess acids.

Respiratory alkalosis results from increased ventilation and reduced partial pressure of carbon dioxide. Diagnosis of the disorder might require evaluation of a sample of arterial blood. By analyzing the pH, $PaCO_2$, and HCO_3 values of the blood sample, clinicians can determine if the underlying cause is metabolic (kidneys) or respiratory (lungs). Then clinicians can determine corrective actions. Respiratory alkalosis rarely results from physical problems associated with the respiratory membranes or passageways. It occurs occasionally in association with neuropsychoses. The disorder is also commonly seen in those ascending to high altitude. High-altitude hypoxia stimulates increased ventilation and loss of carbon dioxide.

TABLE 5.5

Respiratory and renal conditions that lead to imbalances in physiological acid-base homeostasis in mammals

Condition	*Possible causes*	*Resultant changes*
Acidosis		
metabolic	HCO_3^- loss (diarrhea), ketones (diabetes)	reduced plasma pH, reduced plasma bicarbonate ions
respiratory	decreased ventilation, reduced diffusion	reduced plasma pH, elevated plasma bicarbonate ions
Alkalosis		
metabolic	$NaHCO_3^-$ therapy, reduced H^+ (severe vomiting)	elevated plasma pH, elevated plasma bicarbonate ions
respiratory	elevated ventilation, anxiety	elevated plasma pH, reduced plasma PCO_2

Lifestyle and Care of the Lungs

As a child and young man, I had seasonal allergies. I was either born with these or developed them early in childhood. I grew up in rural communities in the arid environments of the western United States. My allergies were most severe during the haying season, when dairy farmers haul and stack baled alfalfa (hay) during the hot dry months of June, July, and August. I was equally allergic to the sagebrush and scrub oak that are omnipresent in many areas of the western United States. Though my allergies manifested themselves in multiple ways during those years, they were most evident in my respiratory system. I had difficulty breathing during the day while working in the hayfields and for several days or weeks after. The severity of my allergies diminished later in life through avoidance of arid environments (I've lived in the humid Midwest or on the East Coast since 1971), treatment with drugs, and, I believe, with increasing age. Perhaps I developed some long-lasting if not irreversible problems in my respiratory system that I am still not aware of. However, I did not exacerbate these by choosing a lifestyle that included smoking cigarettes or abusing other tobacco products.

In the Midwest, and a few years later on the East Coast, my allergies seemed to intensify during the season when ragweed blooms (from about mid-August to the first freeze in October). My nasal passages were often swollen, my eyes watery, and my oral communication difficult and uncomfortable. As a teacher, this was a difficult time of year to suffer from such symptoms. My allergies were so severe at one point that I asked an emergency room physician at Michigan State University to help me find relief. He gave me an intramuscular injection of a steroid drug called Kenalog (triamcinolone acetate). Within a couple days my symptoms were gone, and I felt a tremendous sense of relief. That experience became a permanent part of my memory. A few years later at Rutgers University, I made an appointment with a physician who held both MD and PhD degrees and who specialized in treating allergies. I told him about my history with seasonal allergies, my discomfort with teaching during the ragweed season, and my Kenalog-mediated relief at Michigan State University. I said, "Will you consider experimenting with me?" and he replied, "Yes." During the next few weeks, he exposed me to a series of skin allergens and discovered that I was allergic to almost everything from animal dander to ragweed and grasses. He outlined my treatment options. I rejected, on the basis of time and expense, weekly injections for selected allergens. Instead he prescribed drops for my itching eyes (Vasocon), a spray for inflammation and nasal congestion (Vancenase, used orally to treat asthmatics), and an oral pill for general symptoms (a corticosteroid). He told me to use them only as needed. Within a few days of taking the first doses of these combined medications, I had no further symptoms during that season. By the following spring and summer, I noticed a marked decrease in my responses to allergens. I continued to take the medications only as needed

and very sparingly for the next few years. By about 1980, and ever since, I have been virtually allergy free. What an amazing difference! For the first time since my early childhood, I was freed of the discomforts, inconveniences, and ill health of seasonal allergies.

For students and other readers who have access to good health care, good medical insurance, and who suffer from seasonal allergies, I recommend doing something like I did. Get a referral from your primary care physician. With his/her input, find a physician who holds an MD and PhD, and whose PhD is in the field of immunology. That person should understand science and experimentation better than an allergist who lacks the scientific background and therefore has little or no appreciation for experimentation. Then work with this specialist in designing a course of action suited to your needs. If your outcome is as rewarding as mine was, your quality of respiration and life will be permanently changed.

6

Kidneys and Renal Physiology

Among the many physiological functions of the kidney—including those that are subject to feedback control—one is preeminent and omnipresent from birth to death. This is the need to maintain the homeostasis of body water and body electrolytes. Except for short-lived maladjustments, water and electrolyte balance among intracellular, extracellular, and intravascular spaces must be maintained twenty-four hours a day throughout a lifetime. Challenges to such a balancing act are presented by the daily cycles of hydration and dehydration we all experience. Consider the states of sleep and wakefulness. During the typical six to eight hours of sleep average adults get, we are steadily losing body water. This happens, in part, each time we exhale and lose water vapor to the atmosphere. It also occurs continuously with evaporation of water through the skin. Thus, during our sleep we lose a certain quantity of water but do not replace it. The net result is dehydration.

Normally from the moment we awake until about lunch time, we are consciously or unconsciously rehydrating. This is driven, in part, by eating breakfast, drinking multiple cups of java, and, earlier, by physiological feedback mechanisms promoting thirst: most notably a dry mouth and the afferent sensory signals it generates. During the sleeping hours, we steadily dehydrate and the homeostasis of water and electrolyte balance among the three compartments gets disturbed. Sensory signals from multiple locations throughout the body apprise the kidneys of these changes. The kidneys react via reflex mechanisms by conserving water. Water conservation during sleep takes place primarily at the distal tubules and collecting ducts of the renal nephrons. These tubular structures have specialized cells designed to transport water from the tubular fluid into the blood. In the membranes of such cells are proteins called aquaporins that form water channels. The number of such channels, that is, the production, release, and insertion of aquaporin proteins, is under physiological

control. Peter Agre of Johns Hopkins University shared the Nobel Prize in Chemistry in 2003 for his discovery of water channels.

As the body dehydrates, the osmolality of the blood gradually increases. The increased osmolality is sensed by osmoreceptors in the hypothalamus of the brain. The osmoreceptors cause the production and release of a hormone called antidiuretic hormone (ADH), or arginine vasopressin (AVP), from the posterior portion of the pituitary gland. Once in the systemic circulation, this hormone stimulates distal tubules and collecting ducts in the kidneys to increase the production of aquaporins. Then the kidneys are able to reabsorb more water, thereby excreting less water in the urine. This compensates to a limited extent for the dehydration that increased osmolality. This mechanistic response to imbalances in water and electrolyte distribution suffices until we awake and begin to rehydrate.

Interestingly, we quite often overhydrate in the first several hours of wakefulness. The evidence of this is the number of trips we make to the bathroom between late morning and early afternoon, the progressively lighter hue of our urine as that time period progresses, and the reduced desire to drink fluids (thirst satiety). I will have more to say about this daily cycle later in the chapter. For now, once we have overhydrated, the opposite feedback to the one described above takes over to compensate for the dilution of osmolality in the body water compartments.

Functional Morphology of the Kidneys

As mentioned in chapter 1, the kidneys are among the best examples in humans of the relevance of structure to function. The kidneys are bilateral bean-shaped structures located at the back and outside of the abdominal cavity. The peritoneum is the inner epithelial and connective tissue lining of the abdominal cavity. Abdominal structures like the stomach, liver, and intestines rub against or come in contact with the peritoneum. The kidneys, on the other hand, are located on the outside of the peritoneum and therefore, technically speaking, are not abdominal organs. There are two of them, and they are found dorsally just below the lower ribcage. They are the approximate size of the individual's loosely clenched fist, and in the average 70 kg adult, they weigh about 250 grams each. Each kidney is encapsulated in a rigid sheath made of connective tissue and called the renal capsule. The renal hilus is the point of entry and exit of the ureter, blood vessels, and renal nerves. The top of the kidney where the adrenal gland is attached is called the upper or north pole, and the bottom is the lower or south pole of the kidney.

If a kidney is removed and cut into two halves from top to bottom poles in the sagittal plane (any good medical dictionary will have a diagram), further detail of the kidney's structure is revealed. The innermost area where the ureter and other tubular structures arise is called the renal pelvis. Among its components are the

major and minor calyses (calyx is the singular). The calyses are contiguous with a region called the renal medulla. Just outside the medulla and beneath the renal capsule is the renal cortex. Where the renal cortex and renal medulla meet is a transition zone called the juxtamedullary region. The functional and structural units of the kidney are called nephrons (see chapter 1). There are two populations of nephrons: those found in the cortex (cortical nephrons) and those found in the juxtamedullary region (juxtamedullary nephrons). Both groups of nephrons function similarly, even though the vast majority of nephrons are found in the cortex. The juxtamedullary nephrons play a predominant role in concentrating and diluting the urine; that is, they are important in the conservation of body water.

Each kidney has a million or more nephrons, and each nephron is composed of vascular and tubular components. In order, the physiologically most important vascular components are the afferent arterioles, the glomerular capillaries, and the efferent arterioles. These segmentally arranged blood vessels conduct renal blood flow to and from the nephron, and their hemodynamics of pressure, flow, resistance, and blood viscosity determine filtering efficiency of the individual nephrons. The tubular components of the nephron begin with an inflated, bell-shaped structure that surrounds the glomerular capillaries and is called Bowman's capsule. Bowman's capsule is contiguous with segmental tubular components called the proximal tubule, the loop of Henle, the distal tubule, and the collecting duct. Most of these structures are further subdivided, and the student is referred to any comprehensive textbook of medical physiology to learn about them.

Glomerular Filtration Rate and Urine

Filtration is one of the kidney's primary functions. Filtration of blood takes place in the kidneys at sites called glomeruli. The rate at which all glomeruli of both kidneys filter blood is called the glomerular filtration rate (GFR), expressed in milliliters per minute. It is one of the most important physiological variables that can be evaluated to reveal overall health of kidney function. In the average 70 kg adult human, both kidneys combined receive about 25 percent of the cardiac output. As explained earlier, cardiac output is the volume of blood pumped by either ventricle each minute. In the resting state, cardiac output is about five liters per minute (or 5,000 ml/min) for the average 70 kg person. This means that renal blood flow (RBF) is about 1.25 liters per minute in this average person.

Only plasma gets filtered across the glomerular capillaries of the nephron. From our discussion of hematocrit in chapter 4 you will recall that cellular elements including RBCs, WBCs, and platelets constitute about 45 percent of the total blood volume. The remaining 55 percent is composed of plasma, and most of this is water. Thus, renal plasma flow (RPF) is 55 percent of 1.25 liters/min, or 563 ml/min in the average 70 kg adult. Not all of this plasma gets filtered each minute. Using experimental and clinical techniques, physiologists have discovered that GFR in

a 70 kg adult is about 120 ml/min. If we divide GFR by RPF, 120/563 = 0.21, and multiply this by 100, we find that 21 percent of RPF gets filtered each minute. We call this the filtration fraction.

What is actually getting filtered (that is, removed from the plasma and excreted), and why is this important? By weight or by volume, plasma is greater than 90 percent water. This means that if we eliminate all the organic and inorganic elements from plasma—all the sugars, fats, proteins, ions, electrolytes, and minerals—what is left is water. Alternatively, if we weigh a given volume of plasma and then evaporate the water, the weight of the remaining elements will be about 90 percent less than the original weight of the plasma. I dwell on the composition of plasma to impress the student with how important water is to the process of glomerular filtration. The components of renal plasma that get filtered, including water (the solvent) and solutes dissolved in it, are delivered directly to Bowman's capsule. This capsular fluid is called the ultrafiltrate, and, except for a few constituents, it resembles renal plasma. It differs from plasma primarily in its protein concentration.

Many plasma proteins are large (they have molecular weights greater than 50,000 daltons or molecular radii greater than 42 angstroms)[1] and do not get filtered by the kidneys. The larger the protein or any other plasma constituent, the more difficult it is to get it across the glomerular filtration barrier. One important plasma protein, fibrinogen, is involved in blood clotting and has a molecular weight of about 50,000 daltons. It does not get filtered. Other plasma proteins include the gamma globulins that are important to the immune system. These can have molecular weights in excess of several hundred thousand daltons. They also do not get through the glomerular filtration barrier. Only small molecules are filtered.

Water and the ions and electrolytes that are dissolved in it are small molecules. Glucose, the main blood sugar, has a molecular weight of about 180 daltons. It is also a small molecule and so gets freely filtered. Any constituent of plasma, charged or uncharged, that has a molecular radius less than 20 Å is freely filterable, that is, its filtration is similar to that of the water molecule. Molecules greater than about 42 Å are not filtered. Those between 20 and 42 Å are more or less filterable depending on their charged state. Positively charged molecules are more easily filtered than their negatively charged analogs.

The glomerular filtration barrier is like the respiratory barrier and is composed of three layers of tissue: (1) the glomerular capillary endothelium, (2) a collagenous, proteoglycan basement membrane, and (3) the renal tubular epithelium with accompanying podocytes and pseudopodia. Pseudopodia are elongated, foot- or toelike projections of epithelial cells that circumscribe the glomerular capillaries at their external surface. The basement membrane of the filtration barrier, under physiological conditions, carries a net negative charge. Therefore, negatively charged molecules, even if small, are repelled by the barrier. In some individuals,

including conditioned athletes, a bout of strenuous exercise is accompanied by the appearance of protein in the urine. This is called exercise-induced proteinuria and is thought to be due to the transient removal of electrical charges from the basement membrane. The glomerular filtration barrier is extremely effective in forming glomerular ultrafiltrate.

The structure of the glomerular filtration barrier is analogous to the drain of the kitchen sink. The drain plug and the drain trap lie in close proximity with wastewater passing through the plug first. The plug has small holes in it: about ten to twenty of 4 to 8 millimeters in diameter. The drain trap has fewer, usually about four, and larger holes, often 10 to 15 millimeters in diameter. Functioning together, the two keep dishwater contents larger than 4 millimeters out of the house plumbing system. This is especially important if the system is not equipped with an automatic garbage disposal between the sink drain and the house plumbing.

Physiologists can sample the ultrafiltrate in Bowman's capsule using anesthetized, instrumented animals and micropipette techniques. When they do this, they find that the water content of the ultrafiltrate is similar to that of plasma, as are the concentrations of important electrolytes such as sodium (Na^+), potassium (K^+), calcium (Ca^{2+}), and bicarbonate (HCO_3^-). So ultrafiltrate is said to be osmotically equal to or iso-osmotic with plasma. However, if tubular fluid is sampled downstream to Bowman's capsule (for example, in the proximal tubule, loop of Henle, distal tubule, or collecting ducts), the outcome is quite different. That is because electrolytes, water, and other elements of the ultrafiltrate are processed differently by each segment of this tubular nephron. Most ultrafiltrate gets reabsorbed in the proximal tubule. Reabsorption means the removal of individual elements from the ultrafiltrate and their return to the systemic circulatory system. The reabsorbed chemicals initially get transferred from the tubule into peritubular capillaries. These are renal capillaries that lie immediately downstream of the efferent arterioles. From the peritubular capillaries, the reabsorbed material is delivered to the general systemic circulation. In the kidneys we find a renal-portal, or renoportal, circulatory arrangement. This means that renal arterial blood flow gets delivered to two sets of capillaries that are arranged in series with each other and with other renal vasculature. The renal-portal system is composed of the following structures: afferent arteriole → glomerular capillary → efferent arteriole → peritubular capillary → renal vein. This means that renal blood flow is subject to dual processing via the mechanisms of filtration, reabsorption, and secretion.

The glomerular filtration rate, as mentioned above, is about 120 ml/min. If most of the ultrafiltrate formed each minute got excreted as urine rather than being reabsorbed, then it would take less than one hour to deplete the body of its circulating volume of plasma. By measuring the daily rate of urine excretion, however, we can easily discover that slightly less than 100 percent of the ultrafiltrate gets reabsorbed each day. At 120 ml/min (GFR), about 173 liters/day of ultrafiltrate are produced. Urine excretion in the average 70 kg adult is rarely more than

2 liters/day. Thus 173 − 2 = 171 liters/day are reabsorbed, or about 99 percent of the daily formation of ultrafiltrate. Under physiological conditions, tubular reabsorption is a highly efficient process. It is designed to conserve products the body needs, such as water and electrolytes, and to eliminate those that are potentially harmful, such as urea, ammonia, and other by-products of metabolism.

Regulation of GFR and RPF

The rate of glomerular filtration is a tightly regulated physiological variable: it obeys the law of conservation of mass, and it depends on the intact hemodynamics of the renal circulatory system. Moreover, the physiological control of GFR is intimately linked to the regulation of RPF. The variables that are important in regulating GFR are called Starling forces and include pressures inside and outside glomerular capillaries. There are two pressures in each location. Inside the glomerular capillaries, these pressures are capillary hydrostatic pressure and colloid oncotic pressure. Outside the glomerular capillaries, that is, inside Bowman's capsule, the pressures are capsular hydrostatic pressure and capsular colloid oncotic pressure. Colloid oncotic pressure, sometimes called colloid osmotic pressure or simply oncotic pressure, occurs only in the presence of proteins. Since the plasma inside glomerular capillaries contains proteins but the ultrafiltrate, under physiological conditions, does not, colloid oncotic pressure is regularly much higher inside glomerular capillaries than it is in Bowman's capsule. Hydrostatic pressure is influenced by flow and resistance inside glomerular capillaries and Bowman's capsule. Generally speaking, the greater the volume of flowing fluid, the greater the hydrostatic pressure. Therefore the things that can change Starling forces are those that will subsequently alter GFR and RPF. Among the most important of these are pre- and postcapillary resistances. Precapillary resistance occurs mostly in the afferent arteriole upstream to the glomerular capillaries. Postcapillary resistance is seen mostly in the efferent arterioles that are downstream to the glomerular capillaries. Both sets of arterioles are richly endowed with vascular smooth muscle cells. They are also innervated by motor nerves of the sympathetic nervous system. Thus, both sets of arterioles respond to neurogenic stimuli as well as to vasoactive compounds that are in the general systemic circulation. It is therefore conceivable that either set of stimuli—nerves or circulating compounds—will affect one and not the other set of arterioles.

Imagine the following hypothetical condition. A patient or experimental subject has taken a drug that selectively constricts renal efferent arterioles. As the arterioles constrict, blood flow through them decreases. As a consequence of the reduction in efferent blood flow, the volume of blood inside the upstream glomerular capillaries increases. This causes an increase in capillary hydrostatic pressure. If other Starling forces fail to change correspondingly, then the balance of Starling forces will favor increased loss of fluid and ultrafiltrate into

Bowman's capsule, that is, GFR will increase. Conversely, if precapillary resistance increases and all other Starling forces remain constant, then blood flow through the glomerular capillaries will decrease. This will cause a reduction in hydrostatic pressure and a corresponding decrease in the formation of ultrafiltrate. That is, GFR will decrease. One can see how simple changes in pre- and postcapillary hemodynamics can have a direct effect on the rate of glomerular filtration. Changes in the intravascular concentration of plasma proteins can have an equally significant impact on GFR but such changes are less common than those in pre- and postcapillary resistances.

Another variable that potentially threatens the homeostasis of renal function is renal arterial blood pressure. Changes in renal arterial blood pressure can have direct effects on glomerular capillary blood flow, glomerular capillary hydrostatic pressure, and GFR. Elevations in renal blood pressure (RBP) increase GFR and decrements in RBP decrease GFR. Thus, without a system of checks and balances one can see how GFR would be constantly influenced by changing renal hemodynamics. Fortunately the kidneys were designed with such a system of checks and balances. The system is called pressure-flow autoregulation. It is a marvelous design, one that serves not only to keep renal blood flow nearly constant but to also regulate GFR. Pressure-flow autoregulation in its classic definition simply means the ability of an organ to maintain a relatively constant blood flow over a wide range of blood pressures. Pressure-flow autoregulation is most well developed in the mammalian kidney but is seen also in many other organs such as the heart, brain, skin, and skeletal muscle. This phenomenon has been observed for many decades by multiple investigators under a variety of experimental and clinical conditions.

To explore autoregulation, the investigators use either anesthetized instrumented animals or isolated perfused whole organs. Imagine that you are a research assistant in my laboratory and that we are using an isolated perfused kidney to study the phenomenon of renal pressure-flow autoregulation. The kidney we are investigating has been removed from the animal's body and is being sustained under near-physiological conditions. It must be instrumented to allow us to control renal perfusion pressure while at the same time being able to measure renal blood flow and GFR. Let's also imagine that thirty minutes have passed since we set up the preparation, so these three monitored variables—renal blood pressure, renal blood flow, and GFR—are in their steady states. At this point, we collect some baseline control data, and the numbers look like this: RBP = 100 mmHg, RBF = 80 ml/minute (RPF is about half this number), and GFR = 40 ml/minute (see figure 6.1). At our discretion, we then adjust RBP to equal 150 mmHg (a 50 percent increase) and at the same time continuously monitor RBF and GFR. Almost immediately upon elevating RBP, RBF and GFR increase to about 120 and 60 ml/min, respectively. However, in the next several minutes and while maintaining RBP at 150 mmHg, RBF and GFR steadily decrease to new steady state values

of about 80 and 40 ml/min. In other words, despite the maintenance of elevated blood pressure both blood flow and the rate of glomerular filtration have returned to their baseline, control levels.

At this point in the experiment, we return RBP back to its baseline control level of 100 mmHg and allow the kidney to restabilize for several minutes. Then we collect another set of baseline data and obtain numbers that are the same as those found during the first control period. We then reduce RBP from 100 to about 50 mmHg and find that RBF and GFR also decrease, almost immediately, by about 50 percent each. However, as the low RBP is maintained over the next few minutes, both RBF and GFR return to levels that are not substantially different from their previous baseline values.

From this experiment of elevating and reducing renal blood pressure in an instrumented isolated kidney, one can see that despite increments or decrements in RBP, RBF and GFR in the steady state are maintained at nearly constant levels. This is what renal pressure-flow autoregulation means. The fact that GFR is correspondingly maintained in the steady state means that it too is autoregulated (see figure 6.1).

Pressure-flow autoregulation is a marvelous phenomenon. It involves an even more impressive structure, the justaglomerular apparatus or JGA. The student is encouraged to explore this structure in much greater detail than I describe here.

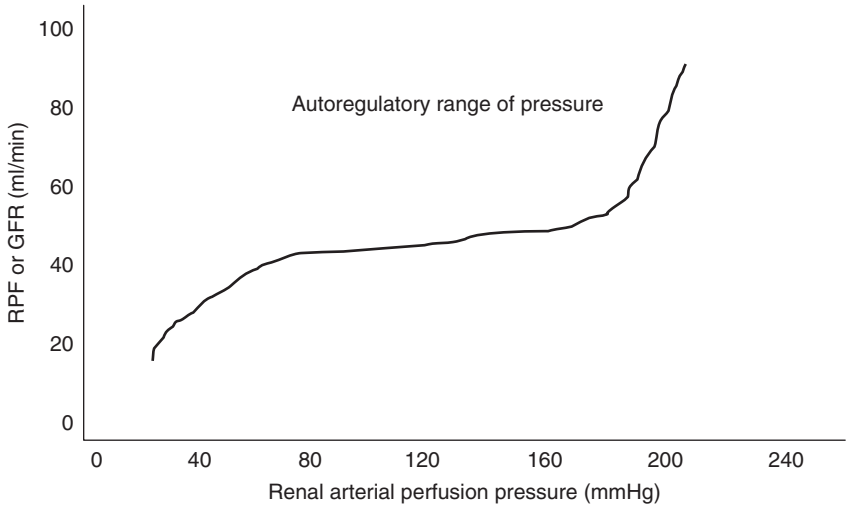

FIGURE 6.1 The blood flow and renal function phenomena of autoregulation (tilda-shaped curve). Note that between renal arterial blood pressures of 0 to 80 mmHg, as pressure increases so does renal plasma flow (RPF) and glomerular filtration rate (GFR). A similar phenomenon is observed above pressures of about 180 mmHg. However, between pressures of 80 and 180, as pressure is elevated (or lowered), RPF and GFR remain relatively constant (autoregulation).

Minimally, the JGA consists of renal tubular and vascular components that lie in close proximity to each other. Some of the cells of the walls of each are specialized to communicate with each other. Tubular epithelial cells detect the concentration of sodium chloride flowing by them. Specialized vascular smooth muscle cells detect changes in renal arterial perfusion pressure. Working cooperatively, these two sets of cells in the JGA as well as other tissues help maintain blood pressure.

The fact that autoregulation occurs in all major mammalian organ systems means that our bodies are able to adjust to stimuli and challenges that have important effects on arterial blood pressure and related physiological variables such as GFR. Some of these stimuli include moderate to heavy exercise, hypotensive crises such as blood loss, hyperthermia, and hydration/dehydration. Several mechanisms have been proposed to explain pressure-flow autoregulation, but only a couple have withstood the test of time. They are the myogenic and metabolic mechanisms, both of which are beyond the scope of this book. *Myogenic* means that vascular smooth muscle cells in the walls of blood vessels have the ability to detect changes in transmural pressure (stretch and destretch of vessel walls). If the transmural pressure increases (that is, if the vessel is stretched by an increase in blood flow), the smooth muscle cells contract, causing luminal diameter to decrease and blood flow to fall. The metabolic mechanism has to do with the rates of production and removal of agents that are vasoactive and can therefore produce changes in blood flow.

Reabsorption, Secretion, and the Formation of Urine

Once the ultrafiltrate leaves Bowman's capsule, it passes sequentially through the proximal tubule, descending then ascending limbs of the loop of Henle, distal tubule, and collecting duct. Beyond the collecting duct, the ultrafiltrate is no longer subject to manipulation by the nephron and is destined to be excreted from the body as urine. Of course even after the collecting ducts, the ultrafiltrate still passes through several tubular and other structures including the ducts of Belloni, the minor and major calyses, the renal pelvis, the ureters, and finally the urethra.

Most of the ultrafiltrate is reabsorbed by the proximal tubule. Absorbed substances include, in equal proportions, both water and solutes. Remember that the ultrafiltrate in Bowman's capsule is iso-osmotic with plasma, that is, both fluids have the same osmolarity of about 280 milliosmoles per liter of fluid (280 mOsm/l). However, depending on the physiological state of the body (for example, whether it is well hydrated or dehydrated), ultrafiltrate collected from locations downstream of the proximal tubules can vary markedly in its osmolarity. In the normally hydrated person, the osmolarity of all body water compartments is about 280 mOsm/l. Under these conditions, fluid in the descending limbs, both thick and thin segments of the loop of Henle, becomes progressively more concentrated. Osmolarity increases because these segments of tubule are

highly permeable to water and water gets reabsorbed. Thus, by the time the ultrafiltrate reaches the tip of the loop in the inner medulla of the kidney, it might have an osmolarity of 600 mOsm/1, or about double that of plasma. The ascending limb of the loop of Henle is impermeable to water both in the absence and presence of antidiuretic hormone (ADH). However, it is highly effective in reabsorbing solutes such as sodium, potassium, and chloride. Thus, even though it receives a more concentrated ultrafiltrate from the descending limb, it sends a dilute ultrafiltrate to the distal tubule.

In the absence of ADH, the distal tubule and collecting ducts are impermeable to water. They too actively transport sodium, chloride, and potassium back into the circulation thus making the tubular fluid even more dilute. This can be seen in an experimental setting where subjects voluntarily drink large volumes of water. If the volume of urine flow and the osmolarity of the urine are simultaneously measured at half-hour intervals before the water is consumed, the urine flow rate might average 0.5 ml/min and the urine might have an osmolarity in excess of 500 mOsm/1. After 60 to 120 minutes of drinking one liter of water, the urine flow rate will increase and the osmolarity will decrease. This is called a water diuresis, and at the peak of such a response, urine flow rate can reach 4 to 6 ml/min or more, and urine can have an osmolarity of less than 100 mOsm/1 (that is, the person produces a dilute urine). Under such conditions, central and peripheral mechanisms block the release of ADH from the pituitary gland thus preventing its actions on the distal tubule and collecting ducts. When a person is seriously dehydrated, the above scenario is reversed. Under these conditions, ADH release might be maximal and the person could excrete a highly concentrated urine, one with an osmolarity that exceeds 1,000 mOsm/1. Urine flow rate in this dehydrated state is minimal, for example, less than 0.1 ml/min. Only enough water is excreted to dissolve the inorganic/organic solutes the body needs to get rid of. That minimal amount of excreted water is called obligatory water loss.

The specific and detailed mechanisms by which the kidneys are able to concentrate and dilute the urine are beyond the scope of this book. However, I urge the interested reader to study these in any of the textbooks of medical physiology listed in the suggested reading.

Renal Health, Hydration, and Urination

The state of hydration of our bodies changes throughout the day. Understandably, hydration state is affected by where we live, what we do, the time of day, and season of the year. Residents of the Southwest, for example, Phoenix, Arizona, consume more water than residents of the Northeast, such as Boston, Massachusetts. In either location residents consume more water in a typical August than in a typical January. This seems like common sense. Nonetheless federal and state agencies try to keep a handle on such information and occasionally report it. Water

consumption patterns quite possibly have been affected by the recently developed bottled and flavored water and sports drinks industries. Nowadays, students with bottled water and discarded water bottles are ubiquitous.

We consume water from three different sources. If we drink water from the tap or from a bottled water cooler, we consume it directly. If we use tap water to make a beverage such as coffee, Postum, hot chocolate, and soda, then we consume water indirectly. Finally, water consumed as a natural component of the foods we eat, or that is found in fruits and vegetables or as a beverage such as cow's milk, is consumed as intrinsic water. Each day, the average American adult consumes approximately two liters of water by these three routes. We also lose about two liters of water each day via the physiological processes of urination, defecation, perspiration, and evaporation. Some of these processes occur on a continuing basis without us being aware of them, for example, expired water vapor or evaporation via the skin. We are better hydrated in the late morning to early evening hours and less hydrated in the late evening to early morning hours, cyclically and diurnally.

Consider the states of hydration of the body upon awakening in the morning and several hours after breakfast. We awake relatively dehydrated. We rehydrate as soon as we begin consuming water by any of the routes defined above. Imagine that you awake at 6 A.M. and go directly to the bathroom to urinate. If you are physiologically normal and observant, you will note that your urine is a darker hue of yellow than it will be at 11 A.M. after eating breakfast, drinking water directly, and having had your second or third cup of coffee. Relative to 6 A.M. when you awake dehydrated, your body is well- if not overhydrated at 11 A.M. When you urinate at 11 A.M., the volume excreted is usually greater than that at 6 A.M. Moreover, the color of urine at 11 A.M. is a lighter hue of yellow. As a rule, the greater the state of body hydration, the lighter the hue of urine. Alternatively, the greater the state of dehydration, the greater the yellow intensity of the urine. All this is affected by the kidney's ability to either conserve or eliminate water.

To teach this concept to upper-level premedical students in an experimental physiology laboratory course at Rutgers University, we divide them into three groups. One group serves as the control or placebo group. The second group consists of overhydrated students, and the third group is composed of students who consume several salt tablets to simulate dehydration and body water conservation. When they enter the laboratory, the students in the three groups go to the bathroom and urinate in graduated beakers. They record the time and urine volume and then measure the osmolarity of the urine samples. Students in the overhydrated group then drink one liter of water as rapidly as possible. Students in the dehydrated group consume several salt tablets. The placebo group students do nothing; all students collect urine and measure its volume and osmolarity each thirty minutes during a 180-minute time period.

The outcome is predictable. Those consuming salt tablets will excrete the lowest urine volumes with the highest osmolarities (due mainly to the high

sodium concentration). The greatest urine volumes and lowest osmolarities will be produced in the group who drank the liter of water. Values in the control group will fall between these two extremes. This is a real yet simple and inexpensive way of teaching students how the kidneys either concentrate urine by conserving body water or dilute it to rid the body of excess water and reestablish homeostatic osmolarity among the various body water compartments. The students who take the salt tablets represent all of us when we awake dehydrated before breakfast. The students who drink the liter of water represent all of us in the late morning to early evening after we have rehydrated.

As another illustration of the physiological principles of hydration/dehydration and osmoregulation, consider beer-drinking fans at sports events and concerts. By weight or volume, beer and many other alcoholic drinks are greater than 90 percent water. Thus after several hours of tailgaiting, beer drinkers are well-hydrated experimental subjects. Equally important, alcohol is a strong inhibitor of the production and release of ADH. In the absence of physiological concentrations of ADH, the distal tubules and collecting ducts of the nephrons cannot reabsorb water with or without the presence of alcohol. Thus, inebriated fans not only challenge their bodies by overhydrating, but they prevent ADH from doing its physiological job.

Monitor your own patterns of output for several days. Record the results in a place where they are secure and permanent. Repeat the observations several years later, then again in a few more years. If this is done over a lifetime, you will not only begin monitoring your kidney health but will learn about the relationships among the kidneys, the circulation, and the respiratory system. On a practical day-to-day basis, you will be able to apply the knowledge gained for your own comfort and convenience. For example, whenever you have mid-morning meetings or other activities that last more than two hours, you will find yourself visiting the restroom before the meeting begins rather than having to interrupt it because of an overfilled urinary bladder. You will be able to drive a distance of one hundred miles or more without stopping at a rest area to urinate. Most importantly, once you have confirmed the diurnal nature of kidney function in your own life, should it ever change without explanation you will be motivated to see your doctor and have her refer you to a nephrologist. You would be less likely to do this if you didn't know about the daily kidney patterns in the first place.

Diabetes and Sugar in the Urine

One of the small organic molecules filtered by the kidneys is blood sugar. Blood sugar or glucose is a monosaccharide, one of the simplest of sugars. In the average adult eating a Western diet of approximately 2,000 calories per day with approximately 30 percent fat, 60 percent carbohydrate, and 10 percent protein, glucose concentrations in the plasma average about 80 to 100 mg/dl in the

postabsorptive state, that is, about four hours after consuming a meal. It might be higher than this during the absorptive state during the actual ingestion of a meal. Under these conditions and in a normally healthy individual, glucose does not appear in the urine. That is because 100 percent of the glucose in the ultrafiltrate gets reabsorbed in the proximal tubule.

The walls of the tubular nephron consist of epithelial cells. The portion of the cell membrane that faces the lumen of the tubule is called the apex of the cell or the apical membrane. The portion of the cell that faces the interstitial spaces and blood vessels (regions of the cell that are not in contact with the tubular lumen), are called the base or lateral walls of the cell (or the basolateral membrane). At each location, there are proteins and membrane channels or pores that extend across the cell membrane from the extracellular tubular lumen to the intracellular or cytosolic spaces. The proteins bind specific components of the ultrafiltrate and transport them from the tubular lumen to the interior of the cell. Other mechanisms move the same components from the interior of the cell, across the basolateral membrane, and into the interstitial and intravascular compartments. One of these specific transmembrane (also called transmural) proteins has two binding sites: one that attaches to sodium ions in the ultrafiltrate and another that binds glucose. The proteins that are specific for glucose and sodium transport across the apical cell membrane are called sodium glucose cotransporters or SGLTs. There are several different SGLTs, and they have different capacities and affinities for binding and transporting glucose and sodium. About 98 percent of glucose is reabsorbed in the first one-third of the proximal tubule. This is called the S1 or straight segment of the proximal tubule. The remainder of the filtered load of glucose is reabsorbed downstream in the S3 or convoluted proximal tubule. In the S1 segment, a high-capacity, low-affinity cotransporter called the $SGLT_2$ mediates reabsorption of glucose. In the S3 segment, a high-affinity, low-capacity cotransporter called $SGLT_1$ is responsible for apical uptake of glucose.

Because most of the glucose is reabsorbed from the early segments of the proximal tubule, where the amounts of glucose in the ultrafiltrate and in the plasma are about equal, there is a steep glucose concentration gradient between tubular fluid and cell interior in the downstream segments of proximal tubule. Thus, the SGLT in that region of the tubule must be able to generate much more energy in order to move glucose over this large uphill gradient. The stoichiometry of the two SGLTs reveals that the downstream cotransporter is 50 to 100 times more energetic than the upstream cotransporter. Once glucose enters the tubular epithelial cell, it exits across the basolateral membrane via a member of the glucose transport family of proteins (GLUTs). Because these transport proteins are not dependent on sodium ions, they move glucose out of the cell by the process of facilitated diffusion. Also, GLUTs do not use energy in the form of ATP to translocate glucose, so they are quite distinct from the SGLTs. However, like

the SGLTs, the GLUTs vary in their affinities and capacities for glucose between the upstream and downstream segments of the proximal tubules.

One can study the influence of changes in plasma glucose concentrations on the renal processes of filtration, reabsorption, and excretion by infusing glucose intravenously and then measuring the concentrations of glucose in the plasma and in the urine. Some results of measurements of this kind are shown in figure 6.2. The three curves in this figure represent (1) the filtered load for glucose or the product of the plasma concentration of glucose and GFR, (2) the rate of glucose reabsorption (middle curve with the rightward inflection followed by a plateau), and (3) the rate of renal excretion of glucose (bottom curve with a leftward inflection near the x-axis).

There are several things to note in figure 6.2. First, the physiological concentrations of glucose in the plasma of healthy humans is about 75 to 125 mg/dl. Under these conditions, there are no molecules of glucose in the urine, that is, renal excretion of glucose is zero (the excretion curve is superimposed on the x-axis). Note also that over the above range of plasma glucose, as the concentration of glucose in the plasma increases (by infusing more glucose into the vein) both the filtered load and rate of reabsorption increase in parallel fashion. Excretion of

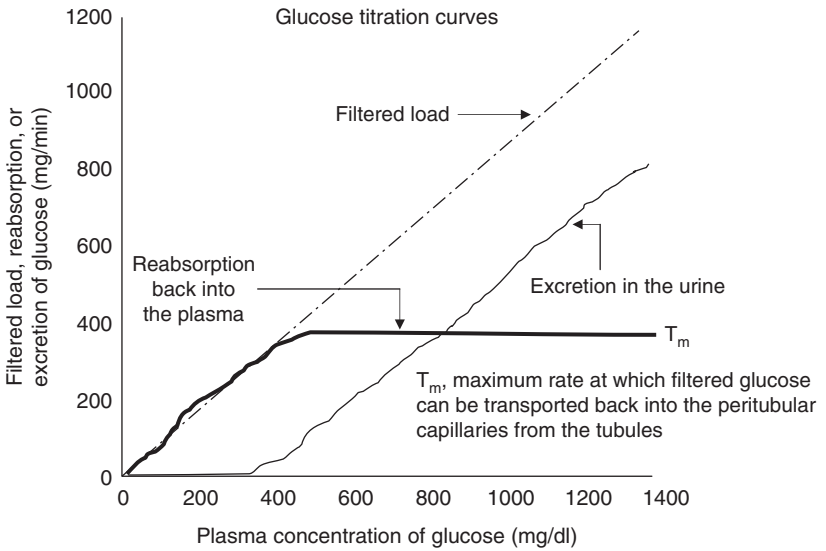

FIGURE 6.2 The influences of circulating plasma concentrations of glucose (blood sugar, abscissa or x-axis) on the filtered load, reabsorption, and excretion of glucose. Note that in a typically healthy person, no excretion of glucose occurs below plasma concentrations of about 300 mg/dl. In other words, below this level, all glucose that gets filtered at the glomeruli is also reabsorbed (see parallel or superimposed curves for filtered load and reabsorption). Clinicians refer to these observations as glucose titration curves.

glucose in the urine does not occur until plasma concentrations of the sugar reach 250 to 300 mg/dl. Also note the splay, or upward-oriented curve away from the x-axis, in the excretion curve over the range of about 250 to 350 mg/dl. This reflects variability among individual proximal tubules and their associated SGLT and GLUT transporters. By comparing all three curves, we can see that as plasma concentrations of glucose reach 250 to 350 mg/dl, tubular reabsorption reaches a maximum—that is, it plateaus despite further increments in concentration—and renal excretion parallels glomerular filtration (filtered load). T_m stands for transport maximum and means that the tubular mechanisms for reabsorbing blood sugar are saturable and have a finite limit. The most common cause of problems related to exceeding T_m in humans is the disease diabetes mellitus or sugar diabetes.

Diabetes mellitus is a serious metabolic disease with many identifying hallmarks, the most broadly recognized being elevated blood sugar. The disease can arise at any time in life and is caused by alterations in the synthesis, release, and tissue actions of insulin. When the disease occurs in children, it is called insulin-dependent diabetes mellitus or IDDM (type I diabetes). When it occurs in adults, particularly the obese, it is designated non-insulin-dependent diabetes mellitus or NIDDM (type II diabetes).

Two hormones, synthesized and released by the pancreas, play the key roles in mediating the metabolism of glucose and related compounds. These are insulin and glucagon. Insulin is produced by the beta cells and glucagon by other cells. In type I diabetes, the body's immune system destroys the beta cells. Thus the pancreas cannot produce insulin, but glucagon continues to be produced and to cause the release of glucose and ketone bodies from the liver. A marked rise in plasma levels of glucose and ketones occurs, producing a tremendous load for the kidneys to filter. This causes an osmotic diuresis or excess loss of water due to the excretion of osmotically active glucose and ketone bodies. Also the ketone bodies are moderate organic acids and cause a marked reduction in plasma pH. Unless corrected, these two conditions, osmotic dehydration and acidosis, will lead to serious tissue and organ damage and death.

Individuals with either type of diabetes appear to be at risk of chronic tissue damage due to elevated levels of glucose. The problem is particularly targeted at the eyes, kidneys, peripheral nerves, and extremities. Use of frequent self-injections of insulin or insulin pumps appear to attenuate the development of the pathologic changes associated with diabetes. This suggests that it is the prolonged elevation of glucose itself that causes the disease. It goes without saying that a lifestyle that is directly correlated to diabetes and kidney dysfunction includes poor diet, lack of exercise, and obesity. Barring genetic and metabolic considerations over which the reader has no control, it is within the grasp of everyone to change these patterns that will improve renal health specifically and other organ and tissue health in general.

7

The Gastrointestinal System

The mammalian gastrointestinal system is also known as the digestive tract or the enteric system. It is a complex system performing mechanical, secretory, digestive, absorptive, and excretory functions. Each of these is under the influence of local gastrointestinal reflexes as well as central feedback control mechanisms. Consider, for example, the medical problems a person might have if she ate three meals per day for several days without having a bowel movement. To prevent the intestinal storage of food wastes and the pathogens they support, a wide variety of gastrointestinal reflexes exist. In the well-tuned, normally functioning mammalian digestive tract, shortly after the consumption of one meal a defecation reflex is activated, causing contraction of the colon and initiating a bowel movement. Nearly simultaneously, mechanical distension of the stomach initiates a gastro-ileal reflex. This causes the ileum to contract and to empty its contents into the colon. Emptying of the ileum makes room upstream as the stomach passes its contents into the duodenum.

In many humans who have bowel movements at approximately twenty-four-hour intervals, the gastro-ileal reflex is activated within thirty to sixty minutes after ingestion of the morning meal. This means the colon has space to accommodate another meal when it is ingested. In all mammals, there is a centrally located satiety center in the brain stem. Usually about four to six hours after the last meal, sensations of hunger, often originating in the stomach, signal to the person that it is time to eat again. Some refer to these abdominal sensations as hunger pains or hunger pangs. Their specific cause is not known. However, consumption of a meal resolves the sensations of hunger, and that individual usually will not be motivated to seek food for another four to six hours.

Of course the normal physiology of hunger and satiety is influenced by disease and other interventions. Any person who has experienced food poisoning such as that caused by *E. coli* or salmonella knows that during the period of

vomiting and diarrhea and for a day or two thereafter the last thing one thinks of is eating more food. Other changes in behavior can also influence digestion and bowel regularity. Examples include travel, fasting and feasting, and prescription as well as over-the-counter medications. Simply modifying one's eating habits can markedly influence behavior of the gastrointestinal system.

My wife and fifteen-year-old son ate considerably less than I did. I decided one time to reduce my daily calorie intake. My plan was to observe them for one month at all meals we ate together. After they had taken their portions, I would take mine based on my visual estimate of theirs. I did this for one month. Many things happened. I lost five pounds and my slacks fit more loosely. Most interestingly, my gastrointestinal functions changed markedly. For example, bowel movements became less frequent and my stools less bulky. I attributed these changes to the reduced daily intake of calories; nothing else had changed in my life during that period. I enjoyed that act of self-discipline and learned a few more things about myself.

Our diets and changes in them affect us differently. The numbers, distributions, and physiological efficacies of membrane-bound digestive enzymes, among other physiological characteristics of the enteric system, differ from person to person. Thus, two people making the same change in their similar diets cannot expect the same outcome in terms of pounds lost.

Components of the Gastrointestinal (GI) or Enteric System

The GI system begins with the oral cavity and ends at the anus (see figure 7.1). The entire system is designed to mix and churn, store and propel, and digest and absorb. These processes begin in the mouth and end in the colon or cecum (for guinea pigs and other mammals). In the healthy individual, they continue for a lifetime. In the person with an unhealthy GI tract, hyperalimentation or feeding via intravenous infusions might be required and can sustain life for prolonged periods.

We can identify points of demarcation along the entire length of the GI tract that help define functions. One such set of points is called sphincters. A sphincter is a band of visceral smooth muscle—sometimes skeletal muscle, other times a mixture of visceral smooth and skeletal muscle—that circumscribes a small but well-defined segment of the GI tract. Sphincters are also comprised of longitudinal and circular smooth muscle cells. These form separate layers of muscle that run the entire length of the GI tract and are arranged in parallel (longitudinal), or lie perpendicular to the long axis of the gut (circular). Six sphincters are of primary importance (see table 7.1). Beginning at the oral and ending at the anal ends of the tract, that is, going in an aboral or caudad direction, these are (1) the upper esophageal sphincter (UES), (2) the lower esophageal sphincter (LES), (3) the pyloric sphincter (PS, located between

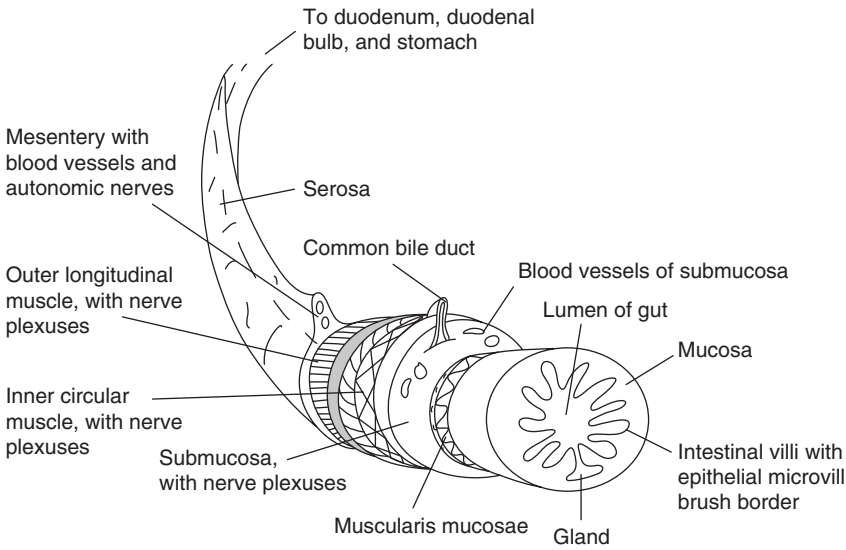

FIGURE 7.1 Anatomic arrangements of the gastrointestinal tract (GIT). Note functions of mesentery (suspends gut loops, carries nerves and blood vessels, keeps other organs in close proximity to gut and pancreas), and locations of muscle layers with corresponding nerve plexuses. Blood vessels simulate arteries, microcirculation, and veins.

(Author's illustration)

TABLE 7.1

Location and function of sphincters in the mammalian gastrointestinal tract

Sphincter	*Location*	*Function*
Upper esophageal sphincter	superior end of esophagus	prevents entry/exit of materials into esophagus if closed
Lower esophageal sphincter	inferior end of esophagus	regulates entry into esophagus, prevents gastric reflux
Pyloric sphincter	antrum and duodenal bulb	helps regulate emptying of the stomach
Ileocecal sphincter	ileum at junction with colon	regulates emptying of contents of ileum into large intestine
Internal anal sphincter	internal anal canal	involuntary control, opens with pressure and contents
External anal sphincter	external anus distal to rectum	voluntary control, opens with conscious movement of bowels

the antrum of the stomach and the duodenal bulb of the small intestine), (4) the ileocecal sphincter (IS, located where the ileum joins the ascending colon), (5) the internal anal sphincter (IAS), and (6) the external anal sphincter (EAS). Between any two adjacent sphincters, the GI tract performs various special functions.For example, the stomach lies between the LES and the PS. It has many special functions, but one important one is the temporary storage of a recently ingested meal.

Consider a cross section of a segment of the small intestine. Imagine you are peering into the lumen. The surface of gut wall that is in immediate contact with the contents of the lumen is called the mucosal side, and the surface that is farthest removed from the luminal contents is called the serosal side of the gut wall. In addition to the longitudinal points of demarcation defined above, these transmural—across the wall—points of demarcation help define the makeup of the gut wall. From mucosal to serosal sides, the layers are (1) the mucosa, (2) the muscularis mucosae, (3) the submucosa, (4) the layer of circular visceral smooth muscle, (5) the layer of longitudinal visceral smooth muscle, and (6) the mesentery, a connective tissue sheath that encloses the GI tract and carries the mesenteric arteries and veins adjacent to the serosal side of the system.

The mucosa is composed of specialized epithelial cells whose collective apical surfaces form the brush borders of the mucosa. Each individual epithelial cell has an apex composed of microvilli, small finger-like projections that protrude into the lumen of the gut. The microvilli increase the surface area for digestion and absorption. Remember the analogy of alveolar surface area and drying towels on a clothesline or cubed versus block ice and their rates of melting. A similar analogy is applicable to the microvilli of the gut wall and maximal area available for secretion, digestion, and absorption.

There are about five different kinds of epithelial cells in the gut wall. Each is designed to perform a different function. For example, some secrete gastric acid (hydrochloric acid), others secrete digestive enzymes, and still others release mucus. There are also many crypts, or glands, in the mucosa. The muscularis mucosa is a layer of visceral smooth muscle that is found between the mucosa and the submucosa. It causes foldings and invaginations of the underlying mucosa thereby increasing exposure of food materials to the secretory, digestive, and absorptive processes taking place there. The submucosa contains many blood vessels, glands, and nerves such as Meissner's plexus and is the point of entry of the common bile duct. The circular and longitudinal layers of smooth muscle are separated by another nerve network called the myenteric or Auerbach's plexus. The Auerbach's and Meissner's networks of nerves constitute the small gut or enteric nervous system (ENS) described earlier in this book. Some authors consider the enteric nervous system to be a third subdivision of the autonomic nervous system.

Mechanics and Reflexes of the GI Tract

Mechanical activities of the GI tract as well as the first gastrointestinal reflex begin in the mouth. The processes of ripping, incising, crushing, and chewing a mouthful of food are collectively called mastication. Masticating food is a mechanical activity. It involves forceful movements of muscles, jaws, and other organs not directly connected to the GI tract. After several seconds or minutes of mastication, the voluntary decision to pinch off and swallow a bolus of food is made. As the tongue presses a quantity of the masticated food against the roof of the mouth, a swallowing reflex that involves multiple coordinated actions of both the GI and respiratory systems is set in motion. As the bolus of food is moved to the back of the mouth, it activates mechanoreceptors that relay sensory signals to the central nervous system. The motor output of this reflex simultaneously relaxes the upper esophageal sphincter, closes the epiglottis over the trachea, and momentarily arrests breathing during expiration.

Once the bolus of food passes the upper esophageal sphincter, distension initiates a primary peristaltic wave of contraction of esophageal muscle that moves the bolus from the upper toward the lower esophageal sphincter and stomach. As the peristaltic wave moves the bolus toward the latter sphincter, it relaxes to allow passage of the food into the stomach, then closes to prevent reflux of gastric contents into the esophagus. If the primary peristaltic wave fails to move the bolus into the stomach, a secondary peristaltic wave is generated that will complete the job. When the bolus of food enters the stomach and joins other ingested components of the current meal, gastric contractions and movements churn and mix the meal creating a semisolid, partially digested substance called *chyme*.

One of the early mechanical actions of the GI tract beyond mastication and swallowing is the generation of peristalsis. This is a coordinated reflex movement of a wave of contraction and relaxation that propels a bolus of food from one location in the GI tract to the next. Consider a bolus of food that is located midway down the esophagus. In order to keep the bolus moving, esophageal smooth muscle immediately upstream to the bolus contracts reflexively as similar muscle immediately downstream relaxes. This creates an upstream-to-downstream pressure gradient in the lumen of the esophagus that propels the bolus forward. This pressure-gradient-like wave moves from the oral to the aboral end of the esophagus ensuring that the bolus does not stop until it reaches the stomach.

For a few hours after a meal is ingested, peristaltic waves move across the stomach in an oral-to-aboral direction. Less frequently, other waves of peristalsis move from the antrum of the stomach toward the esophagus in an aboral-to-oral direction, causing a retropulsion or reverse flow of chyme. On occasion, the forward flow, or antegrade propulsion of chyme, and the retropulsion meet head-on

and an intragastric collision occurs. This is good. The collision causes further churning and mixing of gastric contents thus maximizing exposure of nutrients in the meal to digestive enzymes, acids, and mucus-containing aqueous medium. At other times, forward flow of chyme is sufficiently powerful to force the pyloric sphincter to open. Then chyme is moved from the stomach into the duodenum, the first segment of the small intestine, in a process called gastric emptying. Gastric emptying can be relatively fast or quite slow depending, in part, on the contents of the meal. The greater the fat content, the higher the osmolarity, and the larger the particle sizes, the slower the emptying process. Conversely, all liquid meals are emptied in haste. The pyloric sphincter does not allow particles larger than about two millimeters in diameter to leave the stomach.

Activities of the digestive tract depend to a large extent on the fed versus the fasted state. The period of ingestion digestion and absorption of a meal is called the absorptive state. The semi-fasting and fasting periods between meals are called postabsorptive states. Physiologists use these two terms to describe functions of the gastrointestinal and endocrine systems in health and disease. In the fasting or postabsorptive state, the small intestine is rather quiet. However, quiescence is periodically interrupted by synchronized rhythmic changes in both the electrical and mechanical activity of the gut wall. These rhythmic postabsorptive contractions are called migrating motor complexes (MMC). They consist of several distinct phases, originate in the stomach, and usually terminate in the ileum. MMCs occur at intervals of 90 to 120 minutes, and they achieve net forward movement of residual chyme that exists from a previous meal. Eating a meal terminates the MMCs, and the specific mechanisms governing them are unclear. However, there is good evidence that both the enteric and autonomic nervous systems, as well as hormonal factors, are involved. One of the most important elements is the hormone motilin. It is produced by the duodenum during the MCC, and its appearance and concentrations in the circulating blood coincide with the 90–120 minute intervals that characterize the MMC.

In the healthy individual and speaking generally, there is a high degree of order in the reflex functions of the GI system. In people who eat three meals per day beginning at six or seven A.M. and ending at six or seven P.M., it is common to have a bowel movement about each twenty-four hours. This kind of physiological regularity seldom is seen in those whose eating patterns differ markedly from this. Even for those who follow the twelve-hour meal schedule, any change of routine can have a marked effect on one's regularity.

Bowel movements occur as the result of defecation reflexes. Some of the various reflexes that contribute to bowel movements are duodenocolic (originates in the duodenum and acts at the colon), gastrocolic (originates in the stomach and acts on the colon), gastroileal (originates in the stomach, terminates in the ileum), and enterogastric (originates at many points in the GI tract, terminates at the stomach). Stomach distension following a meal is a common

stimulus for the gastrocolic reflex. The urge to move one's bowels can appear as early as a few minutes to one or two hours following ingestion of the morning meal. Moreover, the composition of our meals can have a pronounced effect on both the timing and the quality of a bowel movement. Diets high in insoluble fiber including vegetables and whole grains such as wheat tend to produce more frequent bowel movements. Moreover, on such diets the stools are commonly bulkier and softer. The converse is true for diets that lack insoluble fiber. Constipation, especially if it occurs regularly, is a disease of the GI tract, and it can predispose the system to more serious diseases like diverticulitis, steatorrhea, and even cancer of the colon.

Ingesting dairy products for those who are lactose intolerant or lactase deficient, or consuming unwashed or contaminated foods containing salmonella or *E. coli* can also stimulate GI reflexes. These reflexes can be explosive and are designed to rid the body of the foreign toxins. Such toxins and the reflexes they evoke can cause serious, even life-threatening disturbances in the homeostasis of water, electrolyte, and acid/base balances. Farmers in third-world countries still fertilize their fruits and vegetables with human feces, which contain *E. coli*. Anyone who has made the mistake of eating a piece of fruit in such a country without first washing it has probably experienced these kinds of reflexes.

Secretions of the GI Tract

All organs and glands associated with the GI tract have secretory capacities. Whether they secrete primarily water and mucus or enzymes and buffers, the secretions have important physiological consequences. Secretions of the digestive tract are under local, reflex, and central homeostatic control. For example, excess acid in a meal acts locally on the duodenal wall to cause release of the hormone secretin. Secretin is absorbed and transported to the pancreas where it stimulates release of bicarbonate. Bicarbonate in turn buffers excess acid in the duodenum. When the acid is buffered, the signal for secretin release is removed and the system returns to its previous steady state.

The primary secretory organs of the mammalian gastrointestinal tract include salivary glands, the stomach, the pancreas, the small intestine, and the gallbladder and liver. Collectively in the adult human, these organs secrete about eight to ten liters of water per day. Water is the universal solvent in which most of the body's organic and inorganic ions and molecules are dissolved. Whatever volume of water the GI tract secretes, it must also absorb. In addition to the eight to ten liters of secretory GI products, the average adult consumes about two liters of water per day. Thus this system must absorb a total of ten to twelve liters of water each day.

One of the most profuse secretory structures of the GI tract is the individual mucous or goblet cell. There are probably billions of these epithelial cells

TABLE 7.2

Properties of mucus that enable it to protect and lubricate

Adherent qualities	adheres to food particles, thin coating of mucosa
Lumen of gut	coats gut wall and prevents direct contact of mucosa with food
Low slippage resistance	foods slide over surface of mucosa with ease
Fecal adherent	binds food matter so it can be eliminated in stool
Resists decomposition	prevents digestive enzymes from destroying it
Amphoteric properties	can buffer small quantities of both acids and bases

distributed throughout the GI mucosa. They have one thing in common: they secrete mucus. Mucus contains mostly water and electrolytes, but it is also composed of macromolecules such as polysaccharides and glycoproteins. The primary functions of mucus are to protect the integrity of the mucosa and to lubricate food. The qualities that enable mucus to perform these functions are summarized in table 7.2.

The mechanisms involved in the secretory activity of the GI cells and glands are under the influence of the PSNS, SNS, and GI hormones. The volume of meal at any point in time and in any given segment of the GI system, as well as the chemical composition of the meal in that location, are basic to activating these mechanisms. For example, stimulating the parasympathetic nerves to the gut wall uniformly increases secretory behavior. The extent of this action depends on the preexisting level of parasympathetic and hormonal stimulation. This is especially true of the salivary glands, the esophagus, the stomach, the proximal segments of the small intestine, and the distal segments of the colon. Conversely, stimulating sympathetic nerves in the gut wall only modestly increases secretory activity. In addition, sympathetic stimulation reduces GI blood flow.

Since all secretory products ultimately come from the circulation, any alterations in blood flow are likely to have important consequences on the secretory behavior of the GI tract. Although many details are still lacking, table 7.3 provides a sequential list of the steps physiologists believe are involved in secretion. This list is based on experimental data and provides our clearest understanding to date.

Salivary Secretions

Daily output of the three pairs of salivary glands in adult humans averages about one liter. Two kinds of secretion are produced. One is mostly composed of water and electrolytes and is called serous fluid. The second contains serous secretions, mucus, and some important organic molecules, most notably ptyalin,

TABLE 7.3

Sequence of events leading to glandular secretions in the GI tract

1. Building materials must be supplied by the circulation at the base of mucosal cells
2. Mitochondria at the base of mucosal cells make ATP as a source of high energy phosphates
3. ATP and circulating nutrients are used to make the organic secretary substances
4. Secretory substances are transferred from the endoplasmic reticulum (ER) to the Golgi complex
5. Secretory vesicles move through cellular cytoplasm to the apical surface of the mucosal cell
6. Stimuli cause fusion of membranes (cell and vesicle) and release of product

an amylase enzyme that helps digest carbohydrates. Secretion of saliva is not only important to the partial digestion of carbohydrate, but it also lubricates, protects, and aids in oral hygiene. The flow of saliva helps to wash out pathogenic bacteria that inhabit the mouth, thus helping to minimize diseases there. Under basal waking conditions, the production and flow of saliva is greater than it is during sleep. Because salivary flow is lower during sleeping hours, the possibility of bacterial accumulation in the mouth increases. This is one reason why it is important to brush teeth and gargle, clean, and rinse the mouth thoroughly before going to bed. Saliva also contains lysozymes and thiocyanate ions, naturally occurring bactericidals. In the absence of salivation, oral disease including dental caries can become rampant and the mouth ulcerated and frequently infected.

Esophageal Secretions

Secretions of the esophagus are not as intensely investigated as secretions of other sections of the GI tract. Esophageal secretions are composed entirely of mucus, that is, the glands in the esophagus are mainly mucous glands. Those in the proximal or upper portion of the organ serve mainly to lubricate food and to coat the surface epithelium. This protects the mucosal surface from damage by newly swallowed food. The mucous glands in the distal or lower portion of the esophagus are more complex but still only secrete mucus. Here the mucus coating exists primarily to protect the epithelium from potential damage due to gastric reflux. Gastric reflux and its connection to esophageal and supra-esophageal erosion and disease have increasingly attracted the attention of the medical establishment in recent years.

Gastric Secretions

Secretions of the stomach are more complex than either salivary or esophageal secretions. There are multiple kinds of secretory glands spread throughout the gastric mucosa. Some are as simple as surface epithelial cells that primarily secrete mucus. Others exist in the form of pits—also called gastric pits or gastric crypts—that invaginate deep into the submucosa. The walls of these pits are lined with a variety of cell types, each designed to perform a different function. From proximal to distal ends, the stomach can be subdivided according to the character of its mucous lining and the glands that are present. The fundus is the most proximal section of the stomach. Near its junction with the lower esophageal sphincter is a region of specialized tissue called the cardia. This tissue acts much like a cardiac pacemaker in the heart in generating periodic depolarizations of the stomach and corresponding waves of contraction. The body or corpus is the main portion of the stomach. It is defined by greater and lesser curvatures and these meet at the third section or antrum. Immediately distal to the antrum is the pylorus, which connects via the pyloric sphincter to the duodenum.

The primary secretory portions of the gastric mucosa are found in the corpus and antrum. These cells line the walls of the gastric pits. There are at least six different cells types. From the surface of the pit to its base, these are (1) superficial epithelial cells, (2) mucous neck cells, (3) regenerative stem cells, (4) parietal or oxyntic cells, (5) chief cells, and (6) endocrine cells. The main secretory products of the corpus are acid, pepsinogens that aid in the digestion of proteins, and intrinsic factor, a glycoprotein molecule that is important in the absorption of vitamin B_{12} and in the formation of red blood cells. Parietal cells secrete acid and intrinsic factor. Chief cells secrete pepsinogens. In addition, gastric crypts contain mucus-secreting cells in the neck of the pit and endocrine cells that secrete histamine.

Glands in the stomach antrum do not contain parietal cells. Therefore the antrum does not secrete acid or intrinsic factor. In addition to chief cells that secrete pepsinogens, the antrum contains superficial epithelial cells that secrete bicarbonate ions that help neutralize gastric acid. The antrum also secretes gastrin and somatostatin. Both gastrin and somatostatin are hormones that have endocrine and paracrine functions in the GI tract. Gastrin release augments secretion of acid. Somatostatin release inhibits production and secretion of acid and gastrin. Both have other important GI functions including promoting growth and differentiation of other gastric epithelial cells.

Pancreatic Secretions

The pancreas has both endocrine and exocrine secretory properties. As we learned in chapter 3, an endocrine gland is one whose secretory products are released into the general systemic circulation. Blood becomes the vehicle of transport to get these molecules to their target organs. An exocrine gland is one

that does not use the circulatory system to deliver its products. Rather, exocrine molecules diffusing through the extracellular spaces are delivered directly to the target tissues. Examples of pancreatic endocrine products are insulin and glucagon, both of which are important to the metabolism of carbohydrates. Examples of exocrine products are bicarbonate ions and pancreatic lipases. Lipases are enzymes that digest dietary fats.

Other Secretions

Other secretory organs of the GI tract include the duodenum and proximal portions of the jejunum. They release secretin, cholecystokinin (CCK), gastric inhibitory peptide (GIP), and vasoactive intestinal peptide (VIP). All can inhibit the secretion of gastric acid, but none has been definitively proven to be the most important. The liver secretes bile acids that are stored during the postabsorptive state in the gallbladder. The gallbladder releases these bile acids during the absorptive state in response to fat in the diet. Bile acids are essential to the physiological processes of emulsification and digestion of fat.

Digestion of Carbohydrates, Fats, and Proteins

Digestion begins with the sight, sound, and smell of food in preparation. Mechanistically, gastrointestinal activities, including digestion and secretion, can be divided into three distinct phases: cephalic, gastric, and intestinal. The sight, sound, and smell of food constitute sensory stimuli that activate electrical, mechanical, and secretory events in the GI tract during the cephalic phase. Once a meal enters either the stomach or the intestine, the second (gastric), and third (intestinal) phases are implemented.

Because the mucosal epithelium is not prepared to absorb food materials as they are ingested, food products must be broken down into smaller units that can be processed by the epithelial cells. This is the function of digestion. While some digestion begins in the mouth and stomach, the vast majority takes place in the small intestine. For those proteins and carbohydrates that escape digestion in the small intestine, colonic bacteria are able to convert them to short-chain fatty acids. Therefore there is some but minimal digestion and absorption in the colon.

In addition to the three phases of GI function, five options for the digestion and absorption of a meal can be defined: (1) No digestion, such as occurs with simple sugars like glucose. Such molecules need no further processing in order to be absorbed, so they are ingested intact. (2) Luminal digestion of large polymers to smaller monomers occurs whenever an ingested nutrient requires enzymatic breakdown inside the lumen of the gut. Examples include the breakdown of proteins to their individual amino acids. (3) Brush-border digestion involves any ingested nutrient that is acted on by enzymes that are integral components

of the apical membranes of mucosal epithelial cells. An example would be the enzyme sucrase that binds sucrose (table sugar) and reduces it to glucose and fructose, its component monosaccharides. (4) Intracellular hydrolysis occurs when di- and tripeptides, for example, are absorbed by the mucosal epithelium, broken down to individual amino acids inside the cell, then absorbed into the bloodstream. (5) Luminal hydrolysis followed by intracellular resynthesis is best illustrated with dietary lipids. The main fat ingredient in our diets is triglyceride. In the lumen of the gut, triglyceride is broken down to glycerol molecules and free fatty acids. These are absorbed by the mucosal epithelium, then reassembled in the form of intracellular chylomicrons. Chylomicrons are absorbed by lacteals (intercellular lymphatic vessels) before they are finally delivered to the systemic circulation.

In the processes of digestion and absorption, surface area is key. The greater the surface area of a particular ingested macromolecule, the greater the probability that digestive enzymes will have access to it. Similarly, in the case of brush-border digestion, the greater the surface area of membrane-bound, brush-border-containing enzyme, the greater the likelihood that ingested energy substrates will make contact with the enzymes. Physiologically this is one of the main purposes achieved by the churning, propelling, and invaginating mechanical actions of the gut wall; they maximize the statistical probabilities that any given food particle will come in contact with essential enzymes in the lumen or at the surface of the mucosa. However, most mixing and churning occurs in or near the central axis of the lumen of the gut. Thus, there is a region a few micrometers in width and adjacent to the surface of the microvilli brush border where mixing and churning do not take place. Physiologists call this region the unstirred layer.

Carbohydrate Digestion and Absorption

Most dietary carbohydrates are consumed in the form of sucrose (from sugar), lactose (from milk), and starches (from plant sources). There are many other carbohydrates in the diet but they are consumed to a much lesser extent. Cellulose, or plant sugar, is also present but there are no enzymes secreted by the human digestive tract to process it. The first enzyme that any ingested carbohydrate is likely to meet is α-amylase—alpha amylase, previously known as ptyalin—a salivary amylase or carbohydrate-digesting enzyme. It is secreted by the salivary glands in response to dietary carbohydrate, therefore it begins to digest carbohydrate in the mouth and stomach. This enzyme remains active in the stomach only until the contents are thoroughly mixed with gastric acid, which inactivates the enzyme. There is then no further digestion of carbohydrates until the chyme reaches the duodenum. Pancreatic α-amylase is much more active than salivary α-amylase, and within a few minutes after arriving in the duodenum, dietary carbohydrate is mostly reduced by this enzyme to small glucose-containing polymers

such as maltose, maltotriose, α-1,4 linked maltooligosaccharides, and α-limit dextrins (five to nine glucose units long).

Glucose is a monosaccharide, and most ingested carbohydrates must be in the form of a monosaccharide before they can be absorbed by the small intestine. A polymer of glucose is a molecule composed of many glucose monomers. Some dietary glucose polymers consist of thousands or even millions of glucose monomers. Other dietary carbohydrates are ingested as simple sugars. For example, sucrose or table sugar is a disaccharide composed of one molecule of glucose and one molecule of fructose. Fructose is a monosaccharide found in many sweet fruits and honey. Lactose, a disaccharide, is composed of one molecule of glucose and one of galactose. It is prevalent in mammalian milk and is used to manufacture infant milk formula. Maltose is a disaccharide composed of two glucose monomers and is the fundamental building block of glycogen, a common animal dietary starch and the principal storage form of glucose in the liver of mammals.

In addition to the salivary and pancreatic amylases that participate in luminal digestion, other carbohydrate-digesting enzymes are bound to the membranes of mucosal cells and participate in membrane or brush-border digestion. Thus, further digestion of disaccharides and oligosaccharides to monosaccharides takes place at the brush border of the duodenal and jejunal microvilli. The major brush-border oligosaccharidases are lactase, sucrase, isomaltase (also called α-dextrinase), and maltase (also called glucoamylase). While these enzyme proteins are integral parts of the mucosal epithelial cell membranes, their active sites face the gut lumen, so they are called ectoenzymes. Their active sites do not reside in the interior of the cell. When lactose meets lactase, lactose is split into one monomer of glucose and another of galactose. Likewise, sucrose is split into glucose and fructose. Thus, the final products of carbohydrate digestion are all monosaccharides with glucose constituting more than 80 percent of the final products of carbohydrate digestion and galactose and fructose less than 10 percent each.

Protein Digestion and Absorption

Less than 20 percent of the protein content of an ingested meal is digested in the stomach. The rest takes place in the duodenum and upper jejunum. The parietal cells of the stomach wall secrete hydrochloric acid at a pH well below 1. By the time this mixes with the contents of a meal and other secretory products of the stomach, gastric pH is elevated to 2 or 3. Chief cells of the stomach wall secrete a protease called pepsinogen. This is inactive until converted by gastric acid to pepsin, a powerful proteolytic enzyme. The pH optimum of pepsin is less than 5, so the gastric juices provide a nice environment for initiation of the digestion of dietary protein. If the pH of stomach acids rises above about 5, conversion of pepsinogen to pepsin, hence the proteolytic actions of pepsin, is

halted. The gastric products of protein digestion are amino acids and small peptides, for example, proteoses, peptones, and polypeptides. One of the proteins that is abundant in meats is collagen. It constitutes an important part of the tough connective tissue that binds other protein elements in meat. It must be broken down in order for the proteolytic enzymes of the pancreas to be effective. Pepsin is one of the few naturally occurring proteases known to do this. People who lack pepsin in the stomach have a more difficult time digesting meat products.

Immediately upon entering the small intestine, whole proteins and the products of gastric digestion of proteins are attacked by pancreatic proteases. Most important among these are trypsin, chymotrypsin, carboxypeptidases (also called carboxypolypeptidase), and elastase (known also as proelastase). These pancreatic enzymes are all stored and secreted in the form of inactive enzymes. This prevents the autolysis of storage vesicles and other cellular components of the pancreas. Once inside the duodenum and jejunum and in the presence of digested proteins, these inactive proteases come in contact with brush-border proteases that activate them. In addition, the brush-border enzymes split partially digested polypeptides into di- and tripeptides and individual amino acids, all of which can be easily absorbed by the enterocytes of the mucosal brush border. Again, there are two sets of enzymes involved in the digestion of proteins and protein products; luminal and membrane-bound proteins.

Fat Digestion and Absorption

Of the ingested energy substrates found in food, fats or lipids are the most complex for the digestive tract to process. Most dietary lipids that humans consume are eaten in the form of triglycerides. A triglyceride is composed of one molecule of glycerol and three molecules of fatty acids attached to it: either short, medium, or long chain fatty acids. These create challenges to the GI tract from the moment they are ingested. Because of their low density and aqueous insolubility, ingested fats and triglycerides settle at the top of the chyme and are usually the last dietary products to be emptied from the stomach.

Lipids are released into the duodenum in the form of fat globules, some large enough to be visible to the unaided eye. Because of their density and water insolubility, the individual fat products inside these globules are not accessible to the lipolytic enzymes. Fat globules must be emulsified to make them available to lipases for further digestion. Emulsification is a complex chemical process involving bile acids, geometry, and the solubilities of fat products in water and lipid. Bile acids are produced by the liver and stored in the gallbladder. Basically, bile acids interpose themselves into the fat globules thereby breaking the globules into smaller fragments. The constituents of these smaller fragments, including cholesterol, free fatty acids, triglycerides, phospholipids, and cholesteryl esters, are assembled with the bile acids into elements called

micelles. The micelles act as transport vehicles to deliver the individual fat products to the unstirred layer. The contents of the micelles of the stirred chyme are in equilibrium with the concentrations of free fatty acids, cholesterol, and phospholipids in the unstirred layer at the brush border. As brush-border lipid products are absorbed and their unstirred layer concentrations decrease, micelles release additional products to maintain the equilibrium.

During the processes of emulsification and micellar transport, bile acids are continuously released into the duodenum. Many of these are absorbed by the brush borders and recirculated to the liver, gallbladder, and upper GI tract to continue participating in the processes of emulsification and fat transport. This recycling process is called the enterohepatic recirculation of bile. The same bile acid can be recirculated several times during the digestion of a single fatty meal.

Once the products of emulsification and lipid digestion are inside the mucosal epithelial cells, lipid digestion continues before absorption takes place. In the enterocytes, free cholesterol, free fatty acids, phospholipids, cholesteryl esters, and mono-, di-, and triglycerides are reassembled into a globular product called a chylomicron. These are too large to be released into the systemic venous circulation via the mesenteric capillaries. Therefore, they are absorbed by lacteals—tiny lymphatic vessels that project into each microvillus—delivered to the general lymphatic circulatory system, and subsequently drained into the venous side of the systemic circulatory system in the large blood vessels of the neck. Chylomicrons are then broken down and absorbed by the liver and other tissues of the body. Chylomicrons and their by-products are subdivided into six classes according to their metabolism and functions in the body. These classes are (1) chylomicrons, (2) chylomicron remnants and the lipoproteins, (3) high-density lipoproteins (HDL, or good cholesterol), (4) intermediate-density lipoproteins (IDL), (5) low-density lipoproteins (LDL, or bad cholesterol), and (6) very-low-density lipoproteins (VLDL).

Health of the Oral Cavity and Vocalization

As mentioned above, the GI tract begins with the oral and nasal passages and ends at the anus. Arguably, it is one of the most important ports of entry into the human body. The foods we eat, the water we drink, the air we breathe, and the pathogens these carry, can all gain access to our internal milieu through the mouth and nose.

About three lectures before the end of spring semester in 2003, I began to lose my voice. I assumed this was a simple case of laryngitis. I basically whispered through those last three lectures and then forgot the ordeal. When the semester ended, I called my doctor and made an appointment. We surmised seasonal allergies and postnasal drip as probable causes of my weakened voice, but the antihistamines and decongestants he prescribed had little or no effect.

Neither did proton pump inhibitors (to check possible gastric reflux) have a beneficial effect; moreover, they caused periodic heart palpitations, so my doctor referred me to an ear, nose, and throat specialist (ENT) and to a gastroenterologist. Other than determining that my vocal cords had no visible tumors and that there were no polyps in my nasal passages—by use of otolaryngoscopy by the ENT—their treatment recommendations were of limited value. We also considered speech therapy, but waiting lists in central New Jersey were one year and growing. Friends I knew with similar voice problems found Botox injections to be of little long-term benefit.

I scrutinized the notes I had taken in my treatment and noticed that the ENT had written in his diagnosis voice abuse. This forced me to acknowledge that I had a persistent habit of excessive and sometimes rigorous throat clearing, especially after ingesting a meal. My habit extended as far back as I could remember. I also began paying closer attention to my habits of vocal communication. I noticed my predisposition to excitement as I spoke of subjects of interest to me. So I tried to control this. I examined my swallowing behavior and noted its relation to timing of meals, content of meals, and to the associated otopharyngeal secretions. These secretions seemed to be the most problematic when I ate dairy products or products with a high sugar content late in the evenings. I realized that I had probably been abusing my voice and vocal cords without even knowing it.

By winter 2006 I had made improvement by treating myself, but I was still experiencing weak vocalization, sore throats, and occasional swollen salivary glands. In early March my primary care physician prescribed azithromycin, an active antibiotic. Shortly thereafter, I felt some improvement. After a week or ten days the soreness and inflammation reappeared, and my doctor prescribed Levaquin. By early May 2006, the rawness in my throat was gone, my glands were not swollen, and my voice seemed to be stabilizing and was stronger. I continue to practice self-prescribed changes in vocal behavior, but the volume and strength of my voice have not reliably returned to pre-2003 levels. I have concluded that the changes are due, among other things, to aging, my teaching career, and longevity of unintentional vocal abuse.

I tell the above story as a backdrop to Barrett's disease or cancer of the upper esophagus and throat. This is a disease that in males has been on the rise in recent years. Nobody knows the cause, but some think it is due to gastric reflux. Reflux of gastric acid, if untreated and prolonged, can cause damage to most soft tissues. Moreover, despite a gastric acid pH that is incompatible with most life, some bacteria are able to colonize the stomach and to even reproduce and thrive there. During the postabsorptive state, stomach pH ranges from 4 to 6. When the HCl-secreting glands are maximally stimulated, as with administration of gastrin or histamine, the pH of gastric contents can be as low as 2. Until the mid-1980s, conventional medical wisdom held that while many bacteria can

and do transit the stomach, none could take up permanent residence, reproduce, and multiply there. That long-held medical dogma has been abandoned since the discovery of the gram-negative, rod-shaped bacterium *Helicobacter pylori* in the human stomach.

The winners of the 2005 Nobel Prize in physiology, Drs. Barry Marshall and Robin Warren, discovered *Helicobacter pylori* in the stomachs of patients in 1984. Consequently, the attention of many in the fields of bacteriology, pathology, physiology, and medicine has been redirected to the examination of the pivotal role of *H. pylori* in the pathogenesis of gastritis and peptic ulcers. The notion that a bacterial infection could set the stage for peptic ulcers, gastritis, and even gastric carcinoma flew in the face of long-standing medical wisdom. Under physiological conditions, there is a gelatinous mucous-like barrier that lines the wall of the stomach and other components of the gastrointestinal tract. This barrier prevents or minimizes the underlying epithelial secretory cells from damage by gastric acid and other potentially harmful agents. *H. pylori*, by the process of acid acclimation, is able to thrive in the stomach by buffering gastric acid once it crosses the outer wall of the bacterium. If the bacterium resides in the stomach or in any other upstream or downstream portion of the GI tract, it can damage or destroy the protective barrier and thus cause gastritis, ulcers, and even cancer.

Almost four decades ago, a physiologist named Horace Davenport proposed a mechanism by which gastric acid, having penetrated the mucosal diffusion barrier, can damage the gastric mucosa. The details of Davenport's proposal have been modified in more recent years, but the underlying basis of his proposal has not changed. Hydrochloric acid, upon contact with the gastric mucosa, damages endocrine-like mast cells thus causing the release of histamine and other mediators of inflammation such as leukotrienes, oxidants, thromboxanes, endothelins, and platelet activating factor. Collectively these compounds reduce mucosal blood flow thereby causing ischemia. Tissue ischemia causes death of the tissue.

Health of the GI Tract and Diet

Good dietary practices and a healthy GI system are inseparable. One of the best things to do for a healthy diet is to follow recently published nutritional guidelines. These include eating a variety of foods; making whole grains, particularly wheat-based products, the foundation of your diet; and eating generous portions of vegetables and fruits of a variety of kinds and colors. To the extent possible, one should replace all beverages, including soft drinks, alcohol, coffee, and tea, with tap water. Remember all such beverages are more than 90 percent water by weight and volume. The remaining 10 percent is mostly weight-enhancing sugars, electrolytes and ions that are readily available in tap water, and unidentified chemicals and toxins.

In addition to controlling your diet, another simple procedure for monitoring GI health is to take note of the frequency of bowel movements and the consistency and composition of the stools. If your diet consists of whole grains, vegetables, and fruits, supplemented with generous quantities of tap water, you should probably have bowel movements on a daily basis. Moreover, with this kind of diet and a well-established daily routine, then your bowel movements will likely occur at nearly the same time every day. If you have more than one bowel movement per day, you are quite possibly overeating. Conversely, even if you practice a healthy diet but lack a daily routine, your inter-stool-interval (ISI) — that is, the hours, days, and weeks between bowel movements—will likely be irregular as well. If you are constipated or have diarrhea on a frequent basis, then you probably lack both a healthy diet and a healthy lifestyle.

Monitoring transit time is another way of measuring the health of your GI tract. Transit time in the gut is the amount of time—minutes, days, and hours—it takes for a meal to pass through the GI tract, or the elapsed time from completion of a meal to elimination of the residue of that meal. The longer the transit time, the more unhealthy the GI system is likely to be; the shorter the transit time, the healthier it is. This is because some of the contents of each meal get temporarily stored in the colon as bacteria and other pathogens. The longer these pathogens remain in the colon, the more likely they are to cause diseases there and elsewhere. A transit time of twenty-four hours or less is optimum.

One can monitor his own transit time as follows. Without changing your diet or lifestyle, make a record of your ISI for several days or for a week or two. Compute the average in hours and minutes. These data will tell you what your basal ISI is. Let's assume that your average ISI is forty-eight hours and thirty minutes. Imagine that for the next four weeks as a measure to improve your GI health, you decide to increase your consumption of insoluble fiber from fifteen grams per day to thirty grams per day. You have chosen to do this, in part, by eating more wheat-based food products such as raisin bran cereal for breakfast and whole wheat bread for sandwiches. After about one month on this diet and having changed nothing else in your lifestyle, you again make a record of your ISI for several days or a couple weeks. You compute the average, and this time discover that it is twenty-six hours and fifteen minutes, a substantial reduction from forty-eight hours and thirty minutes.

If this experiment was conducted in large numbers of people and if similar results were obtained, statistical analysis could be used to establish whether the difference between forty-eight hours and thirty minutes and twenty-six hours and fifteen minutes was statistically significant. If so, the investigators would conclude that doubling one's daily consumption of whole-grain food products causes a significant reduction in ISI. An extension of this conclusion would be that colonic bacteria reside in the lower GI tract for a shorter period of time and

therefore are less likely to cause GI-related problems. That is, the health of one's GI tract has improved by having more whole wheat in the diet.

In order to successfully complete the experiment described above, one must use markers to accurately determine how much time a meal spends in the GI tract. One way of marking the stool is by eating a food product that consistently and reliably colors the feces. There are commercially available dyes for such purposes. These are employed by physiologists, nutritionists, and dieticians in metabolic kitchens. Metabolic kitchens are laboratories where experimental subjects are monitored by scientists. Such labs are not needed to complete the simple experiment I have mentioned above. There are plenty of safe food products readily available in the supermarket that do not affect transit time but do color the stool. These are too numerous to mention, but a short list includes canned spinach, beets, black licorice, and baked brownies. The first two give the stool a dark forest green hue. Brownies obviously color the stool brown. If you are observant, you have noticed other foods that predictably give your stools one or another color. To mark the stool, simply note the day, time, and color of the latest bowel movement. Then eat a moderate quantity of black licorice, such as one hundred grams, and note the time and day. Thereafter, observe the next few stools and determine when the distinctive color first appears. Then data can be collected for the experiment just described.

8

The Reproductive System

The human reproductive system consists of internal and external organs that help identify one's phenotype or degree of maleness or femaleness. The reproductive system is a complex organ system that begins to develop and differentiate early after conception. There are both physical and physiological differences between male and female genders. Males tend to have thick facial hair while women tend to develop breasts. But they are only subtly different or even the same at times in human development and maturation. For example, in the first several weeks postfertilization, the gonads are not yet sexually determined and can potentially develop as either male or female. There are also aspects of reproductive physiology that are much more complex. Answers to questions such as why women experience menopause midway through life while men are able to produce sperm cells to the end of normal life have evaded scientists since the phenomena were first discovered. While reproduction includes physiological concepts such as fertilization, implantation, placentation, parturition, and lactation, and the advanced student and clinician should be familiar with them, only the dynamic physiology of gametogenesis will be discussed here.

Gametogenesis, Genetic, Gonadal, and Phenotypic Sex

Gametes are germ or reproductive cells. The male gamete is the sperm cell and the female gamete is the egg or ovum. Gametogenesis is the process of differentiation, development, and maturation of the germ cells. The dynamic changes both in structure and function of the developing and differentiating gametes is called spermatogenesis in males and oogenesis in females. The mature end products, in males a spermatozoan and in females a graafian follicle, bear little resemblance to their primitive progenitor cells called spermatogonia in males and oogonia in females. Gametogenesis can be divided into the study of male gametes (spermatogenesis) and female gametes (oogenesis).

Determination and differentiation of the genders can be divided into three components: the genetic sex (genotype), the gonadal sex (whether the indifferent embryo develops testes or ovaries), and the phenotypic sex (the physiological/morphological gender, or phenotype). When a sperm cell penetrates and fertilizes an ovum, the new cell is called a zygote. Maleness is determined by the presence of a Y chromosome from the paternal sperm cell. An X chromosome contributed by the father will yield two X chromosomes in the zygote or a genetic female organism. Thus, genetic sex is determined upon fertilization.

For the first several weeks postfertilization, the developing embryo is bipotent, that is, it can become either female or male. Once the mechanisms that determine which gonadal system will develop have been initiated, the gonadal sex of the embryo is set. This happens at about five to eight weeks in the genetic male and several weeks later in the genetic female. During puberty, properly functioning gonads will assist in the development of internal reproductive systems including auxiliary organs such as the prostate gland in males and the endometrium in females. The gonads are also critical to the development of external genitalia. The combination of internal reproductive organs and external genitalia coupled with body morphology set the phenotype or physiological sex of the individual.

In order for a zygote to be produced, copulation and ovulation must be coincidental. There is a relatively narrow window of opportunity lasting approximately twenty-four hours during which a viable sperm and egg must meet. Fertilization of the ovum occurs in the fallopian tube or oviduct where the sperm and egg meet during this window of opportunity. At about this time, commingling of nuclear chromosomes from the sperm and the egg define the end of fertilization and the beginning of embryogenesis (development and differentiation of the embryo). The fertilization process is complex and is divided into at least eight stages all of which must appear on time and in sequence for a viable zygote to be produced.

None of the several million sperm that are released in the male ejaculate are capable of fertilizing an egg as ejaculated. They must undergo further preparation inside the fallopian tube and, as they come in contact with the ovum, a process called capacitation occurs. Under normal conditions, only one of these cells will actually penetrate the egg even though many will reach it. This is indeed a case where the race and prize go to the strong. It is survival and success for the fittest. After the sperm cell penetrates and once the male and female pronuclei (nuclear chromosomes) fuse in stage eight of the fertilization process, a new cell called the zygote is transiently formed, the process of fertilization comes to an end, and embryogenesis begins.

At fertilization, the sperm and egg meet, each contributing 23 chromosomes including one sex-determining chromosome and 22 non-sex chromosomes, to create the zygote. At this time only the genetic sex is determined. The female contributes an X chromosome and the male contributes either an X or a

Y chromosome. If the combination is XY, then the genetic sex is male. If the combination is XX, then the genetic sex is female. This we call the genotype or the genetic contribution to sex determination. However, the maleness or femaleness of a developing fetus is called the phenotype, and this is importantly influenced by events that take place after fertilization. The outcome of this process used to be called the physiological or morphological sex but more recently has become known as sexual differentiation.

During the first three days postfertilization and while still inside the fallopian tube, the embryo develops to a solid mass of twelve or more cells called the morula or conceptus. Strong contractions of the isthmus of the oviduct (the distal end near the uterus) during these three days prevent the developing embryo from entering the uterus. This delay helps facilitate further maturation of the uterine wall in preparation for implantation of the embryo. After the morula moves through the isthmus and into the uterine cavity, it continues to develop to become the blastocyst. The blastocyst is a spherical mass of cells with a fluid-filled inner cavity. Surrounding this cavity is a layer of specialized cells which develop to become the amniotic fluid, the yolk sac, and the fetal portion of the placenta. On one side of the cavity is another cell mass that later becomes the embryo proper. The blastocyst continues to float freely inside the uterus for another three days before it attaches to the endometrium or inner lining of the uterine wall. In other words, implantation of the human embryo takes place approximately six or seven days after ovulation and fertilization.

In the field of reproductive physiology, the biological application of mathematical combinations and permutations finds its niche. The genetic principles of chromosomal crossover, exchange, and recombination collectively determine the genotype and the phenotype of the resultant individual. Not only is her gender determined, but also the color of her eyes, the tone of her skin, and whether or not she will have straight or wavy or blond or brunette hair. As we reflect on our individual character and personalities, the outcomes of another sperm fertilizing the same ovum (or of the same sperm fertilizing a different ovum) in place of our conception in the same parents is almost incomprehensible. Adults fortunate enough to have known the members of their family for three, four, or even five generations should be able to appreciate the contributions of genetics and physiology to the phenotypic outcomes of their descendants.

The Sexually Indifferent Embryonic Gonad

From fertilization to about five to eight weeks postfertilization, the developing embryo contains sexually indifferent gonads. This means the primitive, embryonic gonads are indistinguishable, and their genital tracts are not formed. Between this period of indeterminate status and that of adulthood in either gender, the processes of sexual differentiation take place. Gene complexes on

the sex chromosomes determine whether the primary sex organs differentiate into testes or ovaries. The Y chromosome exerts an important testis-determining effect on the indifferent gonad.

Experimentation and clinical observation have shown that the Y chromosome, with few exceptions, is necessary for development of the male testes. Thus, it was reasonable to speculate that the gene(s) responsible for organogenesis of the testes and ductal system was (were) located on the Y chromosome. It turns out that a single gene, the sex-determining region of the short arm of the Y chromosome (SRY gene), encodes for a testis-determining factor (TDF). This particular transcription factor belongs to the eighty-amino-acid high-mobility group (HMG) superfamily of transcription factors (see Suggested Reading). Together, the SRY gene and its TDF transcription factor determine the genetic sex in the early weeks after fertilization. Later, at the onset of and during puberty, one's phenotypic gender will be determined. The critical preadulthood periods of sexual differentiation and development take place between approximately five and eleven weeks postfertilization and again during puberty (about eleven to fifteen years of age) in both genders, on average.

One of the primordial cell lines in the undifferentiated gonad develops into either Sertoli cells of the male testicular seminiferous tubules or into granulosa cells of the female ovarian follicles. These cell lines have homologous functions. They nourish and support developing germ cells, they foster their growth and morphogenesis, and they guide their movements into the genital ductal systems. They become the major source of estrogens, the female steroidal hormones such as estradiol, which are found to a lesser extent in males. Interstitial cells, another primordial cell line in the undifferentiated gonad, give rise to thecal cells of the ovary and Leydig cells of the testis. Interstitial cells function primarily by secreting androgens (male steroidal hormones such as testosterone), which are found to a lesser extent in females. Androgens are biochemical precursors to the synthesis of estrogens and are used in development of sperm and masculine characteristics in the male and to a lesser extent in the female.

The primitive bipotential gonads are called the wolffian system in the genetic male and the mullerian system in the genetic female. The differentiation and subsequent development of one or the other depends entirely on the presence of the Y chromosome. If the SRY gene is present in the embryo, the wolffian reproductive system will differentiate and develop. This system secretes a substance that causes regression of the mullerian ductal system. The substance is variously called the mullerian-inhibiting factor (MIF), mullerian-inhibiting substance (MIS), and/or mullerian-inhibiting hormone (MIH). Whether MIF is under central feedback control is not known, but it is sure elevated concentrations inhibit female tissues and simultaneously stimulate differentiation and development of male tissues. In the absence of the Y chromosome and by default, the wolffian system fails to differentiate, no MIS is produced (the mullerian

system is not actively inhibited), and, at about eleven weeks postfertilization, the mullerian system begins to develop.

Spermatogenesis and the Male Reproductive System

The process of spermatogenesis begins in the seminiferous tubules inside the bilaterally located testes. The word *tubule* implies a structure with a wall and a lumen something like a garden hose. A blood vessel, a fallopian tube, the small intestine, and surgical tubing are other examples. The wall of a seminiferous tubule is composed of Sertoli cells. The lumen contains secretory products and mature/maturing sperm cells. The sites of differentiation and development of male germ cells (primordial germ cells, immature spermatagonia, primary and secondary spermatocytes, spermatids, and mature sperm cells called spermatozoa) are in the walls of the seminiferous tubules and the lumen.

Because of the fragility of developing spermatocytes, they are separated from the interstitium and other extracellular spaces by a basement membrane that is impermeable to foreign substances such as toxic drugs, viruses, bacteria, and other pathogens. This basement membrane is composed of fibrous connective tissue, charged glycoproteins, and other macromolecules that form a blood-testis barrier. Its function is similar to that of the so-called blood-brain barrier of the central nervous system. Outside the wall of the seminiferous tubules are the interstitial spaces (also called interstices or extracellular space). Here are found the Leydig cells, the site of production and release of the male sex hormones the androgens (testosterone, dihydrotestosterone, androstenedione). Under physiological conditions, little protein is found in the interstitium, but other elements of plasma—water and electrolytes—are in abundance.

At puberty (about thirteen to fourteen years of age in the average male) and under the direction of the appropriate neuroendocrine signals (GnRH, gonadotropin-releasing hormone from the hypothalamus; FSH, follicle-stimulating hormone; and LH, leutinizing hormone from the anterior pituitary gland), spermatogonia begin to divide and proliferate mitotically and to differentiate beneath the basement membrane but between adjacent Sertoli cells in the basal compartment of the seminiferous tubules. Spermatogonia are testicular stem cells, and under physiological conditions there appears to be an endless supply of them. Between the basement membrane and the lumen of the tubules, the Sertoli cells are arranged in columns. The sequence of events in spermatogenesis starts with (1) spermatogonia, or undifferentiated stem cells, which lead to (2) primordial germ cells → (3) primary spermatocytes → (4) secondary spermatocytes → (5) spermatids → (6) mature, fully differentiated spermatozoa (mature sperm cells). These events take place between adjacent columns of Sertoli cells in three distinct locations, the basal compartment, the central (adluminal) compartment, and the tubular lumen. Separating the basal and central compartments are tight

junctions between adjacent Sertoli cells. As spermatogonia differentiate and mature and pass from basal to central compartments, new tight junctions form behind and between adjacent Sertoli cells. Simultaneously, old tight junctions break down in front of the developing spermatogonia.

Transformation of primordial germ cells to functional fully developed spermatozoa varies among individuals and can take from sixty-five to eighty-five days. Each stage of spermatogenesis has a specified duration, with primary spermatocytes and spermatids being the longest (about twenty-three days) and secondary spermatocytes being the shortest (about one day). Germ cells must move forward in space as they develop, that is, from basal to central to luminal compartments, otherwise progress can be impaired. If the environment is not favorable to development and differentiation, impaired spermatocytes degenerate and are eliminated by the process of programmed cell death (apoptosis).

Oogenesis, Folliculogenesis, and the Female Reproductive System

In the female, oogenesis begins in the fetal ovary. Primordial germ cells migrate at five to six weeks after fertilization from the yolk sac of the developing embryo to the genital ridge of the hindgut. By weeks six or seven of intrauterine life, a total of approximately ten thousand primitive oogonia are present in the fetal ovaries. To this point in oogenesis, no atresia (a term reserved for use in females but that refers to apoptosis) has occurred. Because of mitosis, several hundred thousand oogonia are present by eight weeks of gestation. The number of female germ cells peaks at an estimated six to ten million by twenty to twenty-four weeks postfertilization. From about eight weeks of gestation through the remainder of the reproductive lifetime of the average woman, three concurrent processes determine the actual number of germ cells in the ovaries: mitosis (rate of proliferation by cell division), meiosis (nonmitotic development and division of the maturing oogonia), and atresia (rate of programmed cell death, about which we know very little). By birth, only one or two million primary oocytes are present, and at menarche, the first menstrual cycle and onset of puberty, the number has been reduced by atresia to about 400,000.

During fetal life, the outer surface of the ovary is covered by a layer of cells called the germinal epithelium. As the female fetus develops, these primordial stem cells differentiate from this layer of tissue and migrate into the underlying ovarian cortex. In the latter location, each ovum is enshrouded with a coat of cells from the stroma of the cortex (the ovarian supportive tissue). The ovum and this single layer of granulosa cells are now called a primordial follicle. The ovum is still sexually immature and requires two meiotic divisions before its nuclear chromatin material will commingle with that of a sperm cell.

Folliculogenesis is another term used to denote continuation of oogenesis and maturation of the primordial follicle. Primordial follicles begin developing

in the female ovary at about twelve to sixteen weeks postfertilization. Development can be divided into three stages. The first stage develops slowly and parallels the genetic prophase of oocyte development. This stage lasts, under normal circumstances, no less than thirteen years (the age of onset of puberty in an average young woman) but can take up to fifty years (near menopause). As the cell begins to divide meiotically, it induces a layer of spindle-shaped cells from the ovarian stroma to completely surround it. These are precursors to granulosa cells. The primordial oocyte, with its shroud of spindle-shaped epithelial cells, is called the primordial follicle and averages about 25 micrometers in diameter. At about twenty to thirty weeks of gestation, the next phase of stage one occurs when the spindle-shaped cells become more cube-shaped to form the initial layer of granulosa cells. When this happens, the primordial follicle is then called the primary follicle. As the granulosa cells divide, multiple layers of granulosa cells surround the oocyte. These cells also secrete mucopolysaccharides, which form a protective layer around the oocyte called the zona pellucida. Granulosa cells adjacent to the zona pellicuda send cytoplasmic projections into the zona thus creating nutrient-supplying conduits to the developing oocyte. Hence, the cytoplasm of the granulosa cells, like that of the seminiferous tubular Sertoli cells in males, forms a blood-granulosa barrier through which constituents of the blood plasma must pass before coming in contact with the developing female germ cell. Neither the granulosa cells nor the developing oocyte have a direct blood supply at this stage of development. Therefore, exchange of all nutrients between the two tissues occurs solely by diffusion.

The multilayered primary follicle continues to enlarge as granulosa cells proliferate, ultimately reaching a diameter of about 150 micrometers. By this time, the oocyte has reached its maximal size of about 75 micrometers. A secondary follicle is developed when granulosa cells begin secreting fluid that forms small pools between adjacent cells. Formation of this vesicular follicle marks the end of stage one and the maximum degree of development in the prepubertal ovary.

Stage two of folliculogenesis in contrast to stage one is more rapid and requires only seventy-five to eighty-five days. It takes place after menarche. During each month's menstrual cycle, a small number of secondary follicles is recruited for further development, a process that spans two to three menstrual cycles. The small pools of follicular fluid accumulate in a single pool called an antrum. The fluid is known as antral or vesicular fluid, so the follicle at this time is called an antral or vesicular follicle. Antral fluid is an aqueous medium containing an array of dissolved solutes, for example, growth factors, proteins, electrolytes, hormones such as FSH and LH, arginine vasopressin, and the renin-angiotensin system of chemicals. Granulosa cells continue to proliferate, displacing the oocyte into a marginal (peripheral) position and onto a stalk. Angiogenesis (the growth of new blood vessels) gives this follicle direct access to

TABLE 8.1

Some differences between spermotagenesis and oogenesis

1. Onset	spermatogenesis occurs at puberty oogenesis occurs during first trimester in utero
2. Production of primordial germ cells	spermatogonia, millions per day oogonia, 8–10 million in a lifetime
3. Male reproductive lifetime female reproductive lifetime	puberty to senility, arguably 80–90 years menarche to menopause 30–35 years
4. Maturation during gametogenesis	sperm requires oviduct and oocyte ovum requires ovulation, oviduct, sperm
5. Selection process	twice during folliculogenesis in females does not exist in males

blood-borne molecules, gases, and nutrients. The entire process of stage two leads to the creation of the mature antral or vesicular follicle.

Stage three is the shortest and fastest of the three stages of folliculogenesis. About one week after the start of menses, a second selection process occurs. One of a cohort of about fifteen to thirty secondary follicles that had been selected to become antral follicles is selected from among its peer follicles to become the dominant follicle. This process occurs monthly and only in one ovary. The volume of antral fluid in the dominant follicle is markedly increased, but intrafollicular pressure remains relatively constant at about 15 to 20 mmHg. This follicle achieves dimensions of two centimeters in diameter about two weeks before ovulation. At the time of ovulation, the follicle merges with the wall of the ovary, and the two undergo temporary proteolysis to facilitate release of the follicle. The follicle is gently released into the peritoneal cavity during ovulation, and the first meiotic division, which was suspended in utero, is completed. The end products of this period of meiosis are the secondary oocyte (which has been drawn into the fallopian tube for fertilization) and the first polar body (the other decomposed daughter cell that will be discarded). In the fallopian tube, fertilization by a sperm cell induces a second meiotic division in the oocyte yielding a reduced number of chromosomes (twenty-three, the haploid number) and another discarded tissue, the second polar body. Ovulation marks the termination of folliculogenesis.

I have summarized some of the fascinating differences between spermatogenesis and oogenesis in table 8.1. These will serve as a springboard from which the interested student can seek more information.

In most people, genetic, gonadal, and phenotypic gender are consistent with male and female designations. However, in some individuals, deviations and

TABLE 8.2

Examples of disorders in genotypes and phenotypes in humans

Name of disorder	*Basic problems*
True hermaphrodites	possess both male and female sex organs; no Y chromosome, 46 XX chromosomes
Pseudohermaphrodites	one type of gonad with morphology of both sexes, transfer of SRY gene to X chromosome
Mixed gonadal dysgenesis	possess testis plus streak ovary, no Y chromosome, 45 XO chromosomes
Pure gonadal dysgenesis	streak gonads but no body features of XO
Congenital adrenal hyperplasia	abnormal genitalia in genotypic females due to adrenal dysfunction
Impaired androgen actions	enzyme deficiencies, failed responses of target tissues to androgens

disease states, though uncommon, can interfere with the above processes. This chapter is not intended to include a taxonomy of reproductive disorders, nor is the beginning student prepared to learn these. However, for the interested student and practitioner, I have included a sample of several conditions in table 8.2. For more detailed information, see the suggested reading at the end of the book.

Sexually Transmitted Diseases (STDs), Sexual Behavior, and Infertility

Many surveys in the United States suggest that more than three-quarters of teenagers and young adults are sexually active. Among these populations and especially on college and university campuses, there is an alarming incidence of STDs. One study reported by the Albert Einstein College of Medicine found that among 608 young women between the ages of seventeen and twenty-three, the cumulative thirty-six-month incidence of vaginal human papilloma virus (HPV) was 43 percent (see Suggested Reading). During the entire course of the three-year study, as many as 60 percent of the participants were found to have HPV. Among sociocultural and socioeconomic factors that seemed to influence the risk of HPV infection were (1) younger age, (2) increased frequency of sexual activity, (3) increased number of sex partners, (4) sex partners who were not attending college, (5) increased frequency of the consumption of alcohol, (6) minority ethnicity, and (7) engagement in anal sex. The incidence of cervical cancer in women of child-bearing ages thirty to forty years who have had HPV

and other STDs in their adolescence and young adulthood is increased relative to those who were free of these diseases. Having cervical cancer reduces the likelihood that a woman will be able to conceive normally.

In late 2006 and early 2007, a vaccine against HPV was developed by Merck. A national debate began to determine whether or not all young women before reaching the age of puberty should be vaccinated. The governor of Texas made national news in February 2007 by declaring mandatory vaccination of young women in that state. His move was even more controversial, in part, because he apparently bypassed both the Texas state legislature and the public before signing the mandate.

By any definition of good health, a woman who is infected with vaginal HPV does not have a healthy reproductive system. Some of the questions any woman should ask herself at an early age are (1) Do I want to maintain a healthy reproductive system? (2) Would I like to minimize the chances of cervical cancer? (3) Would I like to increase the chances of giving birth to healthy children? If the answers to these and related questions are yes, then a choice must be made. If the young woman chooses to have nonmonogamous sexual relations, then she runs an increased risk of contracting HPV and other STDs. If she wishes at all costs to avoid these diseases, then she must choose to not engage in nonmonogamous sexual activities.

Young men are not freed of the responsibility of asking themselves similarly challenging questions. Perhaps as a young man, your goal is to marry a woman who has never had an STD. If that is the case, you might ask yourself, What is my role in helping to minimize the chances of young women contracting STDs? Am I willing to reduce their risks by controlling myself and reserving sexual activities for marriage?

Male fertility, at least in industrialized nations, is on the decline, and no one knows the cause. There are several ways to quantify male fertility, and the most reliable measure is the number of mature spermatozoa produced per gram of testicular tissue per day. In twenty-year-old men, this rate is approximately 6.5 million; if the average twenty-year-old has ten grams of testicular tissue in each testis, then each day he will produce about 130 million spermatozoa (see Suggested Reading). There is a decline in this rate with age so that it drops to about half that number in older men ages fifty to ninety. Another measure of male reproductive fertility is the number of sperm cells per milliliter of ejaculate, and there is growing concern that this number has been on the decline in recent decades as well. Reproductive biologists who have compared male fertility in the past decade with data gathered about forty years ago have reported an approximately 30 percent reduction. Findings are similar in North America and Western Europe.

9

The Immune System

Our bodies are continuously bombarded by a variety of infectious pathogens including but not limited to bacteria, fungi, molds, parasites, spores, and viruses. Many of these circulate in the atmosphere as airborne matter. Their concentrations and varieties can vary regionally in any country or clime. But they are also normal inhabitants of the skin, the mouth, the respiratory passages, the GI tract, the urinary tract, and the lining membranes of the eyes. They regularly compromise physiological functions of the cells, tissues, and organs, but when invading deeper body tissues en masse they can and do cause serious pathological conditions including morbidity and mortality.

Our body's main defenses against pathogens are circulating white blood cells (leukocytes, also called WBCs) and residential tissue cells that are derived from WBCs such as tissue macrophages. These systems work in coordination by ingesting and destroying foreign materials through phagocytosis and by creating antibodies and sensitized lymphocytes that can either inactivate or destroy the invaders.

Cellular Defenses

Immune implies to defend against infection either by specific or nonspecific mechanisms. The word also has general reference to the immune system or to an immune response. One can refer to any good medical dictionary or textbook of immunology and find tables of immunodeficient disorders, diseases, and syndromes. Additionally, the molecular names and structures of the various immunoglubulins in our blood can be found in such books.

Leukocytes are the body's mobile units of defense. They are formed in the bone marrow but also in the lymphatic system. This is one of the needs for healthy bone marrow and healthy lymphatics. The two leukocyte-forming tissues

are referred to as myeloid (originating in the marrow) and lymphoid (originating in the lymphatics). There are five or six classes of WBCs. Some are characterized by the number of nuclei in the cell, the structure of the nuclei (single-lobed nuclei versus polylobulated nuclei), and their staining qualities (basic, acidic, or neutral). Three classes of polymorphonuclear leukocytes have been described: basophils, eosinophils, and neutrophils. *Polymorphonuclear* means there are many nuclei of differing shapes. The prefixes baso-, eosino-, and neutro- refer to the colors of the granular, cytosolic elements when they are exposed to different stains. The fourth and fifth classes of WBCs are monocytes and lymphocytes. A sixth class, the plasma cells, is less common. Quantitatively, neutrophils are the most abundant. They constitute about 60 percent of all WBCs and are followed, quantitatively, by lymphocytes (about 30 percent), monocytes (about 5 percent), and the others.

Because of their prevalence and from this point forward when referring to granulocytes, I will be considering only neutrophils. Granulocytes are so named because their cytosol contains secretory granules. Once released from the bone marrow into the systemic circulation, granulocytes remain in the vascular compartment for only a few hours. Thereafter and by the process of diapedesis, granulocytes leave the circulation and take up residency in the tissue spaces where they reside for a few days. In times of serious infection, this life cycle is reduced considerably because the granulocytes move rapidly to the site of infection, perform their defensive functions, and are themselves destroyed in the process. Monocytes spend about twice as much time as granulocytes in the circulatory system before they escape into the interstitial spaces. Once outside the vascular compartment, monocytes enlarge considerably to become tissue macrophages. Macrophages can reside in the tissues for weeks or even months unless they are consumed in phagocytic activity shortly after leaving the circulation. The system of monocytes and macrophages has been referred to as the monocyte-macrophage system, the tissue-macrophage system, and the reticuloendothelial system (RES). The last phrase is more entrenched in the medical literature and therefore is more commonly used in general.

Lymphocytes originate in the lymph nodes and other tissues of the lymphatic system. They are released along with other lymphatic products into the systemic circulatory system at selected sites such as the subclavian veins of the neck. After a few hours and when they have performed their phagocytic functions in the blood stream, they leave by diapedesis and enter the tissue spaces. Subsequently, they reenter the lymphatic system and are again discharged into the systemic venous system. Thus, lymphocytes get recycled and are therefore able to perform their defensive actions several times during their life cycles of a few weeks to several months.

Neutrophils and macrophages are among the main combatants of our immune systems. Neutrophils are mature cells that are able to attack and

TABLE 9.1

Cells of the immune system

Cell type	*Functions*	*Characteristics*
Erythrocytes	transport respiratory gases	anucleate, biconcave, discoid, hemoglobin
Granulocytes	phagocytosis, inflammation	enzymatic granules, extravascular mobility
Platelets	hemostasis (blood clotting)	small, anucleate, adherence, aggregation
B-lymphocytes	produce antibodies	in blood, lymph, lymph nodes, tissues
T-lymphocytes	tissue immunity	in blood, lymph, lymph nodes, tissues
Monocytes	tissue immunity	circulate then move into tissues
Macrophages	phagocytosis	released as monocytes, reside in tissues as macrophages

destroy bacteria and other pathogens once they get into the systemic circulation. However, macrophages begin their combat careers as circulating monocytes. They are immature in the monocyte state and are therefore not prepared to fight infection. Once they enter the tissue spaces, their morphology and physiology begins to change, making them more prepared for combat. In the tissue spaces, they immediately begin to enlarge, oftentimes reaching dimensions that are five to ten times their immature size: they reach external diameters of sixty to eighty micrometers as compared to five to fifteen micrometers as monocytes. After achieving the stature of combat soldiers, the former monocytes become extremely potent defenders of the interstitial spaces. Table 9.1 summarizes some of the cells that are involved in the mammalian immune system.

Chemotaxis, Margination, Diapedesis, and Phagocytosis

When injured and inflamed by pathogens, the body's tissues and cells know how to signal for help. Their SOS messages come in the form of multiple chemicals called chemoattractants that are released by the cells. Chemoattractants are most concentrated in the immediate area of release, but by principles of diffusion they can spread to considerable distances from the local site of injury. The process by which they signal for help from macrophages and circulating neutrophils is

called chemotaxis. When a tissue becomes inflamed, dozens if not hundreds of products are released by the injured cells as well as by the offending pathogens. The released factors, chemical signals, and agents include pathogenic toxins released by bacteria and viruses, decomposition products of the injured cells themselves, reaction products of the complement complex (see below), and other substances. Chemoattractants can diffuse to distances of 100 micrometers from the site of injury and inflammation. Since under normal conditions no cells in the body are more than about 50 micrometers from the nearest capillaries, chemotaxis is an extremely effective mechanism for sending messages to nearby macrophages as well as to circulating neutrophils.

As the chemoattractants reach adjacent capillaries, circulating neutrophils are drawn from the central axis of the bloodstream to the inside perimeters of the vessel walls in a process called margination. This enhances the approximation of neutrophils, capillary endothelium, and chemical signals. Because the velocity of blood flow is slowest near the capillary wall compared to the central axis, forward movement of neutrophils is reduced considerably as they approach the wall. Once the cells come in contact with the vascular endothelium, they begin tumbling slowly along the wall until they come to a complete stop. After their forward movement is arrested, they are able to establish contact with chemoattractants. Chemoattractants and other products of tissue injury increase permeability of the capillary wall by increasing the distance between adjacent endothelial cells. This enables neutrophils, by the process of diapedesis, to move through the gaps and into the interstitial spaces. The process is analogous to the movements of an amoeba. Cytoplasmic extrusions are passed into pseudopodia (false footlike projections) that have much narrower external diameters than the body of the cell. When the first pseudopodium has slipped through the inter-endothelial gap, more cytoplasm is extruded into the extravascular component of the neutrophil. Sometimes using multiple pseudopodia, this mechanism continues until the entire cell has passed to the interstitial side of the capillary wall.

Upon approaching an object to be phagocytized, the neutrophil or macrophage first attaches itself to the pathogen. It then distributes pseudopodia around the entire object, and these meet and fuse on opposite sides. This process draws the foreign matter, in an encapsulated form, to the interior of the neutrophil. The encapsulated object is called a phagocytic vesicle or phagosome. Neutrophils are generally able to destroy only a few pathogens, five or ten, by this process. Macrophages, on the other hand, are larger and much more powerful phagocytes. Individually they are able to destroy up to 100 or more pathogens by these phagocytic mechanisms. In addition, because of their size macrophages are able to engulf whole red blood cells and even malarial parasites, whereas neutrophils cannot engulf particles much larger than small bacteria. In both cases, once the immune cells have destroyed several invaders, they wear out and must be phagocytized themselves.

When the phagocytic vesicle is inside the macrophage, the wall of the vesicle comes in contact with lysosomes and other cytoplasmic granules. Membranes fuse, and digestive enzymes, bactericidal agents, and strong oxidants are released from the lysosomes into the phagocytic vesicle. The phagocytic vesicle now becomes a digestive vesicle. Proteolytic enzymes and lipases are part of this arsenal and are needed to digest protein and lipid components of the membranes of the pathogens. Oxidizing agents are among the most effective bactericidal weapons of granulocytes and macrophages. They include superoxide anion, hydrogen peroxide, hydroxyl radicals, and peroxynitrite. Few pathogens are able to survive the lethal effects of these chemicals.

Ports of Entry for Pathogens and the Reticuloendothelial System (RES)

Some of the monocytes that leave circulation and reside in the interstitial spaces remain mobile and free to move about. Others become attached to the tissues and remain fixed in this position for months or even years. As already mentioned, the monocyte-macrophage system is widely distributed throughout the tissues and organs of the human body. The total combination of monocytes, free-ranging and fixed macrophages, and specialized endothelial cells in the bone marrow, spleen, and lymph nodes is called the reticuloendothelial system. One would expect this system to be most abundantly developed at sites where pathogens are expected to enter the body. This is not to say that the system is not found elsewhere. Both the systemic and lymphatic circulatory systems are widely distributed to the cells, tissues, and organs, and infection often spreads through these circulatory systems. But by identifying ports of entry from the external environment, one can more easily understand the natural defenses that help protect the internal milieu. Generally speaking, the skin is a shield against invasion by foreign bodies; under normal circumstances it is impregnable to them. However, this is not true of the compromised skin, that is, an area with an open sore or laceration.

At a family picnic in Utah one summer, my brother killed a large western diamondback rattlesnake. One nephew asked if he could have the rattles and my son wanted to keep the skin to show his friends in New Jersey. I carefully removed the rattles and skinned the snake. To do this I turned the rattler on its back, stretched it out, then made a midline incision from tail to head. I used my forefingers and thumbs to gently separate the skin from the underlying subcutaneous tissue and muscle. It took thirty or forty-five minutes to remove and treat the skin for preservation.

Skinning that snake took place about 5 P.M., I washed my hands thoroughly but noticed the next morning that my right thumb was swollen, reddened, and very tender. The inflammation—edema—was on the medial surface adjacent to my right index finger where the thumb's soft tissue joins the lateral margin of

the thumbnail. The cardinal signs of infection—*calor* (heat), *dolar* (pain), *rubor* (redness), and *tumor* (inflammation and swelling)—intensified for the next couple days. I applied topical antibiotic to little effect, and I determined to go to the emergency room if I saw no improvement by the next day. The following morning the soreness and inflammation had begun to subside, but it took several weeks for my thumb to return to normal. Obviously some pathogens had been able to penetrate the skin of my thumb or to enter through an unnoticed wound. Fortunately for me, my immune system was up to the task of destroying and eliminating the pathogens.

Infections in subcutaneous tissues are generally confined to local regions and are defeated by local macrophages and neutrophils. If the foreign matter escapes destruction locally, it usually does not enter the systemic circulatory system directly. Rather it passes through local lymphatic capillaries and then makes its way to the systemic circulation. This is because lymph flow is unidirectional. On their way through the lymph system, pathogens will come in contact with lymph nodes that are interspersed at regular intervals. These are armed with resident macrophages and lymphocytes that are ready to do combat with the pathogens. Swelling of lymph nodes is one way of detecting systemic infection of the body.

Another port of entry is the respiratory system. Pathogens can enter the respiratory system piggybacked on particulate matter that we routinely inhale. In heavily wooded areas of the country with dense human populations such as Boston to Washington, D.C., the period of mid-April to early June can be especially problematic for the respiratory and immune systems. During this season, trees, bushes, and grasses are beginning to bloom. The increasing concentrations of atmospheric pollen, spores, and molds are vehicles for the transportation of airborne pathogens to human respiratory systems. For asthmatics and other sufferers of allergies, this is a challenging time of the year. The respiratory component of the body's RES must be particularly healthy. Macrophages are found mainly in the walls of the alveoli within the respiratory zone of the lungs. However, they are also present in the conducting zone and the upper respiratory system. As long as the respiratory immune system is in good health, airborne pathogens are of less concern.

The gastrointestinal system is another favorite port of entry for pathogens. The food we consume is one of the main carriers of bacteria and other pathogens. Our mouths and oral cavities always contain multiple strains of well-distributed bacteria including pneumococcal, spirochetal, and streptococcal colonies. These pathogens are also present to a lesser degree in the entire respiratory and gastrointestinal tracts. Sore throats can result from viral infection, bacterial infection, allergies and postnasal drip, and airborne pollutants such as cigarette smoke, chemical fumes, and dusts. More than 90 percent of sore throats are caused by cold and flu viruses, adenoviruses, coxsackie viruses (childhood summer sore throats), and Epstein-Barr viruses. Only 5 to 10 percent are caused by bacteria such as streptococcus and gonococcus.

On occasion people complain of biting their cheeks, lips, and tongues during mealtime. Cankers are prone to appear in or near these wounds, especially when the food has ingredients containing acids, such as walnuts. Any lesion invites invasion and infection by residential bacteria. When these accidents happen or when cankers appear, it is a good practice to rinse the mouth and throat with warm salt water immediately after the meal. One should then brush and floss and rinse and gargle with hydrogen peroxide (1–3 percent solution).

I have found that the following minimizes my suffering and discomfort from sore throats, cankers, and accidental biting of my tongue and cheeks. When the first signs of soreness or injury appear, I avoid eating foods with high fat or sugar content for several hours before bedtime. This reduces late evening to early morning secretions of mucus, helping me to breathe more freely and swallow less frequently during the night. Secondly, I gargle with hydrogen peroxide, then immediately before bedtime I rinse my mouth and gargle with saltwater as warm as I can stand it. I make a solution that is two to three times the physiological salinity of extracellular fluids.[1] The saltwater gargling reduces inflammation and relieves soreness, especially for the early minutes and hours of sleep.

By coating the throat with a solution that has higher salinity than that of the adjacent inflamed cells, the salt solution draws water out of these cells, thus reducing the inflammation. This is part of the homeostasis of water and electrolyte balance discussed in chapter 1. As the osmolarity of extracellular fluids in the inflamed tissues increases, water flows osmotically from cells to the interstitium. Excess water is removed by the circulation. During the time of increased extracellular osmolarity and loss of intracellular water, inflammation decreases. Any net loss of water or net gain of osmolarity will be compensated as homeostatic mechanisms reequilibrate the inter-compartmental distributions of both. These practices help to keep my oral cavity and throat in better repair.

As the food and pathogens we ingest get digested and absorbed, they pass through the liver before reaching the general systemic circulation. They do this through the hepatic-portal circulatory system. Hepatic-portal blood is so named because it originates in the gastrointestinal system (and related visceral organs) and enters the liver as splanchnic venous blood flow. The liver is an effective filtering station. It contains sinusoids that are lined by macrophages called Kupffer cells. The Kupffer cells are so effective at phagocytizing bacteria that many pathogens are destroyed within milliseconds of coming in contact with them. In addition to the immune functions of Kupffer cells, the liver is the site of synthesis of gamma globulins (the antibodies). Antibodies are also known as immunoglobulins (abbreviated Ig). There are five classes: IgA, IgD, IgE, IgG, and IgM.

The immunoglobulins constitute about 20 percent of the circulating plasma proteins. The colloidal properties of plasma are commonly characterized by computing an albumin to globulin ratio (A/G ratio). Globulin refers to the immunoglobulins. These antibodies are large proteins having molecular weights

ranging from a low of about 150,000 up to 1 million daltons. They are composed of light and heavy polypeptide chains each of which has a variable and a constant element. The variable element is capable of attaching to specific antigens. The constant element determines other biological properties such as ease of diffusion, specific sites of attachment within the interstitial spaces, and interactions with the complement system. Quantitatively IgG constitutes about 75 percent of all circulating immunoglobulins and is therefore the main antibody.

Antibodies act in two ways, directly and indirectly. Directly they attach to and destroy foreign invaders. Indirectly they activate the body's complement system, which has at its disposal multiple mechanisms of attack. The key components in this system are a group of proteins designated C1 through C9, B, and D. After binding to an antigen, a certain portion of an antibody's constant domain becomes activated. This allows it to bind to the C1 component of the complement system. A cascade of enzymatic actions is then mobilized, and cellular events that include the following take place: (1) opsonization and phagocytosis, or the uptake of the antigen and antibody complex by activated neutrophils and macrophages, (2) lysis, that is, the disruption of the invading pathogens' membranes, (3) agglutination, or changes in the surface properties of the antigens that cause them to adhere to one another thus minimizing the spread of individuals, (4) neutralization, or rendering invading viruses and bacteria nonvirulent, (5) chemotaxis, or attraction of large numbers of neutrophils and macrophages to the site of invasion, (6) activation of mast cells that release histamine and other factors to increase capillary permeability and augment local blood flow, and (7) inflammation and walling off, that is, the prevention of diffusion of toxins to locations beyond the site of invasion.

Immunity, Regenerative Medicine, and Stem Cells

One's immune responses have a genetic component but are also highly individualized. The prediction is that by the year 2017, treatment of diseases will be tailored to the individual and not to the masses. This means that two people suffering from the same ailment could have markedly individualized treatments beyond the next decade or so. One of the areas of inquiry that leads to such claims is stem cell research.

Despite major differences in the structure and function of cells circulating in the blood and in their longevity inside the vascular compartment, all circulating blood cells derive from pluripotent hematopoietic stem cells (hemangioblasts, see figure 9.1). *Pluripotent* refers to the ability of hemangioblasts, which are the progenitors of blood cells and vascular endothelial cells (those that make new blood vessels in the process called angiogenesis), to give rise to many different cell lines. *Hematopoiesis* means blood forming; *angiogenesis* means blood vessel forming. Important physiological characteristics of hematopoiesis and

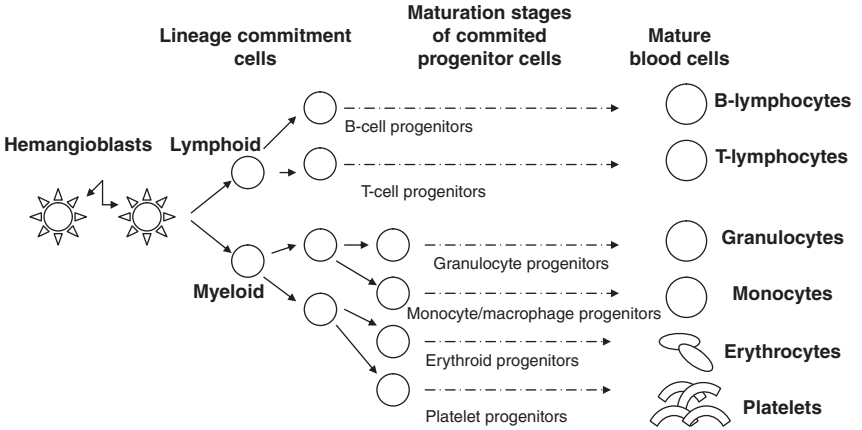

FIGURE 9.1 The processes involved in hemangiopoiesis (formation of new blood cells and new blood vessels from stem cells, hemangioblasts). Note that stem cells can develop into progenitor cells that become definitive cell types (T-lymphocytes, monocytes), or can rejoin the pool of undifferentiated hemangioblasts.

angiogenesis, therefore of hemangioblasts, are self-renewal, proliferation, and differentiation of downstream progeny that lead to production of mature blood cells and new blood vessels.

Prior to the late 1990s, there was little public awareness of stem cells and stem cell research. However, by the end of 2006 reports of finding new stem cells were almost a monthly occurrence. For example, the November 24, 2006, issue of *Science* carried a summary report of two independent groups who simultaneously identified the same stage embryonic mouse stem cells earlier that year. Before these cells differentiate into either cardiac myocytes or vascular smooth muscle cells, they progress through an intermediate stage called "triple positive cells." This means that at that stage of development, the embryonic cells express three specific genes as markers of cellular development. Similar cells were found by both groups in mouse embryos at precisely day eight of in utero development.

In figure 9.1, a daughter cell, one that arose from a hematopoietic stem cell, has begun to differentiate. Because it can no longer duplicate itself, it is called a committed progenitor cell. Only a few of the many stages of its differentiation are illustrated in figure 9.1. In reality, there is a continuum of differentiation with multiple divisions between each of the adjacent stages. In the differentiation scheme that is illustrated, a few stem cells give rise through successive divisions to increasingly large numbers of daughter cells. Each successive stage has less opportunity for differentiation and is therefore more restricted in potential than the preceding stage. The pluripotent hematopoietic stem cell is elusive. No pure samples have been harvested to date.

The field holds the promise of regenerative medicine, that administration of stem cells to damaged tissues can lead to their replacement with healthy cells. Already, both experimentally induced and naturally occurring heart attacks (myocardial infarctions, or MI) in both animals and humans have been treated with stem cells. Replacement of necrotic myocardial tissue with regenerated tissue has been reported with varying degrees of success. This work is progressing, and scientists and clinicians hope that all disease—cancers and diabetes as well as cardiovascular ones—will be treatable using stem cells. To this end, several states have announced stem cell research initiatives.

Myocardial infarction is a complex disease state that involves more than the immune system. Both reversible and irreversible heart damage can be caused simply by occluding blood flow to the myocardium. If blood flow is occluded for a few minutes and then restored, myocardial function during the postocclusion reperfusion state can be compromised and not restored fully for the next twenty-four to forty-eight hours. The impeded recovery is called myocardial stunning. If blood flow is deprived for nearly an hour instead of a few minutes, some of the cardiac muscle cells become apoptotic and others necrotic (this ischemic damage is discussed further below). When blood supply is disrupted for more than forty-five to sixty minutes and then restored, damage in addition to that caused by cessation of blood flow is superimposed on the ischemic damage. This is called reperfusion injury (also discussed below). Thus, during a heart attack, tissue damage can be caused by the disruption of blood flow, by the restoration of blood flow, and by a combination of the two.

The immune event that occurs during myocardial ischemia is the activation of an immunologically important chemical system called complement. Complement and reactive oxygen species (ROS) activate a downstream cascade involving release of tumor necrosis factor (TNF-α), which in turn activates cellular cytokines. Individually and collectively, these tissue factors do damage to subcellular and cellular organelles and thus to tissues, whole organs, and organ systems. (In all physiological states, some agents get down-regulated while others are up-regulated. For example, down-regulating angiogenic growth factors and simultaneously up-regulating anti-angiogenic factors can lead to reductions in the size of a solid tumor.) Beyond ischemic injury is the damage caused during reperfusion. Reperfusion leads to inflammation, and inflammation involves release of more damaging cytokines and chemokines. Inflammation also disrupts the homeostasis of important Starling forces in the microcirculation.

Ernest Starling (1866–1927) was a British physiologist whose interests included the mammalian microcirculation. Starling identified four variables inside and outside of capillaries that influence homeostasis of blood volume. The four variables are capillary hydrostatic pressure and capillary oncotic pressure (also called plasma oncotic pressure, plasma protein oncotic pressure, protein oncotic pressure, or colloid osmotic pressure) and their corresponding extravascular counterparts,

interstitial hydrostatic pressure and interstitial colloid osmotic pressure. These four variables are the direct determinants of intravascular blood volume and of extravascular interstitial volume. They are affected by two indirect determinants, pre- and postcapillary resistance. When the steady state of these variables is disrupted, edema and inflammation occur. Therefore, inflammation is a circulatory problem, and all stimuli and stressors must operate through the microcirculation in order to cause inflammation. This includes the activities of cellular and molecular signals and factors that continue to be identified.

Imagine you are outside painting the trim on your shed, working in the garden, or participating in a family picnic. You unsuspectingly disturb a hornet's nest and get stung on the wrist. Prior to the accident, microcirculation, Starling forces, and the turnover of cellular cytokines and immunoglobulins in the skin and subcutaneous tissues of your wrist were in their physiological steady states. However, the hornet's injectate contains proteins and other products—perhaps plant antigens—that are foreign to the tissues of your body. Within minutes of the sting and beginning at the site of injection, you note the classic signs of infection and inflammation: calor, dolar, rubor, and tumor. These symptoms and the subsequent swelling of the entire forearm during the next several hours and days are consequences of the foreign injectate activating the tissue's immune system and disrupting Starling forces and microcirculatory homeostasis.

When tissue injury occurs, regardless of the cause, multiple chemicals and molecules are released by the injured tissues (table 9.2). These cause marked changes not only in the injured tissues but in the surrounding uninjured tissues as well. Many changes that occur subsequent to the initial injury contribute to inflammation. Inflammation is characterized by local vasodilation, increased blood flow, increased capillary permeability, and loss of potentially large volumes of fluid from the intra- to the extravascular spaces. Exudation of plasma proteins including fibrinogen from the vascular to the extravascular spaces also occurs. Exudation can lead to clotting of interstitial fluid and migration of monocytes and neutrophils into the tissue spaces. Some of the cellular chemicals that cause the above responses include histamine, bradykinin, serotonin, prostaglandins, and products of the complement and blood clotting systems. Several dozen additional factors have been identified by cellular and molecular biologists. While a comprehensive list is beyond the scope of this book, table 9.2 presents a partial alphabetical list. All such factors are potential targets for future treatment of tissue damage and inflammation.

Within minutes after inflammation begins, macrophages, already present in the tissues, begin their phagocytic function. In response to infectious signals and inflammation, these cells hypertrophy and become mobile. While their numbers might not be large, this first line of defense is often lifesaving. In the next hour or so, large numbers of circulating neutrophils begin to concentrate in the injured area. Products released from the inflamed tissues change the

TABLE 9.2

Some of the factors involved in tissue injury and inflammatory responses

Acronym	*Term (phrase)*
α-SMAc	α-smooth muscle action
bFGF	basic fibroblast growth factor
CSIF	cytokine synthesis inhibiting factor
DCHF	dichlorofluorescein
ICAM-I	intercellular adhesion molecule-1
IL	interleukin
IP-10	interferon-γ-inducible protein-10
LFA-1	leukocyte function antigen-1
LPS	lipopolysaccharide
LTB-4	leukotriene B4
LTC-4	leukotriene C4
MCP-1	monocyte chemoattractant protein-1
M-CSF	macrophage-colony stimulating factor
MMP	matrix metalloproteinase
NADP	nicotineamide-adenine dinucleotide phosphate
NF-κβ	nuclear factor-κβ
PAF	platelet activating factor
PAF-AH	platelet activating factor-acetylhydrolase
PDGF	platelet-derived growth factor
PSGL-1	P-selectin glycoprotein ligand-1
ROS	reactive oxygen species
SCF	stem cell factor
sCR1	soluble complement receptor type I
Smemb	embryonic smooth muscle myosin
SOD	superoxide dismutase
TIMP-1	tissue inhibitor of metalloproteinases-1
TGF-β	transforming growth factor-β
TNF-α	tumor necrosis factor-α
TNFR	tumor necrosis factor receptor
u-PA	urokinase type plasminogen activator
VEGF	vascular endothelial growth factor
VLA-5	very late antigen-5

luminal surface of capillary endothelial cells, causing the neutrophils to adhere to them as part of the process of margination. Secondly, similar locally released factors cause the intercellular attachments between adjacent endothelial cells of capillaries and postcapillary venules to loosen. This causes openings large enough for neutrophils to pass through by the processes of extravasation and diapedesis. The third line of defense is the chemical attraction of newly arrived interstitial neutrophils to the actual site of tissue injury.

Because circulating neutrophils have already reached their maturity within the vascular compartment, they are immediately able to go to work killing pathogens and scavenging other foreign products that have invaded the tissue spaces. While the need exists, the number of circulating neutrophils can increase from the physiologically normal 3,000 to 5,000 cells per microliter of whole blood to 15,000 to 25,000 per microliter, a hyperplastic process known as neutrophilia. Although several dozen factors have been implicated in the control of the macrophage and neutrophil response to inflammation and injury, five are believed to play dominant roles (see table 9.2). They are TNF-α, interleukin-1, and colony-stimulating factors such as granulocyte-monocyte colony-stimulating factor (GM-CSF), monocyte colony-stimulating factor (M-CSF), and granulocyte colony-stimulating factor (G-CSF). All are produced and released by activated macrophages in the injured tissues.

10

Muscle Function

Muscle Diversification

In chapter 1, I described the relationship between structure and function using two examples, muscles and kidneys, to illustrate. Structurally, muscle can be broadly classified as either striated or nonstriated. The two kinds of striated muscle are skeletal and cardiac. Nonstriated muscle is further characterized as visceral smooth muscle or vascular smooth muscle. Smooth muscle does not contain well-defined striations and is most commonly found in hollow tubular structures such as blood vessels, gut walls, and fallopian tubes. As well as the above distinctions, scientists classify different muscle types according to both physiological function and metabolic pathways for energy production and consumption.

Some skeletal muscles do not need to contract rapidly, but they do need to resist fatigue and to maintain tension for long periods of time. An example of this kind of muscle is the soleus of the lower leg. On the other hand, some muscles need to contract rapidly but at infrequent intervals. Examples of these kinds of muscle include the extraocular muscles of the eye and the extensor digitorum longus of the wrists and ankles, fingers and toes. Striated muscles are further characterized according to the rate at which they develop force once stimulated, the rate at which they shorten to lift a load, the rate at which they fatigue upon repetitive stimulation, and the rate at which they recover from fatigue. Muscles are therefore referred to as *slow twitch* (type I fibers whose rates of contraction upon stimulation are relatively slow) and *fast twitch* (type II fibers, with relatively fast rates of contraction). Fast-twitch fibers are further classified as type IIa (fatigue resistant or indefatigable) and type IIb (fatigable).

In their cytosol, slow-twitch muscles have a high content of the oxygen-binding protein myoglobin, they have many capillary networks surrounding each individual fiber, and they have a stable supply of oxygen. For these and related reasons (for example, many mitochondria per fiber), they are called

oxidative fibers: they have little capacity to store glycogen and must rely on a steady supply of oxygen and metabolic substrate in order to function.

Fast-twitch fibers differ widely among themselves. As already mentioned, they consist of both fatigue-resistant and fatigue-sensitive muscle fibers. They have a relatively lower blood supply and mitochondrial content than slow-twitch fibers because they rely less on oxidative metabolism for contraction. Fast-twitch fibers are generally much larger than slow-twitch fibers, they have a more extensive sarcoplasmic reticulum (SR) for the rapid release of calcium ions, and they have a high content of glycolytic enzymes so they are able to release energy for glycolytic processes rapidly. Because fast-twitch fibers have less myoglobin, they are lighter in color than slow-twitch fibers. Thus, slow twitch are also called red muscle fibers and fast twitch are known as white muscle fibers.

Slow- and fast-twitch fibers represent the extremes of a continuum of muscle fiber types. Each whole muscle (biceps or gastrocnemius, for example) has multiple types of fibers, but one kind tends to predominate. Differences in muscle fiber types derive in large part from the expression of various isoforms of the contractile and regulatory proteins that are present in each fiber. For example, myosin ATPase activity might be directly correlated with the rate of contraction, and there have been more than a dozen isoforms of myosin heavy chains identified to date. At least four isoforms of the myosin heavy chain (MHC-I, MHC-IIa, MHC-IIb, and MHC-IIx/d) are expressed in skeletal muscle. Differences in the rates and strengths of contraction might also reflect differences in various isoforms of the myosin light-chain proteins, as well as differences in sarcolemmal and sarcoplasmic reticulum calcium channels, calcium pumps, and calcium-sequestering proteins.

Like the broad diversity among striated skeletal muscle fiber types, there are similar kinds of differences among cardiac striated and smooth muscle cell types. Cardiac muscle cells can be broadly classified as contractile versus conductile. Specialized muscle cells of the sinoatrial and atrioventricular nodes as well as those in the bundles of His, the right and left bundle branches, and the Purkinje fibers differ in size, conduction velocities, and refractory periods. The degree of autonomic innervation of each of these specialized regions of the mammalian conduction system is also markedly different. The right vagus nerve, for example, predominantly innervates the sinoatrial node while the left vagus nerve innervates the atrioventricular node and bundles of His. The sympathetic innervation of these same tissues shows similar differential distribution. Stimulating the right branches of cardiac sympathetic nerves is accompanied by a marked increase in heart rate with a modest increase in ventricular contractility. Stimulation of the left cardiac nerves causes little change in heart rate but significant increments in cardiac contractility.

There are even greater differences in the functions of different types of vascular and visceral smooth muscle. For example, the vascular smooth muscle responses to the same drug in two different arterioles can be different. In the

same segment of microcirculatory vessels, the arteriole upstream and the venule downstream might respond differently to the same agent. In addition, the same arteriole will respond differently to the same agent under differing physical and chemical conditions. For example, adenosine is a well-known naturally occurring coronary vasodilator. The vascular smooth muscle of coronary arterioles in the heart will relax considerably more to the same concentration of adenosine when the perfusate pH is 7.0 (acidic) than when it is 7.4 (neutral).

Excitation-Contraction Coupling in Muscle

Excitation-contraction coupling (E-C coupling, or electromechanical induction) is the physiological relation between an action potential that depolarizes a muscle cell membrane and the subsequent development of force and shortening (contraction) in that same cell. This phenomenon has been thoroughly investigated in both cardiac and skeletal striated muscle. It begins with an action potential depolarizing the sarcolemma (plasma membrane) of any striated muscle cell (see figure 10.1). The change in voltage of the membrane where it invaginates into the fiber at the t-tubules causes the opening of voltage-dependent calcium channels. Because

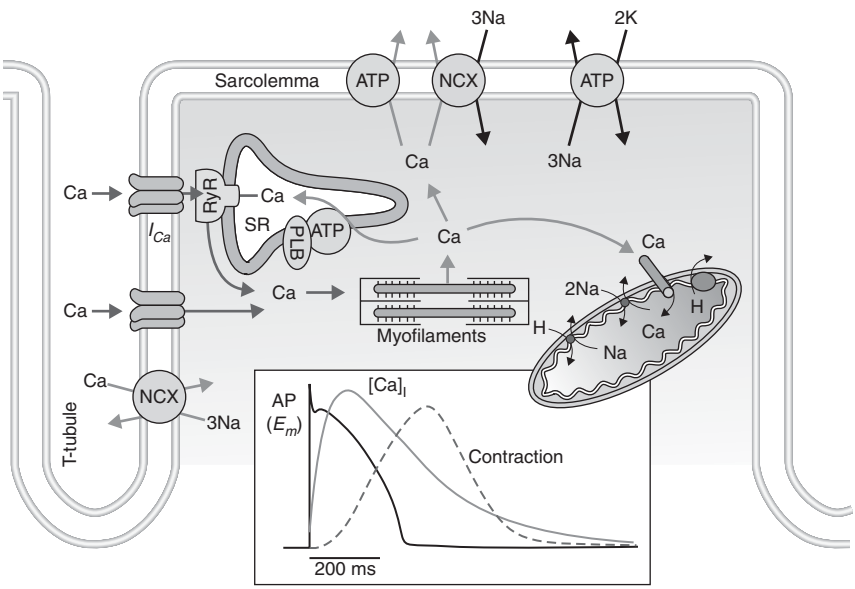

FIGURE 10.1 Processes involved in striated muscle excitation-contraction coupling. As an action potential depolarizes the sarcolemma, calcium channels are opened and extracellular calcium flows into the cell. Note location and positions of calcium receptors and calcium channels in the t-tubule and sarcoplasmic reticulum.

(From http://edoc.hu-berlin.de/dissertationen/abdelaziz-ahmed-ihab-2004-09-20/HTML/abdelaziz.html via images.google.com)

these particular calcium channels are sensitive to the drug dihydropyridine, the calcium receptor that is associated with the channel is called the dihydropyridine receptor (DHPR). Voltage-mediated activation of the dihydropyridine receptor opens the channel and allows extracellular calcium to flow into the cell (excitation). This calcium stimulates release of more calcium from the sarcoplasmic reticulum (SR) by opening a second set of calcium channels. Calcium receptors that are associated with this second set of calcium channels are called ryanodine receptors (RYR). The dihydropyridine receptor is named to reflect a class of calcium channel blockers that act at this site and that are called dihydropyridine calcium antagonists (nifedipine and nisoldipine were early examples of the class). The process by which the DHPR and RYR calcium channels interact during E-C coupling is called calcium-induced calcium release. As the SR releases its calcium, the intracellular concentrations of the ion increase to an extent that is sufficient to initiate the mechanical processes of contraction (see inset in figure 10.1).

As noted, contraction of striated muscle requires an elevation in myoplasmic concentrations of calcium. In addition, the process of contraction is regulated by the thin filament actin. The development of force by the myofilaments increases as the intracellular concentrations of calcium increase. The curve relating force development to intracellular concentrations of calcium is sigmoidal (S-shaped). Between intracellular calcium concentrations from one hundredth to one tenth micromolar, there is no change in force. As calcium increases above one tenth micromolar, force begins to increase, reaching a half-maximal response at about one micromolar. Under physiological conditions, maximal force is reached at an intracellular calcium concentration of about ten to one hundred micromolar.

Mechanistically, calcium released from the SR binds to the thin filament protein troponin-C. The calcium/troponin-C complex causes another protein, tropomysin, to move toward the actin filament. This movement exposes the critical myosin-binding site on the actin filament, which allows actin (thin filament) to bind to myosin (thick filament) and form a cross-bridge between the two filaments. There are several calcium-binding sites on troponin-C. Two of these have a high affinity for calcium ions but also bind magnesium ions at rest. There are additional troponin subunits as well, troponin-I and troponin-T. The two high-affinity calcium sites on troponin-C seem to be responsible for regulating interactions between troponin-I and troponin-T subunits. Binding of myosin to actin causes a further movement of tropomysin. Although a given tropomysin molecule extends over several actin filaments, it is hypothesized that the binding of myosin to actin exposes further binding sites on as many as fourteen actin molecules.

Cycling of Cross-Bridges and Shortening of Sarcomeres

Once activated by intracellular calcium ions, the binding sites on the actin filament become attached to the heads of the cross-bridges on the myosin filaments.

After actin and myosin are bound, ATP-dependent conformational changes in the myosin molecule result in the movement of the actin filaments toward the center of the sarcomere. This shortens the length of the sarcomere and thereby contracts the muscle fiber. The mechanism by which myosin produces force and shortens the sarcomere is hypothesized to involve four basic steps that collectively are called the cross-bridge cycle. In the resting state, myosin is thought to have partially hydrolyzed ATP (step 1; also called state a). When calcium is released from the SR, it binds to troponin-C, which exposes an active site allowing energized myosin to bind to actin (step 2, state b). Myosin then undergoes a conformational change called a ratchet action that draws actin toward the center of the sarcomere (step 3, state c). ATP and inorganic phosphate are released, resulting in the dissociation of myosin from actin (step 4, state d). Acting as an ATPase, myosin then hydrolyzes ATP using some of the released energy to recock the myosin head and to return myosin to the resting state.

If intracellular calcium levels are still elevated, myosin will undergo another cross-bridge cycle producing further contraction of the muscle. Each ratchet action of the cross bridge is capable of moving the thin filament approximately ten nanometers. The cycle continues until the sarcoplasmic reticulum calcium pumps (SERCA) remove cytosolic calcium back into the SR. As the cytosolic calcium levels fall, calcium dissociates from troponin-C and the troponin-tropomyosin complex moves and blocks the myosin binding sites on the actin filament. If the supply of ATP is exhausted, as occurs with death, the cycle stops at step 3 with the formation of permanent actin-myosin complexes. In this state, the muscle becomes rigid and is in a state called "rigor mortis."

When a muscle contracts, work is performed and energy is expended. Considerable quantities of ATP are reduced to ADP during the contraction process, and the greater the amount of work, the greater the demand for ATP cleavage, known as the Fenn effect. Following is the sequence of events thought to be responsible for the Fenn effect. (1) Before contraction begins, myosin heads bind ATP, which confers ATPase activity on them. The by-products of ATP hydrolysis, ADP and inorganic phosphate, remain attached to the myosin head. In this state, the myosin head is perpendicular to the actin filament but is not bound to it. (2) As the troponin-tropomyosin complex moves and exposes the myosin-binding site on actin, the myosin head binds to it. (3) In a power-generating movement, the myosin head rotates from the perpendicular position toward the arm of the cross bridge thereby drawing the thin filament over the thick filament. (4) As the filaments move over each other, the previously bound ADP and inorganic phosphate are released, and a new molecule of ATP binds to the vacated site on the myosin head. Binding of ATP causes detachment of the myosin head from the actin filament (see figure 10.2). (5) Detachment of myosin from actin cleaves the new molecule of ATP and (6) the cycle is ready to be repeated.

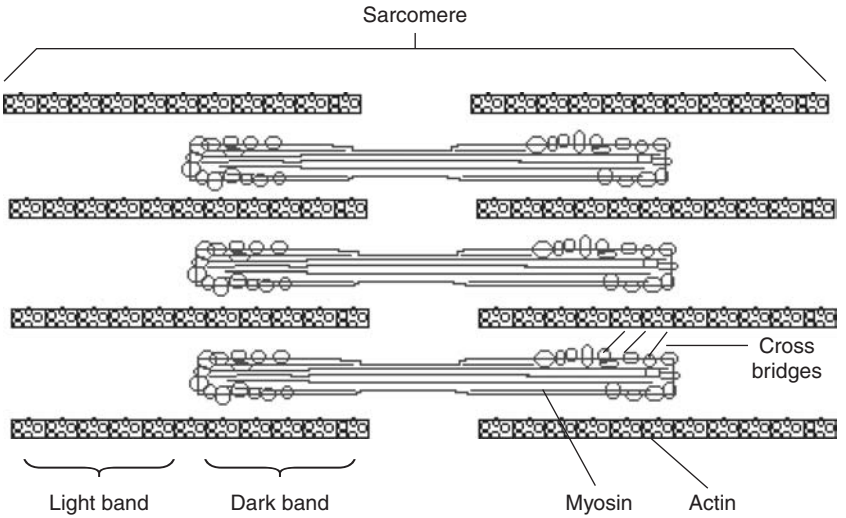

FIGURE 10.2 Cross bridges in striated skeletal and cardiac muscle. The cross bridges are part of the thick filament myosin molecule that attach to tropomysin binding sites on thin filaments.

(From http://www.tiem.utk.edu/~gross/bioed/webmodules/musclefig.gif via images.google.com)

Muscle Levers, Hypertrophy, and Atrophy

Striated muscles (particularly skeletal muscles) operate by applying tension to their points of insertion into bones. The bones form different types of lever systems. The larger the cross-sectional area of a muscle, the greater the maximum force of contraction. Someone whose weight training has produced biceps with cross sections of six square inches should be able to bench press about 300 pounds. One whose biceps are only half that size will be able to press only about 150 pounds. When the forearm is at right angles to the upper arm, the tendon attachment of the biceps is a couple inches anterior to the fulcrum at the elbow, and the total length of the forearm lever is about fourteen inches. Because of the physics and mathematics, the lifting power of the arm under these conditions would be about one-seventh of 300, or approximately 43 pounds.

The total mass of a muscle can increase under a variety of conditions. The condition that most readers are familiar with is weight training. When the mass of a muscle increases, it is called muscle hypertrophy. When the mass of a muscle decreases, it is called muscle dystrophy or muscle atrophy. Virtually all muscle hypertrophy is the result of an increase in the number of actin and myosin filaments in each muscle fiber. This results in an overall enlargement in the size of the individual muscle. Muscles hypertrophy more rapidly when they are loaded (are bearing weight) prior to contraction. The simplest evidence of

this concept is to try it. Measure the diameter of your biceps either in the tightly flexed (uplifted forearm at right angles to the biceps) or loosely relaxed position (forearm extended relative to the relaxed biceps). Find a room and some weights and begin bench pressing a reasonable amount of weight for six to ten weeks. Repeat this activity three or four times per week doing three sets of ten to twelve repetitions per set. After six to ten weeks, repeat the measurements on your flexed and relaxed biceps and note the difference.

The specific mechanisms by which forceful contractions (for example, bench pressing) lead to muscle hypertrophy are not known despite the years and money that have been spent on the investigation. Some evidence suggests changes in motor nerves and neuromuscular junctions (synapses) that favor muscle growth. Other evidence focuses solely on the muscle cell. We do know that the rate of synthesis of muscle contractile proteins is much greater during the development of hypertrophy. This leads to greater numbers of both actin and myosin filaments in muscle fibers, sometimes increasing by greater than 50 percent. It makes little sense to think that muscle hypertrophy could occur without a corresponding increase in the enzymatic and biochemical processes needed to provide energy to the growing muscle mass. This is especially true of the glycolytic enzymes and biochemical pathways.

Another type of muscle hypertrophy occurs when muscles are stretched to greater than normal length. This causes new sarcomeres to be added to the ends of the fibers where they attach to the tendons. When muscle is shortened to less than its resting length for extended periods, sarcomeres at the ends can actually disappear. These processes of lengthening and shortening in skeletal muscle are called muscle remodeling. They occur to ensure proper length for the physiological function of muscle.

The heart can also hypertrophy. Elite athletes and others who are persistent at training and conditioning often develop exercise-induced cardiac hypertrophy. This is good for the competitive athlete: enlargement of the heart means greater contractile power, which translates to improved cardiac output (blood flow). Having an increased cardiac output is one means of providing a greater supply of oxygen to exercising muscles and other tissues. But there is another kind of cardiac hypertrophy that is undesirable and a hallmark of ill health and impending death. This is cardiac hypertrophy secondary to congestive heart failure. In this case, the thickness of the ventricular walls does not increase but can actually decrease. The increased size of the heart reflects engorgement of the ventricular chambers with a greater than normal volume of blood. Because the ventricular muscle lacks strength of contraction, the chambers progressively overfill, cardiac output is reduced, and the homeostasis of both the systemic and pulmonary circulatory systems is compromised. When severe and terminal, cardiac hypertrophy due to congestive heart failure can only be corrected by cardiac transplantation.

Muscle Diseases and Conditions

Malignant hyperthermia is a serious disease that affects one in several thousand Americans per year. Along with central core disease and Brody disease, it is one of several genetic disorders that cause disturbances in calcium homeostasis in striated skeletal muscle. Affected persons are at potential risk when they are exposed to any of the volatile surgical anesthetics such as halothane and isoflurane. Administration of muscle relaxants such as succinylcholine can also induce and/or exaggerate malignant hyperthermia. Succinylcholine is a short-acting acetylcholine receptor antagonist that acts by briefly opening the ion channel associated with the receptor and then closing it. This results in a burst of muscle activity, including transient whole-body clonic contractions followed by muscle paralysis.

In surgery, malignant hyperthermia is accompanied by rapid breathing, rapid heart rate, low oxygen and high carbon dioxide contents in the circulating blood, and rising body temperatures (hyperthermia). The condition is also accompanied by marked swings in blood pressure, muscle rigidity, and sweating. The patient's temperature can rise as rapidly as one degree centigrade each several seconds. While the onset of the disease usually occurs during anesthesia, it can be initiated up to several hours after an operation. If left untreated, the patient can develop respiratory and lactic acidosis and muscle breakdown. The damaged muscle tissue releases excess potassium, and a profound circulating hyperkalemia ensues (elevated potassium in the blood). This is dangerous because hyperkalemia causes disturbances in the generation and conduction of cardiac action potentials. Under healthy physiological conditions, the concentration of potassium in the plasma is about four millimoles per liter of blood. Potassium concentrations of six to eight millimoles per liter can cause ventricular arrhythmias. When plasma levels reach twelve to sixteen millimoles, they can induce sustained ventricular tachycardia, ventricular fibrillation, and sudden cardiac death. In an experimental laboratory setting, circulating concentrations of potassium are often elevated to test the safety and efficacy of new anti-arrhythmic medications.

The major features of malignant hyperthermia—hyperthermia, muscle rigidity, and increased metabolism—encouraged early investigators to conclude that the disease represented aphysiological regulation of muscle contraction. Further support for this hypothesis came from studies using biopsy of muscle samples from susceptible and nonsusceptible humans and animals. Muscle samples from susceptible individuals contracted with greater vigor when exposed to the volatile anesthetics than did muscle samples from nonsusceptible subjects. Moreover, calcium-induced calcium release in the sarcoplasmic reticulum of susceptible patients was enhanced when compared with unaffected subjects (similar results have been obtained in experimental animals). Administration

of caffeine, which causes the SR calcium channels to open, produces greater muscle contraction in affected than in unaffected people. Taken together, these observations suggest that malignant hyperthermia results from an aphysiological response of the calcium-release channel in the sarcoplasmic reticulum. More work is needed to confirm these hypotheses.

Brody disease is characterized by painless muscle cramping and impaired muscle contraction during exercise. While climbing stairs rapidly, for example, muscle might cramp up and stiffen and be temporarily out of commission. This particular muscle abnormality is seen in the arms, legs, and eyelids and seems to be temperature dependent (it is intensified by cold weather). Brody disease can be either a recessive or a dominant trait and might involve mutations in several genes, but it is a rare malady occurring only in about one in 10 million births. Mechanistically, it appears that Brody disease results from a reduction in the activity of calcium pumps located in fast-twitch muscle.

Central core disease is a rare autosomal dominant trait that results in general muscular weakness from both disintegration of contractile filaments and loss of mitochondria in the central core of large muscles. The condition is also thought to result from mutations in the ryanodine receptor. As a result, release and uptake of calcium from the SR is impaired. This places greater responsibility on mitochondria to sequester calcium, leading to calcium overload and subsequent loss of mitochondrial function.

In my laboratory, when we compare cardiac mitochondria from injured hearts (using electron microscopy and biochemistry), we find general preservation of both structure and function when experimental animals were pretreated with acetaminophen versus those pretreated with placebo (acetaminophen solvent). Mitochondria have both an outer and inner system of membranes with ion channels traversing both membranes. One of the more important channels is called the mitochondrial permeability transition pore (MPTP). It regulates the transfer of protons such as hydrogen ions into and out of the mitochondrion. Scientists hypothesize that calcium overload, hypoxia and reoxygenation, and ischemia and reperfusion injury disrupt function of the MPTP. We hypothesize that acetaminophen protects the MPTP from such damage (see Suggested Reading).

Writer's cramp is a phrase I was first exposed to as a graduate student. It is a condition of the writing hand and arm brought about by excess use of flexor muscles and simultaneous disuse of corresponding extensor muscles. I experienced this phenomenon as a graduate student and more recently while writing this book. When writing my doctoral dissertation in 1975, I lived in my laboratory for the two months it took to finish my dissertation. There were no personal computers then. All my writing was done longhand. Because I had a deadline to meet, I did few things during those weeks except eat, sleep, and write. Except for meals, I wrote steadily for twelve to fourteen hours per day. I slept on a cot in the lab at the end of my workday.

One night I must have turned too vigorously in my sleeping bag, and I fell off the cot. As I tumbled to the floor, I extended my right arm to brace for the fall. As I extended the arm, a bolt of lightning-like heat and pain shot from my wrist to the shoulder. I fully awoke with a cry of anguish. After several minutes the pain subsided, and I began experimenting with the arm. I discovered that I could not extend it fully from a 90 to 180 degree angle. At about 125 to 135 degrees the pain reappeared, and I could not bear to extend the arm further.

It took the better part of six months to regain full use of the extensor muscles in my right arm. I later calculated that for 700 to 850 hours in that two-month period, my right arm remained flexed between 90 and 135 degrees. My extensor muscles were progressively atrophying without my being aware of it.

Experimenting with Muscle

As mentioned earlier in this chapter, the strength of contraction in striated muscle is directly proportional to the strength of a stimulus. I teach college students this principle each year in an experimental physiology laboratory. The gastrocnemius muscle (large calf muscle) is removed from the hind limbs of euthanized bullfrogs. The muscles are mounted in an apparatus that allows the students to stimulate them electrically while simultaneously recording contractile responses. The equipment and responses are attached to a modern data acquisition system (computers, transducers, A-D converters, amplifiers, and monitors). As the magnitude of the electrical stimulus is doubled, tripled, and quadrupled, the strength of contractions increases correspondingly. Successive contractions become stronger because more and more motor units (motor nerves plus the muscle fibers they innervate) are recruited as the voltage in the stimulator is increased. When the contractile responses no longer increase, a plateau is reached.

During one course, we divided the muscles into two groups, those treated with standard physiological salt solution (the control group) and others placed in standard physiological salt solution containing acetaminophen (the experimental group). We obtained data from thirty to thirty-five control and thirty to thirty-five acetaminophen-treated muscles. The plateau for maximal contractile strength at a given voltage was more than double in the acetaminophen-treated muscles than in the control group. This was an unexpected but interesting experimental finding. It revealed that by mechanisms we have not yet explored acetaminophen enhanced the function of striated skeletal muscle. Of course, these experiments were conducted by students in a teaching setting using amphibians. We did not argue that the findings were relevant to mammals.

Less than one year later, we were discussing these results with colleagues from a nearby medical school. They were muscle biologists trained in physiology and using modern molecular techniques to study skeletal muscle function in

mice. The mouse is a commonly used experimental animal today because it is easy to produce mutations in genes and the proteins they express. My colleagues invited us to collaborate with them to see if such acetaminophen-mediated effects could be replicated in a scientific setting in mammals. These investigators were expert in isolating, removing, and instrumenting both fast- (EDL) and slow-twitch (soleus) mouse muscle fibers. We tested the following hypothesis: acetaminophen will reduce fatigue and improve recovery in both fast and slow twitch murine (mouse) muscle fibers.

To test our hypothesis, my colleagues isolated and suspended soleus and EDL muscle fibers on an apparatus where the muscles could be maintained in conditions that were as near physiological as possible. The muscles were submerged in physiological salt solutions that replicated blood plasma. This solution was gently oxygenated and the pH, partial pressures of oxygen and carbon dioxide, and temperature were regulated. Two sets of muscles were tested: those submerged in a regular physiological salt solution, and those placed in a similar solution containing acetaminophen. Once muscles were in a resting steady state (judged by baseline resting tension), they were stimulated to exhaustion. This means a repetitive series of electrical stimuli were administered until contractile strength reached a maximum and then progressively dwindled to near zero. In these pilot experiments, we made four observations about the acetaminophen-treated muscles: (1) they reached a greater maximal contractile strength during stimulation (just like the bullfrogs' gastrocnemius muscle), (2) they fatigued less rapidly, (3) their contractile strength upon fatiguing was considerably greater than that of the untreated muscles, and (4) they recovered more rapidly from fatigue.

Although my colleagues have accepted positions in another institution, we have agreed to continue this investigation. Based on the results in mice, we believe acetaminophen might affect calcium release and/or the contractile machinery itself (function of contractile proteins such as troponin, tropomysin, and others). Time and experimentation will tell.

11

Integrated Physiological Responses

At the turn of the twenty-first century, all things physiology were about integration. This means understanding mechanisms from molecular to whole animal levels. Such knowledge allows science to be quickly transferred from the laboratory bench to the hospital bed (translational physiology). Regulation of circulating blood volume is one such topic.

Hypovolemic Hypotension

Hemorrhagic shock such as occurs in combat or in automobile accidents reduces the circulating volume of blood and causes hypovolemic hypotension. This has far-reaching consequences for the body and survival. It also exemplifies how multiple organs and systems respond in a coordinated fashion when the physiological homeostasis of the cardiovascular system is upset. In the experimental physiology laboratory, hemorrhagic shock and hypovolemic hypotension have been among the time-tested classic experiments used to teach advanced students about the integration of organ systems physiology.

Imagine an adult whose organ systems are in the physiological steady state under resting conditions. Assume his heart rate, cardiac output, and mean systemic arterial blood pressure are 75 beats per minute, 5 liters per minute, and 100 mmHg. All is well because his cardiovascular system is in a state of homeostasis. Now consider the same person after an emergency crisis brought on by the precipitous loss of 30 to 40 percent of his circulating blood volume. There are two phases of response to the crisis. The first could be characterized in the early seconds to minutes after the onset of hypovolemia. The second, a considerably different picture, would take place hours to days later and only if the victim was still alive.

Among other changes, the physiological responses to the early stages of hypovolemic shock involve cardiovascular reflexes that make emergency adjustments to the crisis. The second phase involves delayed, more long-term adaptations by the kidneys, the adrenal glands, and higher brain structures such as the hypothalamus. During these two periods, this person's body will have gone from a state of physiological harmony—one where vitals signs are at steady state and equilibrium, gradients and feedback control are in balance, and homeostasis is maintained—to a near-death state of disharmony and commotion.

Baroreceptor Reflexes and Cardiac Output

In the above example, hypovolemia occurs after blood volume has declined from approximately five to about three liters. Correspondingly, blood pressure will plummet, from 100 mmHg to perhaps as low as 40 or 50 mmHg. This will cause havoc for blood perfusion of most organs and tissues. That is, the homeostasis of oxygen and nutrient supply and demand will be upset, and the survival of organs—the victim's life—put in jeopardy.

The volume of blood in the vascular compartment is one of the physical determinants of blood pressure. When blood volume declines so does blood pressure, and when blood volume increases blood pressure follows. A reduction in blood volume coupled with hypotension are sensory signals that activate several cardiovascular reflexes. One of the most important of these is the baroreceptor reflex. The sensory receptors of this five-component reflex arc are located, bilaterally, in the carotid sinuses near the bifurcation of the common carotid arteries into their internal and external carotid branches. Carotid sinus baroreceptors—also called pressure receptors, pressoreceptors, or mechanoreceptors—are designed to detect both mean systemic arterial blood pressure and beat-to-beat pulsatile blood pressure. Deviations in either or both of these from their physiological values either activate or inactivate the receptors.

In our subject, as mean and pulsatile blood pressures decline, stretch or deformation of the walls of the carotid sinuses is reduced. This results in a reduced frequency of afferent sensory signals being transmitted to the brain stem each second. The action potentials are conducted via the sinus or Hering nerves to the pons and the medulla oblongata. These two brain stem structures and adjacent tissues constitute the location of many of the body's important regulatory control centers, such as the cardiovascular, respiratory, thermoregulatory, and satiety centers.

The cardiovascular control centers of the medulla subserve both cardiac and vascular functions. When activated, one set of neurons within the cardiac center stimulates an increase in heart rate and is called the cardioacceleratory center. Other neurons cause a decrease in heart rate when stimulated and are referred to as the cardioinhibitory center. Similar classifications apply to the

vasomotor centers that are divided into vasopressor (which constricts blood vessels) and vasodepressor (dilates vessels) regions.

The efferent motor action potentials coming out of the cardiovascular control centers to the heart and vasculature are carried by both sympathetic and parasympathetic branches of the autonomic nervous system. Under normotensive conditions, both sets of autonomic nerves discharge tonically, at a rate that depends on the individual. This means that under normotensive conditions, the autonomic nervous system contributes to the basal resting heart rate, blood pressure, and other cardiovascular variables that influence heart rate and blood pressure.

As pressure in the carotid sinuses falls, the frequency of action potentials traveling to the control centers is reduced. Subsequently, tonic discharge of motor neurons originating in the cardioinhibitory center is inhibited while neurons of the cardioacceleratory center are stimulated. The net effect of these changes is an acceleration of heart rate. Thus, our subject's heart rate, which was 75 beats per minute in the basal state, might now increase to 125 or even 150 beats per minute in the acute stage of hypotension. Other sympathetic nerves innervating the heart are activated to cause concomitant increments in cardiac contractility. Tachycardia is part of the more extensive reflex response that is designed to help blood pressure in the sinus and elsewhere return to normal. The reflex includes changes in peripheral resistance, venous blood volume, and cardiac output.

The physiological state of one's baroreceptor reflex can be tested using a variety of clinical/experimental techniques. For example, the clinician can place an inflatable collar around the subject's neck and inflate the collar to a pressure that will occlude blood flow through the carotid arteries. This is analogous to an inflatable cuff being used to occlude blood flow to the forearm when measuring blood pressure. When the neck collar is inflated and carotid blood flow is reduced, pressure in the carotid sinus will fall. The baroreceptors will detect this hypotensive signal and will cause a reflex increase in heart rate that the clinician can monitor.

Cardiac output is determined directly by the product of heart rate and stroke volume. There are other indirect determinants of cardiac output, including arterial blood pressure and cardiac contractility, but the net effect of an increase in heart rate and an accompanying increase in cardiac contractility or stroke volume is an increase in cardiac output. Under resting conditions and when our subject had a circulating blood volume of 5 liters, his cardiac output was about 5 liters per minute. That cardiac output was produced by a heart rate of 75 beats per minute and a stroke volume of about 65 to 70 milliliters per cardiac cycle. During the acute stage of hemorrhage, a reduction in blood volume will cause a corresponding decrease in cardiac output. However, as described above, the reduction in cardiac output will be opposed, reflexively, by an increase in heart rate and cardiac contractility whose combined effects are to minimize

any decrement in cardiac output. Thus, the cardiovascular control system was designed to maintain homeostasis of cardiac output and therefore organ perfusion. As a rule, cardiac output in such a victim of hypovolemia will be sustained near baseline levels or it might be modestly decreased. Rarely, however, will cardiac output increase above basal levels under these conditions.

Vasomotor Responses and Blood Pressure

Concurrent with the acute changes in heart rate, cardiac contractility, and cardiac output is a coordinated and efficient vasomotor response. The same sensory afferent signals that led to changes in heart rate and cardiac contractility will activate vascular responses. The vascular responses originate in the vasopressor and vasodepressor centers of the brain stem's medulla. Prior to hemorrhage, physiological comparators in these control centers were accustomed to evaluating afferent inputs reflecting systemic mean arterial pressure of 100 mmHg and pulsatile pressure of about 40 mmHg. During the hypotensive crisis, the comparator experiences a mean arterial pressure of no more than 60 mmHg and a pulsatile pressure much lower than 40 mmHg. These two sets of values—normotensive verses hypotensive—get compared, and adjustments are made by relaying efferent action potentials to the systemic vasculature, causing them to constrict.

Any vasodilator response in the presence of reduced blood volume would only complicate matters by causing a further decline in systemic arterial blood pressure. Thus, as the vasopressor centers are stimulated, the vasodepressor centers are simultaneously inhibited. The consequences of increased vasopressor action potentials include arterial, venous, and arteriolar vasoconstriction. Increased arterial vasoconstriction will prevent a further decline in systemic arterial blood pressure. In conjunction with the maintenance of cardiac output (relative to its nadir during hypovolemia), arterial vasoconstriction will begin to reverse hypotension thus helping to restore blood pressure to its pre-hemorrhage, baseline levels.

Centrally mediated vasoconstriction of large veins effectively shifts the distribution of circulating blood volume from the venous to the arterial side of the circulatory system. Large veins act like blood reservoirs. Under normal cardiovascular conditions, about 60 percent, or 3 of 5 liters, of total blood volume are found in the veins. When the veins constrict, their compliance decreases and they cannot accommodate such a large volume of blood. The excess volume must be shifted to some other segment of the vascular system, and the fortunate recipients under conditions of hemorrhage are the arteries. This physiological transfusion of a large fraction of blood to the arterial vasculature elevates arterial blood pressure and helps sustain life.

A final component of the systemic vascular response to acute carotid sinus hypotension is the change in capillary Starling forces. These microcirculatory

adjustments are brought about by corresponding alterations in upstream and downstream hemodynamics. The relevant up- and downstream sites are the precapillary arterioles and postcapillary venules. Homeostasis between hydrostatic and oncotic pressures inside and outside the capillary, the segment of vasculature interposed between arterioles and venules, ultimately determines net water loss or gain by the entire vascular compartment. The coordinated vasomotor output of the brain stem control centers during hypovolemic hypotension leads to arteriolar vasoconstriction and simultaneous venular vasodilation. The net effect of such changes is a reduction in capillary blood volume and, correspondingly, a drop in capillary hydrostatic pressure. This causes a decrease in loss of water from the capillaries or the preservation of circulating blood volume. Net loss of water from the capillaries, seen only under pathophysiological conditions such as inflammation and edemogenesis, promotes a proportional reduction in arterial blood pressure. Prevention of water loss negates this action.

The effectiveness of the baroreceptor control system to maintain cardiovascular homeostasis is determined by the gain of the negative feedback. In our example, the negative feedback to the brain stem control centers was a reduction in blood pressure inside the carotid sinus baroreceptors. To understand the concept of physiological gain, imagine that a large volume of blood is transferred into a person whose carotid sinus receptor control system is not functioning properly and whose arterial blood pressure rises from 100 to 175 mmHg. Now imagine that the same volume of blood is transfused into the same person when his baroreceptor control system is working and pressure rises by only 25 mmHg. In the latter condition, his feedback control system has caused a correction of −50 mmHg, that is, an adjustment from 175 to 125 mmHg. There still remains a correction of about 25 mmHg, called the *error*, which means that the control system, even though functioning, is not operating at 100 percent effectiveness. The gain of this person's functional control system is gain = correction/error. In this example, the correction is −50 mmHg and the persistent error is 25 mmHg, so the gain is −50 divided by +25, or −2. This means that a disturbance that decreases or increases this person's blood pressure does so only by one-third as much as would happen if the baroreceptor system was not working.[1]

Renal Response and the Renin-Angiotensin System

In addition to the carotid sinus baroreceptors, there is a renal baroreceptor system strategically located at the junction of the renal afferent and efferent arterioles where they come in close proximity to the distal tubules of the nephron. As mentioned earlier, this structure is called the juxtaglomerular apparatus or JGA. Specialized vascular smooth muscle cells found in the walls of the renal afferent arterioles of the JGA are designed to detect renal arterial blood pressure and changes in it. Any decrease in renal pressure such as would occur with

hemorrhagic hypotension causes these specialized cells to release renin. Renin is a proteolytic enzyme that, when released into the systemic circulation, comes in contact with a macromolecule called angiotensinogen. Angiotensinogen is produced by the liver and is a physiological component of the plasma. When renin and angiotensinogen meet, renin clips off a small chain of ten amino acids called angiotensin I or AI. Angiotensin I is a physiologically inert decapeptide.

After its formation and while it circulates in the plasma, AI comes in contact with an enzyme found in the lungs called angiotensin-converting enzyme or ACE. Angiotensin-converting enzyme is an integral component of the vascular endothelial cells that line the pulmonary vasculature. This enzyme binds to AI and removes two more amino acids, leaving an octapeptide called angiotensin II or AII. Angiotensin II is anything but physiologically inert; that is, AII has multiple physiological actions. Some of its most important target tissues include arteriolar vascular smooth muscle cells, the adrenal cortex, and the hypothalamus.

After the renin-angiotensin system, as it is called, is activated subsequent to hypovolemia, AII binds to its receptors on vascular smooth muscle cells, causing the cells to contract. This leads to vasoconstriction. During a hypotensive crisis, AII-mediated vasoconstriction helps prevent further decrements in systemic arterial blood pressure as well as restore a normotensive state. AII is one of the most active naturally occurring physiological vasoconstrictors identified to date. The field of research on this topic is broad and over the past few decades has lead to an entirely new branch of pharmacology and medicine: the development and use of ACE inhibitors to help control hypertension.

In addition to its vasoconstrictor effects on vascular smooth muscle, AII stimulates production and release of aldosterone by the adrenal cortex. Aldosterone is a mineralocorticoid or steroid hormone that helps regulate the balance of body water and salts. One of its main actions is on the tubules of the nephron where it promotes increased reabsorption and therefore decreased excretion of sodium ions. When sodium ions are reabsorbed, water molecules follow passively. Thus, the renin-angiotensin system, acting through the release of aldosterone, minimizes loss of body water by renal excretion and thereby conserves body water and circulating blood volume. During a hypotensive crisis, this translates into a salutary effect on systemic arterial blood pressure.

Finally, AII has been shown to promote production and release of arginine vasopressin (AVP) or antidiuretic hormone (ADH) by the hypothalamus and posterior pituitary gland. Arginine vasopressin is not only another potent vasoconstrictor, it also increases the numbers and activities of aquaporin channels in renal tubular epithelial cells. By its vasoconstrictor properties, seen mainly at the arterioles, AVP increases blood pressure. Its renal tubular properties maximize reabsorption of water thereby minimizing water diuresis or renal loss of water by excretion. This also helps to expand the circulating blood volume thus restoring blood pressure to normotensive levels.

Blood Flow and Its Distribution and Redistribution during Hypotension

The loss of circulating blood volume creates a crisis for the body much in the same way that a loss of revenue creates a crisis for government. The loss could be evenly distributed across all parts, or some parts could be affected disproportionately to others. In the cardiovascular steady state when blood volume is constant, all things cardiovascular are in homeostasis. Specifically each major organ system such as the heart, brain, kidneys, liver, and gut get a share of the 5 liters per minute of blood flow that is proportional to their metabolic needs and oxygen demands. We call this relationship tissue oxygen supply and demand, and the two are in a state of physiological harmony when cardiovascular homeostasis prevails.

The distribution of blood flow to the various organs and tissues of the body is different during a hemorrhagic crisis. The mammalian cardiovascular system was designed to redistribute blood flow away from less vital and toward more vital tissues during hypovolemic crises. Specifically, the heart, brain, liver, and adrenal glands are more critical to survival than are kidneys, skin, skeletal muscle, and gut. In the physiological normotensive state, the kidneys and gut, for example, receive a combined 50 percent of the blood flow. The brain and heart combined receive less than 20 percent. At the peak of a hypovolemic crisis, it is not unusual for the kidneys and gut to receive less than 10 percent of the available cardiac output while the heart and brain receive more than 20 percent. In terms of the actual volume of blood flow these vital organs receive each minute (ml/min), this still might not be enough to maintain CNS and cardiac function when large volumes of blood are lost. After the crisis and only when circulating blood volume is completely restored will the pre-hemorrhagic allocation of blood flow be restored to all the organs.

Long-Term Responses to Hypotension

Each of the cardiac and vascular events described above is initiated within seconds or minutes after the onset of hemorrhage. Other mechanisms take longer to institute. These play a more important physiological role after hours, days, and even during a lifetime. Preeminent among them is the protracted homeostatic regulation of body water and body sodium.

Including the vascular compartment, sodium ion (Na^+) is the most abundant cation in the extracellular fluids. It is responsible for about half of the osmolarity there. Extracellular osmolarity determines intracellular water volume, and sodium largely determines extracellular osmolarity. Therefore, extracellular sodium importantly determines intracellular water, the largest of all the water compartments of the body. The ability of water to move from the intracellular spaces and into the vascular compartment means that intracellular water is a

vital resource in times of hypotensive crises. Sodium's ability to influence this can hardly be overstated.

The long-term physiological regulation of blood pressure is more importantly determined by sodium balance than by any other contributing entity. Therefore, the student must understand salt and water balance in order to fully appreciate the cardiovascular system and the kidney's contributions to its regulation. If the student's daily consumption of salt in all forms exceeds his daily excretion of salt, he is said to be in a state of positive sodium balance. Conversely, if one's daily excretion of salt exceeds her daily consumption of salt, she is in a state of negative salt balance. Over the long-term, both conditions are unhealthy and can have a devastating influence on blood pressure and its regulation. The more salt there is in the vascular compartment, the more water there will be. The more water, the higher the blood pressure.

Interesting and clinically applicable physiological experiments have been done in this arena for nearly half a century. For example, if one places an animal on a diet that includes excess salt, after several days the animal's weight will increase (due to consumption and retention of excess water). Osmoreceptors in the hypothalamus as well as thirst sensors in the mouth encourage the animal to drink more water to compensate for the increased load of salt in the body. If the animal is sustained on the high salt diet within several days to a few weeks, systemic arterial blood pressure will begin to rise. The animal becomes hypertensive. The experimentally induced hypertension can be corrected by reducing the amount of salt in the animal's diet. Return of blood pressure to a normotensive state also takes several days to a few weeks.

Similar experiments with similar results have been done in humans. Unfortunately, essential hypertension, or elevated blood pressure of an unknown etiology, constitutes the vast majority of all cases of high blood pressure in developed nations: we do not know the cause. The outcome of any treatment, therefore, is uncertain. One of the best ways to control arterial blood pressure is through one's lifestyle. Having a healthy diet that is low in salt and fats, avoiding habits such as smoking and chewing tobacco, and being dynamically active on a regular basis are proven measures that yield health, vitality, and longevity. Obesity and excess weight are more highly correlated with hypertension, elevated circulating lipids, and diabetes than any other physiological parameter. The reader should take this to heart (pun intended) and begin practicing it yesterday.

Remember that physiology is the basis of medicine, historically and today. By understanding physiology and medicine's dependence on it one should live a qualitatively healthier and more productive life. As the student's knowledge of physiology increases, so should her ability to give good advice to future patients and clients. Those who are wise will listen to this advice and will adjust their lifestyles accordingly. Here is to our marvelous bodies and to the physiological knowledge that will help the student, her patients, and clients take better care of them.

12

For the Record

It has been said that records are made to be broken. Of course this has general reference to sports and athletics. Still, it can apply to physiology, to your own health records, and to the records you teach your patients and clients to keep.

Physiology, like all sciences, is a science of record keeping. Some of the records are known as data. Above all else that they do, physiologists are first and foremost writers. Through the centuries, they have collected their records in different forms such as handwritten on paper, ink-drawn polygraphs, heat-inscribed tracings, and electronic files. Physiologists call the primary source of their records original data. After several weeks of experimentation, the physiologist's original data begin to accumulate. Before they can publish their observations, physiologists have to extract selected pieces of information from each original record. That information is reduced, subjected to rigorous statistical analysis, then transformed into figures, tables, images, and a textual manuscript. The manuscript is submitted to a scientific journal where it gets reviewed by peers, rejected by editors, resubmitted by authors, and finally published and made available to the interested public. Records and the reports they lead to are an indispensable component of physiology and life. The student should learn this early and practice it continuously.

When my young family and I were struggling on an assistant professor's salary, an older friend and colleague gave us a used car. As helpful as the second car was, my friend gave me another gift that was almost as important. It was a file folder filled with all the original service and maintenance records for that car. These records became indispensable to me when I assumed responsibility for this car. The folder and my friend's act taught me to keep records on all my future cars.

Imagine that as a health care provider you agree to take on a new patient. Suppose that this patient, during his first office visit, hands you a file folder

containing all his past medical records. Having this information to supplement what the patient describes as his current health would be immeasurably important to you as his new health care provider.

Forwarding medical records to specialists in advance of appointments with them can save time and improve their diagnostic abilities. I was once referred to an orthopedist to evaluate a shoulder injury. First in the process was an MRI, and I asked the technician to make a copy of all my images so I could mail them to the orthopedist. Then he was able to review my records before he saw me.

I believe one of the wisest lessons a patient or a patient's health care provider can practice is to learn how and when to keep and to evaluate medical records. To teach this is one of your jobs as a practitioner. I also believe that records kept in the clinic should be made available to the patient for keeping at home. Such a trained patient will be motivated to refer to these periodically. He will thereby become a better patient because he is more educated and more health conscious.

In 1976 my family and I joined one of the first health maintenance organizations (HMOs) in central New Jersey. After a couple of changes in ownership, we were notified that the physical facilities would be moving to a new location. I was concerned about my medical records, so I made an appointment to photocopy them. When I reviewed them, I discovered records of pregnancy and maternity. Obviously they were not mine. Someone else's records had been misfiled in my folder. I wondered how many of my records had been mistakenly filed in the folders of other patients. I also began wondering if data for my blood chemistry and blood pressure had been mistakenly recorded onto someone else's chart and vice versa. Nowadays physicians see patients at the clip of three to five per hour. Nurses and others are interposed between patients, and prescriptions and notes are often written in haste. Whatever the circumstances for the medical personnel, however, no one should be more interested in the accuracy of records than the patient. Her physician or other health care provider should be the next most interested party. The above experience convinced me of the need to keep my own medical records, which I have done ever since.

Records can be kept in multiple forms. I encourage the student to find a form that fits her style and to begin using it immediately. Tabular formats have proven most useful for me. I try to record dates; names of doctors, nurses, or technicians who collect the data; and physiological variables such as body temperature, weight, and heart rate. Then on a spread sheet I arrange that information into rows and columns so I can apply simple statistical formulas at will. Table 12.1 is one such example I have developed for collecting my own raw clinical data. It can be adapted to the student's needs or for the clinician's purposes.

TABLE 12.1

Sample chart for recording heart rate and blood pressure over time

Date	*Physician (nurse, PA)*	*Clinic/ Laboratory*	*Heart Rate (cpm)*	*Systolic Blood Pressure (mmHg)*	*Diastolic Blood Pressure (mmHg)*
n					
mean					
s.e.m.					

Blood Pressure and Your Health

From the student's earliest experience as an athlete, scout, or summer camper, he knows that a physical or medical examination includes, among other things, measurements of one's blood pressure, heart rate, and body temperature. Data for these variables are always recorded on a chart. The wise student will ask for a copy of the chart. The concerned practitioner will volunteer to give a copy of the data to the patient. If copies are not available, the patient should ask the nurse/doctor to record the values on a piece of paper and to give them to her. For me as a physiologist, there are two good reasons for doing this. First I desire the numbers for my personal records. Secondly, when it comes to my blood pressure, I have minimal confidence in the accuracy of individual measurements made in haste in physician's offices, but I do believe in statistical averages.

On occasion a nurse will place the inflatable cuff on my right arm and tell me my blood pressure is, for example, 130 over 85. I will ask her to place the cuff on my left arm and take the measurements. Not often but occasionally the two sets of values are different. The two sets of values should not be different, and if there is a difference, it is caused by human error. Either the nurse did not listen through the stethoscope with equal care, or she did not apply equal precision when noting the systolic and diastolic pressures on her mercury manometer. There are few if any other correct possibilities for explaining why blood pressures in the left versus right arms of the same person on the same occasion would differ. Thus, when that particular nurse measured my blood pressure I did not leave the clinic believing the exact values she had recorded. However, I added her values to those taken by other nurses to compute a statistical mean and standard error of the mean for that set of numbers. The outcome, I believe, is more representative of my actual systolic and diastolic blood pressures.

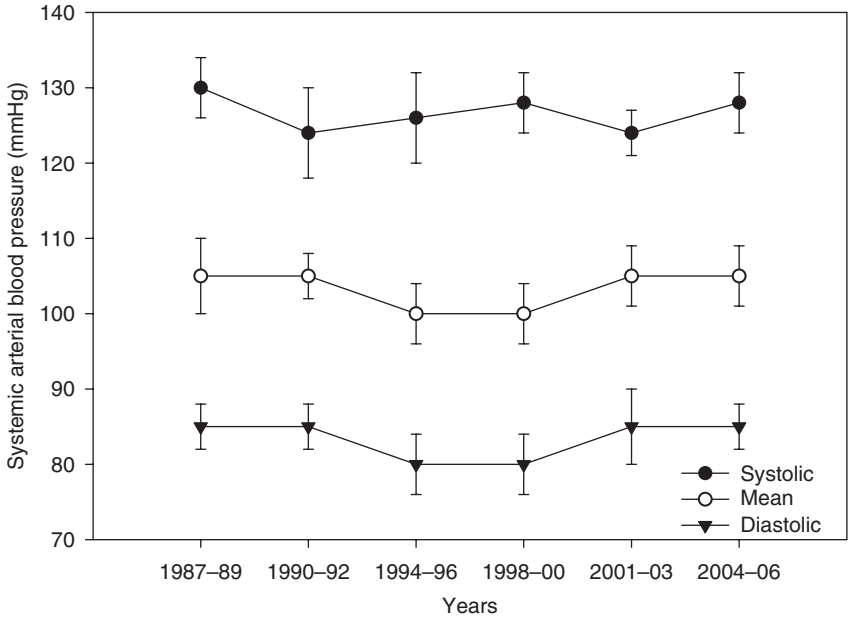

FIGURE 12.1 Author's blood pressure (y-axis) plotted over time (x-axis, years). Data are means (symbols) plus or minus one standard error of the mean (vertical bars attached to symbols) for systolic (closed circles), mean (open circles), and diastolic (closed triangles) pressures collected between 1987 and 2006. Each mean represents several measurements.

It is next to impossible for a nurse or anyone to record a patient's blood pressure that is accurate to the nearest 1 mmHg using indirect techniques. Use of an inflatable cuff and stethoscope is an indirect technique. Obtaining accurate blood pressure values such as 124 over 81 or 172 over 94 is unlikely. Thus, I always round the numbers in increments/decrements of five. In this example, I would round to 125 over 80 or 170 over 95. The averages of these numbers monitored over several months to several years will be more representative of the actual blood pressure. The only way blood pressure can be measured accurately is by use of direct techniques. This means implanting a catheter in an artery or vein and attaching it to sensitive electronic equipment designed to measure pressures. Of course, this is an invasive technique and cannot be done routinely in clinical settings. Even when pressures are measured directly, it is best to round them to the nearest 5 mmHg.

Figure 12.1 is a sample of my arterial blood pressure recorded at different clinics during a period of about thirty years. I recorded the numbers and averaged them statistically. The open and filled symbols represent means for several different measurements during the two-year periods of time shown on the x-axis. The vertical bars associated with each of the symbols are called the standard errors of

the mean (s.e.m.) and reflect the dispersion of the data or physiological variability. Two or three things are worth noting in this figure. First, my data mirror normal values for the population of American males of my age collected over time. That is, there is nothing outstanding here that would cause concern for my physician or me. Secondly, my systolic blood pressure (top line, closed circles) averaged about 125 or 130 mmHg during that period of time and was below the 140 mmHg that defines systolic hypertension or elevated blood pressure during systole. Similarly, my diastolic blood pressure was about 80 or 85 mmHg on the average. This is also below the danger value of 90 mmHg used by the American Heart Association and other organizations to define diastolic hypertension or elevated blood pressure during diastole. Still, my systolic blood pressure in the past several years reveals a trend towards the 140 mmHg that defines hypertension. This is cause for me to be concerned about those things that influence systolic blood pressure. I have not begun smoking, so that cannot be a cause. Perhaps I have not been as vigilant in watching my diet and exercise program as I should have been.

Keeping tabs on arterial blood pressure is an important exercise in the care and maintenance of one's health. Elevated arterial blood pressure is one of the hallmarks of cardiovascular disease in western societies. I have already explained the physiological determinants of diastolic and systolic arterial blood pressure and why elevated levels are dangerous (see chapter 5). From that discussion, readers will be able to understand how what they eat and do over an extended period of time affects blood pressure and health.

Blood Lipids and Physical Activity

Another important record to keep is one's circulating concentrations of fats or blood lipids. In the mid 1980s through the early 1990s, there was a coordinated effort by the federal government and other public and private health agencies to increase public awareness of diet, blood lipids, and health called the National Cholesterol Education Program. This came about the time the Nobel Prize in physiology was awarded to Michael Brown and Joseph Goldstein for their work on cholesterol and its metabolism (1985). Today public health and nutritional messages focus on the dangers of trans fats. For example, in December 2006, New York City banned use of trans fats in food preparation in all restaurants there. Soon after, other cities began to follow suit.

During the earlier cholesterol campaign, I began monitoring my own blood lipids on a regular basis. Every time I went for a physical examination, I asked my doctor to include blood lipid screens. Specifically, the lipids I am most concerned with are triglycerides (TRG), low-density lipoproteins (LDL), high-density lipoproteins (HDL), and total cholesterol. The LDLs and HDLs are often referred to as "bad" and "good" cholesterols, respectively. Total cholesterol comes from

the summation of LDL, HDL, and a couple other classes of circulating lipoproteins, most notably the intermediate-density lipoproteins (IDL) and the very-low-density lipoproteins (VLDL). From a dietary basis, TRGs are the most important because they are the dietary source of the bulk of fat in our circulatory systems.

In the bloodstream, fats are frequently bound to other molecules such as proteins; they are not free in the circulation. When a lipid is bound to a protein and transported in this fashion, it is called a lipoprotein. Lipoproteins are important because the lipid portions get released in the tissues including the walls of blood vessels. In the walls of blood vessels, they accumulate as fat and mineral deposits, eventually becoming atherosclerotic plaques. Such plaques tend to narrow the intraluminal diameter of blood vessels. As the inside diameter of the vessel narrows, blood flow is impeded and the prospects of blood clotting increase. Blood clots, also known as emboli, thrombi, or thrombo-emboli, as well as fragments of atherosclerotic plaques can dislodge from vessel walls and flow downstream where they become entrapped by any blood vessels that have an inside diameter narrower than the diameter of the clot or plaque fragment. This is usually though not always the cause of myocardial infarction and sudden coronary death as well as cerebral ischemia, infarction, and stroke.

Figure 12.2 illustrates my blood lipid levels monitored during a twenty-five-year period. In each case the open and closed symbols represent the averages for several separate clinical measurements of that particular variable. The vertical lines passing through the symbols represent the s.e.m. This record contains information only for total cholesterol (total-C), HDL (HDL-C), and LDL (LDL-C). Remember that total cholesterol is the sum of LDL, HDL and to a lesser degree IDL and VLDL. (My triglycerides are not shown in this figure but I have records for them just the same.) As in figure 12.1, there are several things to be gleaned from these data. First, note the curve illustrating my total cholesterol. The highest mean value is about 150 mg/dl and the lowest is about 130 mg/dl.[1] To put these data in context, the average value for total cholesterol among graduating high school seniors in the United States is about 150 mg/dl. Many of these students are physically active during their high school years. They play football, basketball, soccer, softball, baseball, lacrosse, and a variety of other sports. In addition, many have parents who help regulate or otherwise monitor the daily diets of their children. Granted, a growing number of American teenagers and their parents do not fit this mold of activity, diet, and childcare. Still, my circulating concentrations of total cholesterol are equal to those of fit graduating high school students. My numbers are equal to theirs, in part, because I have led an active lifestyle over the course of fifty years.

Sadly, many of these youth, including students in the health sciences, transform into sedate college juniors and seniors. They leave the care of their parents, their diets get neglected (the freshman fifteen), physicians' examinations

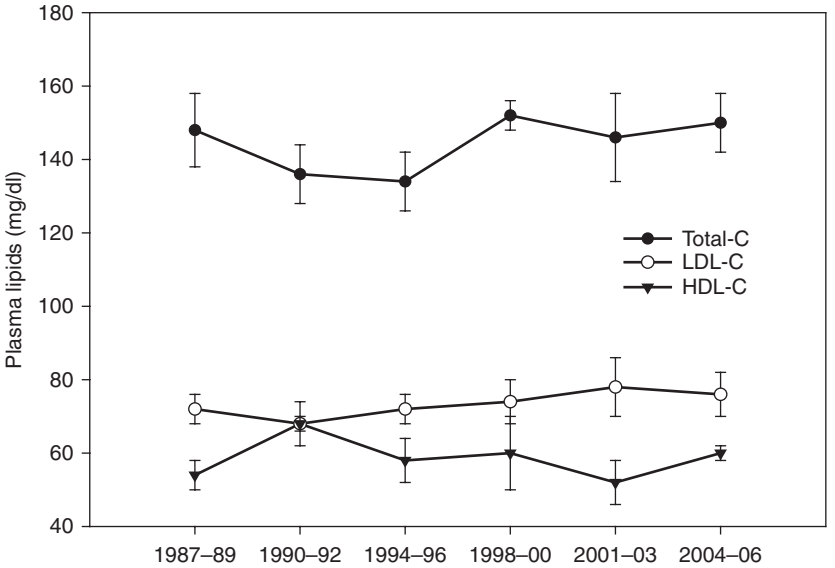

FIGURE 12.2 Author's circulating blood (plasma) lipids (y-axis) plotted over time (x-axis, years). Data are means (symbols) plus or minus one standard error of the mean (vertical bars attached to symbols) for total cholesterol (closed circles), low-density lipoprotein cholesterol (bad cholesterol, open circles), and high-density lipoprotein cholesterol (good cholesterol, closed triangles) collected between 1987 and 2006. Each mean represents several measurements.

are not scheduled, and they replace a physically active behavior with a sedentary, late-night, party-centered lifestyle. Within a relatively few years, their circulating concentrations of cholesterol have risen significantly above 150 mg/dl. This is a major health concern for at least two medical reasons. First, circulating concentrations of blood cholesterol are most closely correlated with body weight. That is, the heavier one is, the higher his cholesterol levels. The freshman fifteen refers to the average number of pounds of weight American teenagers put on during their first year in college. If that rate continued, students would gain in excess of fifty pounds during their four years of college. Secondly, high levels of cholesterol put one at risk of early death from cardiovascular disease. To help guide Americans in practices of good health, the American Heart Association and other agencies have established guidelines for cholesterol and other lipids in the blood. A range of 200 to 239 mg/dl is defined as borderline high risk, and values greater than 240 mg/dl are considered high risk. Barring genetic predispositions that prevent an individual from influencing her circulating concentrations of lipids, the best rule of thumb is to keep total cholesterol, LDL cholesterol, and triglycerides as low as possible. This can be done mainly through eating a healthy variety of foods on a daily basis and by establishing regular programs of daily physical activity.

From figure 12.2, note that my plasma HDL levels are about 50–60 mg/dl and my LDL levels are about 70–80 mg/dl. While not shown here, there have been times during those twenty-five years when my HDL levels were higher than my LDL levels. This happens mainly when I am involved in intense regular exercise. Another general rule of thumb is to keep the HDL levels as high as possible and the LDL levels as low as possible. Diet and exercise are again important although individual results will vary from person to person and from time to time even in the same person.

From the late 1980s to the early 1990s, my HDL concentrations rose while my LDL levels dropped until the two met. This is unusual but good and there is a reasonable explanation for the phenomenon. During that time, I had a graduate student who was determined and competitive. A new recreation facility had just been built at Rutgers University. My student and I decided to take advantage of these facilities and began playing racquetball together. We agreed to have physical examinations by our physicians including complete blood screens, to play intense competitive racquetball for some weeks, and then to have follow-up physical examinations. What was planned as a one-semester activity turned into several years of healthy competition. My student and I met on the courts three times weekly and played vigorously for nearly an hour. Neither one of us cared much about the etiquette of racquetball. We were more concerned with movement and physical activity on the court. The net result for blood lipids were noticeable elevations in our HDL cholesterol concentrations and reductions in total cholesterol, LDL cholesterol, and triglycerides.

Before and since that time I have engaged in dynamic physical activities including running, cycling, and swimming, and have checked the effects with physicians' examinations and blood screens before and after each period of thirty to ninety days. None were as markedly effective for me and my circulating lipids as the racquetball that I played with my competitive graduate student and friend. I recall that period with great fondness. I suppose that if friends could exercise together, long-standing lipid-lowering health benefits could result. I have not tested this idea further but would be happy to hear from any reader who has.

Finally, I apologize if I have sounded either boastful or condescending. If your blood lipids are not in line with good health, take heart. Without prescription medication you can do something about this. Prepare another table (see table 12.2) and then go for a physical examination including a full blood lipid screen. Ask for a copy of the results. Then begin a self-determined, inexpensive program of weight loss, diet control, dynamic activity, and persistence. After eight to twelve weeks (and when you have shed eight to ten pounds), go for a second physical examination and blood lipid screen. Compare the two sets of results. Most likely you will see positive improvements. These will motivate you to even higher levels of achievement.

TABLE 12.2

Sample chart for recording blood lipids over time

Date	*Physician (nurse, PA)*	*Clinic/ Laboratory*	*Triglycerides*	*Total Cholesterol*	*HDL-C*	*LDL-C*
n						
mean						
s.e.m.						

Blood Cells and Good Health

From the late 1990s, stem cell research has captured the interests of science, medicine, and the public. Stem cells can be collected from several sources including umbilical cord blood, bone marrow, embryos, and adults. Stem cells are often referred to as unipotent or pluri- or multipotent. Unipotent means that a particular line of stem cells is destined to become a single differentiated cell type, for example, a nerve cell, muscle cell, or endocrine cell. Pluripotent means that a particular line of stem cells can become any of several cell types including those just mentioned. If the stem cells are destined to become circulating blood cells, they are referred to as hematopoietic stem cells. Such primordial (ancestral) hematopoietic stem cells can differentiate into red blood cells (erythrocytes), white blood cells (leukocytes), or blood platelets (cell fragments). Their differentiation depends in part on chemicals that direct growth and development such as growth factors. Nerve growth factors (NGF) contribute to the undifferentiated cell becoming a neuron while angiogenic growth factors (AGF) direct differentiation of stem cells into blood vessels. All of this depends critically on the health and well-being of the organs and tissues from which the ancestral stem cells arise.

The concentrations of blood cells in circulation at any point in time are importantly influenced, among other things, by diet, exercise, and lifestyle. For example, we know of communicable diseases that compromise the immune system such as HIV. Compromised immune systems mean poorer health in general. The immune system, in part, is comprised of circulating white blood cells such as lymphocytes, neutrophils, and macrophages. Diet and exercise habits have an effect on immune system health as well although all the mechanisms are not known. At one point in my adult life, I was exercising three or four times per week and donating blood at least quarterly. Then I decided to increase my daily consumption of insoluble fiber via a hot breakfast cereal. After a routine physical examination, my physician reported that my circulating leukocyte

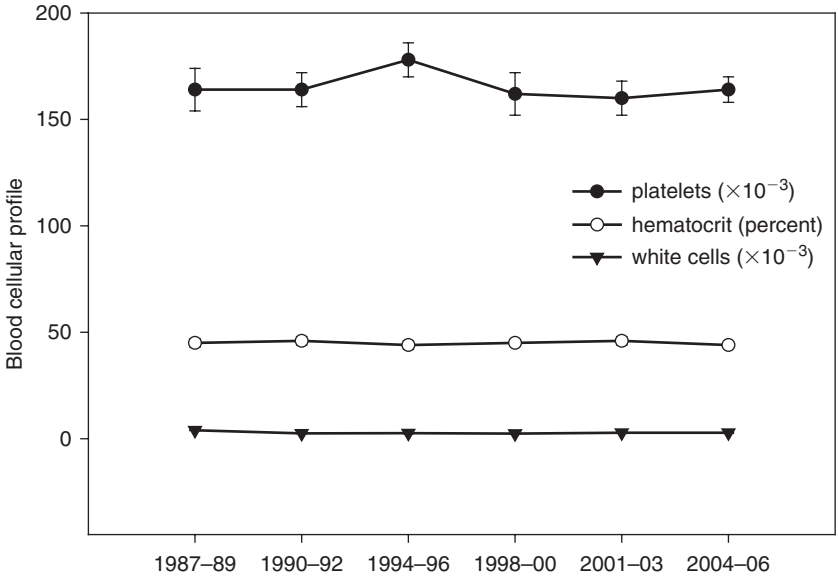

FIGURE 12.3 Author's profile of circulating white blood cells ($\times 10^{-3}$ per microliter of whole blood, closed triangles), platelets ($\times 10^{-3}$ per microliter of whole blood, closed circles), and hematocrit (percent, open circles) (y-axis) plotted over time (x-axis, years). Data are means (symbols) plus or minus one standard error of the mean (vertical bars attached to symbols). Each mean represents several measurements.

levels were on the low side of normal, a condition he called leukopenia. I explained to him what I was doing with regular exercise, high fiber diet, and blood donations, and he suggested additional blood screens and laboratory tests and that I see an oncologist, which I did. Over the course of several months, she did additional tests trying to isolate the cause of my leukopenia. Tests included examinations of my liver, spleen, and bone marrow. She could find no clues to explain the decline in circulating white blood cells.

I told her about my high fiber diet, regular exercise, and quarterly blood donations. She shrugged off the notion of contributions from my diet and exercise but suggested that I discontinue donating blood so frequently. Soon thereafter I also modified my diet to reduce my daily consumption of fiber. During the oncologist's evaluations, I learned that high dietary fiber can interfere with digestion and absorption of iron and other minerals that are important to hematopoiesis. I modified my lifestyle and have continued to monitor my white blood cells, but they have never returned to their preexperimental values. Figure 12.3 illustrates some of my data before, during, and after that period of time. The average adult concentration of leukocytes is about 5,000 cells per cubic microliter (or 5,000 WBC/μl^3). Neither my family physician nor the oncologist were able to tell me what had caused my leukopenia. As a physiologist, I reasoned

that the combination of donating blood, consuming excess fiber, and engaging in moderate to heavy exercise had, by mechanisms that I cannot explain, compromised the white cell-producing capacity of my hematopoietic tissue. Alternatively, my combined diet, exercise, and blood donations might have permanently accelerated the rates of destruction of my white blood cells. If the rate of degradation of a product exceeds its rate of synthesis, then the circulating concentrations of that product will be reduced.

One of the reasons monitoring white blood cells is important is that they constitute an important component of the body's immune system. Reductions in circulating and noncirculating white blood cells potentially mean a compromise in the immune system's capacity to fight disease. With my physician's help, I have continued to keep records of my blood chemistry. If my white cell count declines below the range of 2.8–3.0, I will ask for referrals to see other oncologists. Medicine is a dynamic art, and it continues to change as physiology and other medical sciences advance. Doubtless, for diagnosing and treating paroxysmal leukopenia (an unexplained decline in the circulating concentrations of white blood cells), there are procedures and medications available today that were not available ten or twenty years ago. Moreover, as stem cell research moves forward, new discoveries for treating leukopenia will be made. At the time of this writing (December 2006), Jon Corzine, the governor of New Jersey, had just signed legislation creating the Stem Cell Research Initiative of New Jersey. Hundreds of millions of dollars will be spent to build and staff the Stem Cell Research Institute of New Jersey, a joint effort between Rutgers University and the University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School.

Blood Sugar, Diabetes, and Metabolic Syndrome

While watching a softball game one spring, I noticed a young boy about six or seven years old. While playing catch with his mother, he made several runs to a nearby restroom. Later, in a conversation with the mother, I commented about her energetic little boy. She said something to the effect that all he does is eat, drink, and urinate. I asked her if she had ever heard the phrase the *trilogy of diabetes*. She said no. There are three well-defined symptoms of undiagnosed, therefore untreated, diabetes. From the Greek they are called *polyphagia, polydypsia,* and *polyuria*. The prefix *poly* means "many or multiple." *Phagia, dypsia,* and *uria,* loosely translated, mean "hunger," "thirst," and "urination." In other words, the symptoms of undiagnosed diabetes, which are familiar to any health care professional who has spent reasonable time studying physiology, are excess hunger, excess thirst, and excess urination. The boy's mother said she would arrange for a physical examination for her son and thanked me for the information.

When my sons were young, I was a scoutmaster. Over a couple of years camping, hiking and working with these young men, I noticed a boy about age

twelve who exhibited the trilogy of diabetes. He was also extremely slim and had a pale, sickly skin tone. I mentioned my observations to his mother and father and described the symptoms of diabetes to them. Several weeks later they told me that they had taken their son for a physical and blood work. The doctor had recommended a glucose tolerance test. Subsequently this boy was diagnosed with Type I insulin-dependent diabetes mellitus (IDDM) and was placed on a special diet and regimen of insulin injections.

A glucose tolerance test is one of the definitive measures physicians use for detecting diabetes. The patient is brought to the clinic in a fasting or semi-fasting state and water is given at will but food is withheld for one or two consecutive meals. This means that patient has not eaten for sixteen to twenty-four hours prior to the test. She is asked to drink a flavored liquid containing a high concentration of sugar. Urine and/or blood samples are collected every thirty minutes over the course of two to three hours. On a graph, the blood or urine concentrations of sugar are plotted as a function of time. In a nondiabetic person, blood sugar concentrations rise rapidly to a known level and then decline almost as rapidly. In a diabetic individual, blood sugar levels rise considerably above the level for nondiabetics but decline only slowly. Thus, the shapes of the two curves for diabetic and nondiabetic are characteristically distinguishable. This is the basis of the diagnosis of diabetes. Anyone you know, young or old, who displays the trilogy of diabetes and has not been screened should be.

Diabetes is the most common serious metabolic disease in humans. One variant is Type I diabetes or insulin-dependent diabetes mellitus, which results from an immune-mediated selective destruction of the pancreatic beta cells (those that produce and secrete insulin). It is also called early-onset or childhood diabetes. In the absence of insulin and in the presence of another pancreatic hormone, glucagon, circulating concentrations of glucose and ketone bodies accumulate. In excreting these, the kidneys cause an osmotic diuresis. The pathophysiological consequences that follow can lead to acidosis, dehydration, and death from diabetic ketoacidosis; treatment is to supply insulin by injection. Type II diabetes, or non-insulin-dependent diabetes mellitus (NIDDM), is more complex. The beta cells still produce insulin but are not as responsive to elevated concentrations of glucose as are the beta cells of normal subjects. Moreover, even in the presence of insulin, the metabolism of glucose is compromised. These patients have developed insulin resistance.

The story of the discovery of insulin by Frederick Banting and Charles Best and colleagues is among the great collaborations between experimental physiology and the practice of human medicine. The University of Toronto in Canada became the site of perhaps the most celebrated example of a discovery made in a physiological laboratory that had immediate therapeutic implications. Insulin was discovered, isolated, purified and made ready for clinical use in the brief period of a couple years (1921–1922) by the research team of Frederick Banting, a

professor of physiology, Charles Best, his graduate student in physiology, John James Macleod, a professor of biochemistry, and James Collip, his graduate student in biochemistry. In 1923 Banting and Macleod became the first North Americans to win the Nobel Prize in physiology or medicine. They shared their fractions of the prize with Charles Best and James Collip. I recommend reading the physiological and biochemical experimental components of this story to any student or clinician interested in the relation between physiology and medicine. It is one of the most striking examples of translational physiology that I can think of.

A detailed discussion of the physiology and endocrinology of diabetes is beyond the scope of this book. However, a few points should be made. In normal persons, insulin counterbalances the actions of glucagon by promoting the uptake of circulating glucose by the tissues. In most of us, increased glucose in the blood, such as after ingestion of a meal or during a glucose tolerance test, is accompanied by increased secretion of pancreatic insulin. This does not happen in Type I diabetes where, in the absence of insulin, glucagon can cause sufficient release of glucose to elevate plasma concentrations five- to tenfold. In the presence of excess blood sugar and excess blood ketone bodies, the kidneys must work extra hard to process these metabolites. That causes excretion of osmotically active chemicals and excess loss of water. Ketone bodies are strong organic acids, so their excess can lead to severe metabolic acidosis, the accumulation of excess hydrogen ions (reduced pH) in both the intra- and extracellular spaces. Metabolic acidosis coupled with severe dehydration is a formula for disaster including potential coma and death. These can and do occur if the diabetic is not treated with insulin.

One of the modern scourges of industrialized societies is the obesity epidemic. The increased prevalence of obesity among children is directly associated with an increase in Type I diabetes. An association between metabolic disorders and cardiovascular disease has been known since the late 1940s. In the 1980s, the association became clearer and the term metabolic syndrome was coined to designate a cluster of metabolic factors that put an individual at increased risk of early death from cardiovascular-related disease. The root causes of metabolic syndrome as well as the basic definition depend on who is defining the disorder. For example, the National Cholesterol Education Program defines the syndrome more conservatively than does the World Health Organization. They do agree, however, on a few defining characteristics. One is a plasma triglyceride concentration that is greater than or equal to 150 mg/dl. Another is a waist circumference of 35 to 40 inches (lower numbers for women and higher numbers for men). HDL concentrations of 35 to 40 mg/dl or lower is a third indicator of metabolic syndrome. The World Health Organization is more liberal in regard to blood pressure, as they define the syndrome as 140/90 mmHg or above. The National Cholesterol Education Program's blood pressure values

of 130/85 are too conservative in my view and are discordant with most other professional guidelines, such as those recommended by the American Heart Association, the National Institutes of Health, and the World Health Organization.

The physiological range for glucose either in whole blood or plasma should be about 80–100 mg/dl. Anytime the student or patient has a physical examination and blood work and glucose concentrations are appreciably outside this range, you should seek an explanation. You do this by first analyzing any changes in eating habits or daily routines of physical activity that might have occurred hours or days before withdrawal of the blood sample. As a practitioner, you should take adequate time with your patient to discuss the data. Between the patient and the clinician and any specialists who might be involved, all parties should come to a satisfactory conclusion about the probable causes of the elevated or depressed blood sugar. Moreover, both the student and the practitioner should be aware that technical mistakes can be made in the laboratory. As either a patient or one treating a patient, it is acceptable to ask for a reanalysis of blood glucose. If the source of the problem was identified and corrected between the times of the first and second blood samples, then the second test results should reveal blood glucose levels once again in the range of 80–100 mg/dl. Over the course of several decades and as the student or patient accumulates her medical records, she should note that circulating concentrations of blood sugar have remained relatively stable at these levels.

GLOSSARY

absorptive state. the period during which a meal is being ingested, digested, and absorbed.

acrosome. the head of a mature sperm cell

actin. a muscle protein of relatively low molecular weight; a main constituent of the myofibrillar thin filament

activators. cells, tissues, and organs that are innervated by efferent motor neurons

adenylyl cylase. an enzyme that converts adenosine triphosphate (ATP) to cyclic adenosine monophosphate (cAMP)

adrenocorticoids. steroid hormones produced by the adrenal glands

afferent arterioles. small renal arteries that give rise to glomerular capillaries

agglutination. changes in the surface properties of antigens that cause them to adhere to one another

airborne pathogens. disease-causing agents transported in the air by any means including on pollen, spores, and molds

akinesia. difficulty initiating movement

aldosterone. an important salt-regulating hormone produced by the adrenal cortex

alveolar ventilation. abbreviated VA, difference between minute ventilation and dead space ventilation

androgens. male gonadal steroid hormones

anemia. reduced number of red blood cells in the circulation

angiogenesis. the process of forming new blood vessels

angiotensin I. AI; a decapeptide clipped off of angiotensinogen by renin that is physiologically inert

angiotensin II. AII; an octapeptide created when pulmonary angiotensin-converting enzyme (ACE) removes two amino acids from AI

angiotensinogen. also called renin substrate; a circulating macromolecule produced by the liver that serves as substrate for renin

anterior. in the upright individual, the front of the body

antidiuretic hormone. ADH, a hypothalamic/pituitary hormone that minimizes renal excretion of water

aortic bodies. peripheral chemoreceptors located on or near the aorta

aquaporins. proteins that make water channels in cell membranes

arginine vasopressin. AVP, vasopressin; an acronym for ADH

atresia. loss of female oogonia during embryogenesis

autocrine. a hormone released by cell A that acts on cell A

autoregulation. also called pressure-flow autoregulation; the ability of an organ to maintain a relative constant flow despite changes in pressure

baroreceptor reflex. a cardiovascular reflex that begins at the carotid sinus mechanoreceptors

biconcave disk. the shape of the red blood cell with concavity on both sides of the cell

bilateral. on both sides, such as the kidneys are bilateral organs

bipotential gonads. early in embryogenesis, gonads that have the potential to become either male or female

blastocyst. a spherical mass of cells with fluid-filled cavity that continues to develop inside the uterus

Bowman's capsule. a renal capsular structure containing glomerular capillaries

Bowman's space. the site of formation of ultrafiltrate; space between glomerular capillaries and the interior of Bowman's capsule

bradycardia. slow heart rate, fifty beats per minute or less

calorie. a basic unit of heat (the amount of heat needed to raise the temperature of water one degree centigrade)

capacitation. influence of the ovum (egg) on the sperm cell that enables it (the sperm) to fertilize the egg

cardioaccelerator. an agent or action that increases heart rate

cardioaccelerator center. neurons in the brain stem that cause heart rate to increase

cardioinhibitor center. neurons in the brain stem that cause heart rate to decrease

cardioinhibitory. an agent or action that inhibits cardiac function

carotid baroreceptor. a neuronal structure that detects changes in blood pressure (mechanoreceptor, pressoreceptor, or pressure receptor)

carotid bifurcation. division of the main carotid artery, left or right, into internal and external branches

carotid bodies. peripheral chemoreceptors located on or near the bifurcation of both common carotid arteries

central pattern generator. neurons in the respiratory control centers whose cyclic discharge drives the inspiratory and expiratory cycles

chemoattractants. chemicals released by injured cells or tissues that act as signals for other chemical/cellular/metabolic processes

chemoreceptors. in respiratory physiology, sensory receptors—both central and peripheral—that detect changes in PO_2, PCO_2, and H^+ content of body fluids

chemotaxis. process by which one chemical or molecule signals another to draw near

chronotropic response. a change in heart rate caused by an action or agent; can be either positive or negative

chylomicron. reconstituted triglycerides and other fat products inside mucosal epithelial cells

CICR. calcium-induced calcium release; interaction of dihydropyridine and ryanodine receptors that leads to calcium release

colloid oncotic pressure. the ability of a protein-containing aqueous fluid to attract water and cause pressure

comparator. the component of a physiological control system that compares two sets of values such as actual blood pressure and physiological values for blood pressure

conducting zone. trachea, bronchi, and nonrespiratory bronchioles

conduction. movement of an action potential down an axon (neuron)

controllers. neurons, usually found in discrete locations of the brain stem (and other higher brain centers) that exert regulatory control

cross bridge. an extension of the myosin thick filament that binds to the actin thin filament

dead space. volume of air in the conducting zone; the portion of tidal volume that does not exchange with blood because it fails to reach the respiratory zone

dead space ventilation. air that ventilates the conducting zone but not the respiratory zone

depolarization. departure, in the positive direction, from a negative resting membrane potential

diapedesis. movement of neutrophils from the vascular lumen to the interstitial spaces

dihydropyridine receptors. receptors associated with sarcolemmal calcium channels

dorsal respiratory group. neurons that are part of the medullary respiratory control centers

dromotropic response. a change in atrioventricular conduction velocity caused by an agent or an action; can be positive or negative

E-C coupling. excitation-contraction events in the heart that link membrane depolarization to contractile activity

edema. the swelling of tissues caused by accumulation of extracellular water

effectors. another term for *activators*

efferent arterioles. small renal arteries immediately downstream to glomerular capillaries

electrolyte balance. maintenance of the constancy of sodium, potassium, calcium, and other electrolytes in body water compartments

embryogenesis. growth, development, and differentiation of the maturing embryo

emulsification. the role played by bile acids in breaking down and digesting ingested fats

endocardium. innermost cell layer of the heart; cells that are in contact with ventricular blood

endocrine. a hormone released by a gland then transported by the circulation to a remote site of action

endothelium. cells that constitute the interface between circulating blood and the vessel wall

energy balance. calories consumed are equal to calories expended (burned)

enteric system. gastrointestinal tract or system

epicardium. outermost cell layer of the heart; cells that are in contact with the pericardium

equilibrium. physiological steady state, such as when the rate of inflow equals the rate of outflow

error signal. difference between physiological value and the adjustment made to compensate for a deviation

erythropoiesis. process of forming and releasing new red blood cells

estrogens. female gonadal steroid hormones

exercise proteinuria. protein in the urine that accompanies strenuous exercise

exocrine. a hormone or other product released from a cell into a duct or tubule as a means of reaching its site of action

expiration. the process of exhaling a tidal volume of air

external respiration. exchange of respiratory gases between air and blood, and blood and tissues

extravascular. outside the blood vessel lumen

fast-twitch fibers. muscle fibers with low blood flow and mitochondrial content but that are larger and have more extensive SR than slow-twitch fibers

feedback. physiological information relayed by one component of a control system to another

Fenn effect. in muscle, the relationship between the amount of work done and the supply of ATP to support it

filtration barrier. components of walls of glomeruli that impede or prevent ultrafiltration

filtration fraction. the portion of renal plasma flow (RPF) that gets filtered by the kidneys each minute

folliculogenesis. development and maturation of a primordial follicle

forced expiratory volume. FEV1, the volume of air a patient can forcefully expel from the lungs in one second; a measure of pulmonary function

frontal. in the upright individual, a section that divides the anterior from the posterior

gameteogenesis. growth and development of the male and female gametes (germ cells)

genetic gender. the gender that is determined when the male and female pronuclei meet during fertilization

genotype. the genetic compliment of a zygote

glial cells. cells in the central and peripheral nervous systems that support neurons

glomerular capillaries. renal capillaries that filter blood plasma and are located inside Bowman's capsule

glomerular filtration rate. the simultaneous rate of formation of ultrafiltrate in both kidneys

glomerulus. the component of a nephron that contains the glomerular capillaries

gradient. physiological differences in a variable, such as between arterial and venous blood pressure

growth factors. proteins and peptides that promote proliferation and differentiation of cells and tissues

half life. the time it takes an initial concentration of a hormone to be reduced to half its original concentration

hemangioblasts. primitive stem cells that give rise to blood- and vessel-forming cells

hematocrit. red blood cell fraction of a centrifuged sample of whole blood

hematopoiesis. the process of forming blood cells, particularly red blood cells (see also *erythropoiesis*)

hemodynamics. physiological variables that influence blood flow, such as pressure, resistance, vascular geometry, and viscosity

hemorrhagic shock. state of shock caused by loss of blood volume

homeostasis. physiological balance in a cell, tissue, organ, or whole animal

hyperalimentation. intravenous infusion of energy substrates in a patient that is unable to consume food orally

hyperplasia. increased number of cells in a tissue

hypertension. elevation in blood pressure

hypertrophy. increase in size of a cell/tissue

hypophysectomy. removal of the pituitary gland and/or a tumor thereof

hypotension. reduction in blood pressure

hypovolemic hypotension. reduced blood pressure secondary to an abrupt loss of blood volume

inferior. in the upright individual, a landmark that is towards the feet

inotropic response. a change in contractility of the heart caused by an action or an agent; can be positive or negative

inspiration. the process of inhaling a tidal volume of air

interatrial septum. a wall separating the two atrial chambers in mammalian hearts

internal respiration. use of oxygen by cells and subcellular organelles

internodes. space (distance) between two adjacent nodes in a myelinated axon

interstitial. outside cells and blood vessels, the spaces between the cells and tissues

interventricular septum. a wall separating the two ventricular chambers in mammalian hearts

intra-alveolar. inside the alveoli; used interchangeably with *intrapulmonary*

intrapleural space. potential space between the visceral and parietal pleural membranes

intrapulmonary. inside the lungs; used interchangeably with *intra-alveolar*

intravascular. inside the blood vessel lumen

ischemia. reduced tissue blood flow whether transient or chronic

JGA. juxtaglomerular apparatus; morphologic proximity of key renal structures to each other

lacteals. end lymphatic capillaries that absorb chylomicrons and other fats products

lateral. in the upright individual, the side of the body or moving away from the center line (midline)

luminal digestion. breakdown of food particles inside the lumen of the gut

lymphocytes. immune cells produced in the lymph nodes, released into the circulation and tissues, then recycled to the lymphatic system

lymphoid. originating in the lymphatics

malignant hyperthermia. a condition induced by volatile anesthetics in a surgical setting that affects muscle tone and is potentially life-threatening

margination. movement of cells from central axis of flow to vessel wall in preparation to enter tissue spaces

mastication. the oral processes of chewing, crushing, incising, and ripping food

medial. in the upright individual, the center line of the body such as through the sternum and navel

membrane digestion. also called brush-border digestion; digestion of food stuffs by enzymes that are imbedded in cell membranes

menarch. first female menstrual cycle, marks the onset of puberty

menopause. end of the menstrual portion of the female reproductive life cycle

menses. monthly menstrual cycles in females

microcirculation. smallest of the blood vessels including, among others, arterioles, capillaries, and venules

midmyocardium. region of the heart muscle midway between the epicardium and endocardium

MIF. mullerian-inhibiting factor, an agent secreted by embryonic male gonads that prevents female gonads from developing

minute ventilation. VE, the total volume of air moved in and out of the lungs via both conducting and respiratory zones each minute

morphology. the basic shape or structure of a thing (cell, organ, animal)

morula. also known as *conceptus*; development of the zygote to a mass of twelve or more cells during the first three days postfertilization

motor unit. a motor nerve and all the muscle fibers it innervates

myelin. lipid-like insulating material that circumscribes axons

myeloid. originating in the bone marrow

myocardial infarction. a heart attack; reduced flow of blood to heart wall leading to damage and death

myocardium. muscle cells that compose the wall of the ventricles and atria

myocyte. a muscle cell

myosin. a muscle protein of relatively high molecular weight; the main component of myofibrillar thick filaments

nephron. functional structure of the kidneys consisting of vascular and tubular elements

neuroendocrine axis. organs involved in neuroendocrine regulation or physiology including the CNS, hypothalamus, pituitary, adrenal gland, and target tissues

neutrophilia. an increased number of circulating white blood cells secondary to infection and inflammation

neutrophils. mature white blood cells that can attack and destroy whole bacteria and other pathogens

nonstriated muscles. muscle fibers typically lacking striations; examples include vascular and visceral smooth muscle

obligatory water loss. the minimum amount of urinary water necessary to dilute and excrete unwanted solutes during states of dehydration

osmoreceptors. sensory neurons that detect osmolality (osmolarity), or concentrations of solutes, of body water

osmosis. the influence of charged ions on the movement of water

oxyhemoglobin. hemoglobin with oxygen molecules bound to it

$PaCO_2$. partial pressure of carbon dioxide in arterial blood

$PACO_2$. partial pressure of carbon dioxide in alveolar air

PaO_2. partial pressure of oxygen in arterial blood

PAO_2. partial pressure of oxygen in alveolar air

paracrine. a hormone released by cell P that acts on an adjacent cell Q

partial pressure. a fraction of a total pressure, for example, the partial pressure of oxygen is only a fraction of total barometric pressure

PCO_2. partial pressure of carbon dioxide in an unspecified location

pericardium. membranous, connective-tissue-like sac in which the heart is contained

perimenopausal. the period surrounding menopause (lasting several months to several years depending on the individual)

peritoneum. epithelial tissue lining of abdominal cavity

phagocytosis. ability of one cell to engulf and destroy another

phagosome. an encapsulated pathogen once it has been engulfed by a neutrophil or macrophage

phenotype. the degree of maleness or femaleness of a developing human

platelets. anucleated, disk-like blood cells designed to prevent the loss of blood from damaged vessels

pneumothorax. puncture of the thoracic cavity, loss of subatmospheric intrapleural pressure with possible collapse of lungs (uninflatable until repaired)

PO_2. partial pressure of oxygen in an unspecified location

polycythemia. excess red blood cells in the circulation

postabsorptive state. the time after which most of a meal is digested and absorbed (about four or five hours after a typical meal in a healthy person)

postcapillary resistance. the resistance to blood flow in small blood vessels just downstream to the capillaries

posterior. in the upright individual, the back of the body

precapillary resistance. the resistance to blood flow in small blood vessels just upstream to the capillaries

primordial germ cells. gonadal undifferentiated cells that give rise to male sertoli and female granulosa cells

proteinuria. protein in the urine (an aphysiological finding under most conditions)

pulsatile blood pressure. changes, throughout the cardiac cycle, in arterial blood pressure

PVO_2. partial pressure of oxygen in mixed venous blood (inside the right atrium, right ventricle, or pulmonary artery)

RAAA system. renin, angiotensin, aldosterone, antidiuretic hormone regulatory system; helps govern body water, electrolytes, and blood pressure

reactive hyperemia. excess blood flow that occurs upon release of a mechanically occluded blood vessel

reactive oxygen species. chemicals such as superoxide anion, hydroxyl radical, and peroxynitrite that damage biologically important processes or structures

receptors. docking stations made of proteins that are integral to all cell membranes

recruitment. an increasing number of motor units functioning as stimulus strength increases

reduced hemoglobin. HbH, hemoglobin carrying excess acid, such as that leaving the systemic capillaries en route to the lungs

reflex arc. an anatomical structure of five components needed to convert a sensory signal into a motor action

refractory period. the period of time following initial depolarization during which a second stimulus will not cause an action potential

regenerative medicine. the predicted future cure of disease and ailments based on the use of stem cells

renal baroreceptors. pressure sensors in the renal afferent arteriole that are part of the juxtaglomerular apparatus

renal capsule. a rigid connective tissue sheath encasing each kidney

renal hilus. the point of entry or exit of renal blood vessels, nerves, and ureters in each kidney

renal pelvis. inner area of kidney that gives rise to the ureters

renal portal. renoportal; the arrangement of two sets of renal capillaries (glomerular and peritubular) in series with each other

renin. a proteolytic enzyme that is released by the renal JGA when renal arterial perfusion pressure declines

reperfusion. restoration of tissue blood flow following a period of ischemia

repolarization. return from the depolarized state to a resting (negative) membrane potential

RES system. reticuloendothelial system; enlarged extravascular monocytes and macrophages that are important parts of the immune system

respiratory zone. respiratory bronchioles, alveolar ducts, alveolar sacs, and alveoli

ryanodine receptor. intracellular calcium receptors associated with release of calcium from the SR

saggital. in the upright individual, a section that divides the body in two using the midline

sarcolemma. the cell membrane of a myocyte, particularly striated muscle cells

sarcomere. the basic functional and structural unit of striated muscle cells

sarcoplasmic reticulum. SR; an intracellular system of tubules in myocytes that store and release calcium

set points. genetically (physiologically) determined value for a variable

slow-twitch fibers. small fibers with an extensive capillary network (high blood flow) and a high content of cytosolic myoglobin

spermatogenesis. growth and development of the maturing sperm cells

spermatogonium. a developing sperm cell

spirometry. method used to test the functional capacity of the respiratory system

SRY gene. sex-determining region of the Y chromosome; helps determine maleness

Starling forces. hemodynamic variables that influence the formation of ultrafiltrate

STDs. sexually transmitted diseases

steady state. a physiologic state when a variable is constant

striated muscles. muscle fibers with characteristic alternating light and dark bands (striations); examples are cardiac and skeletal muscle

superior. in the upright individual, a landmark that is towards the head

synapse. space between a pre- and postsynaptic membrane

tachycardia. fast heart rate, one hundred beats per minute or greater

TDF. testis-determining factor; a transcription element encoded by the SRY gene

thrombocytes. another term for *platelets*

tidal volume. volume of air inhaled or exhaled during a single respiratory cycle

tissue macrophages. circulating monocytes and macrophages that have transformed and taken up residence in the interstitial spaces

trans-alveolar. across the wall of an alveolus; the wall of the alveolus separates the intra-alveolar and intrapleural spaces

translational physiology. experimental findings that are quickly transferable to a clinical setting

transmission. movement of an action potential across a synapse by chemical or electrical means

transport maximum. T_m, the point at which all protein carriers in an epithelial cell are occupied by ligand

transverse. in the upright individual, a section that divides the body into superior and inferior halves, such as at the pelvic girdle

tropomysin. a thin filament protein that regulates cross-bridge formation between myosin and actin

troponin-C. a thin filament calcium-binding protein

t-tubules. indentations (invaginations) of the sarcolemma that plunge deep into the myocyte but still represent extracellular space

ultrafiltrate. aqueous (water-like) solvent and solutes that leave glomerular capillaries and enter Bowman's space

vasodepressor center. neurons in the brain stem that cause blood pressure to decrease

vasomotor. having to do with the regulation of internal radius of a blood vessel (motor tone)

vasopressor center. neurons in the brain stem that cause blood pressure to increase

ventilation. movements of lungs and thorax that cause respiratory airflow

ventral respiratory group. neurons that are part of the medullary respiratory control centers but are located inferiorally relative to the superior dorsal neurons

vital signs. heart rate, blood pressure, respiratory rate, body temperature

walling off. the prevention of diffusion of toxins to locations beyond the site of invasion

writer's cramp. pain and tissue dystrophy that accompany excess use of flexor muscles and simultaneous disuse of extensors in the arm and hand

NOTES AND SUGGESTED READING

PREFACE

Note

1. AC Guyton, JE Hall, *Textbook of Medical Physiology,* 11th ed. (Philadelphia: Elsevier Saunders, 2006).

CHAPTER 1 THE FOUNDATION

Suggested Reading

Angelucci A, Bressloff PC. 2006. Contribution of feedforward, lateral, and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Prog Brain Res.* 154:93–120, Review.

Aston-Jones G, Cohen JD. 2005. An integrative theory of locus coeruleus norepinephrine function: adaptive gain and optimal performance. *Ann Rev Neurosci.* 28:403–50, Review.

Berne RM, Levy MN, Koeppen BM, Stanton BA. 2004. *Physiology*. 5th ed. St. Louis: Mosby.

Boron WF, Boulpaep EL. 2005. *Medical Physiology: A Cellular and Molecular Approach*, updated ed. Philadelphia: Elsevier Saunders.

Cannon WB. 1939. *The Wisdom of the Body*. New York: W. W. Norton.

Dolman NJ, Tepikin AV. 2006. Calcium gradients and the Golgi. *Cell Calcium*. 40:505–12, Review.

Guyton AC, Hall JE. 2006. *Textbook of Medical Physiology*, 11th ed. Philadelphia: Elsevier Saunders.

Nguyen DC, Sainte-Marie Y, Jaisser F. 2007. Animal models in cardiovascular diseases: new insights from conditional models. *Handbook Exp Pharmacol.* 178:377–405, Review.

Tin C, Poon CS. 2005. Internal models in sensorimotor integration: perspectives from adaptive control theory. *J Neural Eng.* 2:S147–63, Review.

CHAPTER 2 UNDERSTANDING THE MAMMALIAN NERVOUS SYSTEM

Suggested Reading

Ahs F, Furmark T, Michelgard A, Langstrom B, Appel L. 2006. Hypothalamic blood flow correlates positively with stress-induced cortisol levels in subjects with social anxiety disorder. *Psychosom Med.* 68:859–862.

Borodinsky LN, Spitzer NC. 2007. Activity-dependent neurotransmitter-receptor matching at the neuromuscular junction. *Proc Natl Acad Sci USA.* 104:335–340.

Germann WJ, Stanfield CL. 2005. *Principles of Human Physiology,* 2nd ed. San Francisco: Pearson Benjamin Cummings.

Gilbert DL. 1983. The first documented report of mountain sickness: the China or Headache Mountain story. *Respir Physiol* 52:315–326.

Hackett PH, Rennie D, Levine HD. 1976. The incidence, importance, and prophylaxis of acute mountain sickness. *Lancet* 2:1149–1155.

Harris NS, Wenzel RP, Thomas SH. 2003. High altitude headache: efficacy of acetaminophen vs. ibuprofen in a randomized, controlled trial. *J Emerg Med* 24:383–387.

Honigman B, Theis MK, Kozoil-McLain J. 1993. Acute mountain sickness in a general tourist population at moderate altitudes. *Ann Intern Med* 118:587–592.

Ingvar DH. 1976. Functional landscapes of the dominant hemisphere. *Brain Research* 107:181–197.

Johnson TS, Rock PB. 1988. Acute mountain sickness. *N Engl J Med* 319:841–845.

Marcum JA. 2006. "Soup" vs. "sparks": Alexander Forbes and the synaptic transmission controversy. *Ann Sci.* 63:139–56.

Meier C, Dermietzel R. 2006. Electrical synapses: gap junctions in the brain. *Results Prob Cell Differ.* 43:99–128, Review.

Montgomery AB, Mills J, Luce JM. 1989. Incidence of acute mountain sickness at intermediate altitude. *JAMA* 261:732–734.

Newberg AB, Wintering NA, Morgan D, Waldman MR. 2006. The measurement of regional cerebral blood flow during glossolalia: a preliminary SPECT study. *Psychiatry Res.* 148:67–71.

Puzzo D, Palmeri A, Arancio O. 2006. Involvement of the nitric oxide pathway in synaptic dysfunction following amyloid elevation in Alzheimer's disease. *Rev Neurosci.* 17:497–523, Review.

Rork TH, Hadzimichalis NM, Baliga SS, Golfetti R, Merrill GF. 2006. New perspectives on acetaminophen. *Curr. Cardiol. Rev.* 2:131–146, Review.

Rork TH, Hadzimichalis NM, Merrill GF. 2006. Acetaminophen attenuates peroxynitrite-activated matrix metalloproteinase-2-mediated troponin I cleavage in the isolated guinea pig myocardium. *J Mol Cell Cardiol.* 40:553–561.

Spiler NM, Rork TH, Merrill GF. 2005. An old drug with a new purpose: cardiovascular actions of acetaminophen (paracetamol). *Curr. Drug Targets—Cardiovasc Hemat Disorders.* 5(5):419–429, Review.

CHAPTER 3 THE ENDOCRINE SYSTEM AND PHYSIOLOGICAL COMMUNICATION

Suggested Reading

Barrett, EJ. 2005. Organization of endocrine control. In *Medical Physiology*, by Boron WF, Boulpaep EL, updated ed. Philadelphia: Elsevier Saunders.

Bellino FL. 2006. Advances in endocrinology of aging research, 2005–2006. *Exp Gerontol.* 41:1228–1233.

Beuschlein F, Reincke M. 2006. Adrenocortical tumorigenesis. *Ann NY Acad Sci.* 1088:319–334.

Chahal H, Drake W. 2007. The endocrine system and ageing. *J Pathol.* 211:173–180.

Farwell AP, Braverman LE. 1996. Thyroid and antithyroid drugs. In *The Pharmacological Basis of Therapeutics*, by Goodman LS, Gilman AG, 9th ed. New York: McGraw-Hill.

Gulcelik MA, Dinc S, Kuru B. 2006. The incidence of hyperthyroidism in patients with thyroid cancer in an area of iodine deficiency. *J Surg Oncol.* 94:35–39.

Mormede P, Andanson S, Auperin B, Beerda B, Guemene D. 2007. Exploration of the hypothalamic-pituitary-adrenal function as a tool to evaluate animal welfare. *Physiol Behav.* Jan 15; Epub ahead of print.

Pellecchia MT, Pivonello R, Colao A, Barone P. 2006. Growth hormone stimulation tests in the differential diagnosis of Parkinson's disease. *Clin Med Res.* 4:322–325.

Reddy SS. 2006. Endocrinology update 2006. *Cleve Clin J Med.* 73:1019–1024, Review.

Rodondi N, Aujesky D. 2006. Subclinical hypothyroidism and the risk of coronary heart disease: a meta-analysis. *Am J Med.* 119:541–551.

Supomsilchai V, Soder O, Svechnikov K. 2007. Stimulation of the pituitary-adrenal axis and of adrenocortical steroidogenesis ex vivo by administration of di-2-ethylhexyl phthalate to prepubertal male rats. *J Endocrinol.* 192:33–39.

CHAPTER 4 THE CARDIOVASCULAR SYSTEM AND THE BLOOD

Suggested Reading

Barclay JK, Stainsby WN. 2004. Perfused skeletal muscle: an experimental preparation for many questions. *Can J Appl Physiol.* 29:123–138, Review.

Ellison GM, Torella D, Karakikes I, Nadal-Ginard B. 2007. Myocyte death and renewal: modern concepts of cardiac cellular homeostasis. *Nat Clin Pract Cardiovasc Med.* 4 Suppl 1(S1):S52-S59.

Federici AB. 2006. Management of inherited von Willebrand disease in 2006. *Semin Thromb Hemost.* 32:616–620.

Fliedner TM. 2006. Nuclear terrorism: the role of hematology in coping with its health consequences. *Curr Opin Hematol.* 13:436–444, Review.

Fujiwara Y, Kurokawa S, Asakura Y. 2007. Correlation between heart rate variability and haemodynamic fluctuation during induction of general anaesthesia: comparison between linear and non-linear analysis. *Anaesthesia.* 62:117–121.

Hayes SN. 2006. Preventing cardiovascular disease in women. *Am Fam Physician* 74:1331–1340, Review.

Itshaki S, Lavie L, Pillar G, Tal G, Lavie P. 2005. Endothelial dysfunction in obstructive sleep apnea measured by peripheral arterial tone response in the finger to reactive hyperemia. *Sleep* 28:594–600.

Klieman L, Hyde S, Berra K. 2006. Cardiovascular disease risk reduction in older adults. *J Cardiovasc Nurs.* 21(5 Suppl 1):S27-S39, Review.

Lafontant PJ, Field LJ. 2006. The cardiomyocyte cell cycle. *Novartis Found Symp.* 274:196–207; discussion 208–213, 272–276, Review.

Loukogeorgakis SP, Panagiotidou AT, Yellon DM, Deanfield JE, MacAllister RJ. 2006. Postconditioning protects against endothelial ischemia-reperfusion injury in the human forearm. *Circulation* 113:1015–1019.

McTigue KM, Hess R, Ziouras J. 2006. Obesity in older adults: a systematic review of the evidence for diagnosis and treatment. *Obesity* (Silver Spring) 14:1485–1497, Review.

Nakamura M, Han B, Nunobiki O, Kakudo K. 2006. Adrenomedullin: a tumor progression factor via angiogenic control. *Curr Cancer Drug Targets* 6:635–43, Review.

National Technical Information Service. 2005. *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2005*. Springfield, VA: National Technical Information Service.

Pritchard KI, Abramson BL. 2006. Cardiovascular health and aromatase inhibitors. *Drugs* 66:1727–1740, Review.

Terashvili M, Pratt PF, Harder DR. 2006. Reactive oxygen species cerebral autoregulation in health and disease. *Pediatr Clin North Am.* 53:1029–1037, Review.

USDA, Human Nutrition Information Service. 1982. *Nutritive Value of Foods, 1982*. Home and Garden Bulletin no. 72. Washington, DC: U.S. Government Printing Office.

Zandecki M, Genevieve F, Gerard J, Godon A. 2007. Spurious counts and spurious results on haematology analyzers: a review. Part I: platelets. *Clin Lab Haematol.* 29:4–20.

Zandecki M, Genevieve F, Gerard J, Godon A. 2007. Spurious counts and spurious results on haematology analyzers: a review. Part II: white blood cells, red blood cells, haemoglobin, red cell indices and reticulocytes. *Clin Lab Haematol.* 29:21–41.

CHAPTER 5 HEALTH AND THE RESPIRATORY SYSTEM

Suggested Reading

Doelken P, Abreu R, Sahn SA, Mayo PH. 2006. Effect of thoracentesis on respiratory mechanics and gas exchange in the patient receiving mechanical ventilation. *Chest* 130:1354–1361.

Findlay G, Wise M. 2006. Acute hypercapnia and gas exchange in ARDS. *Chest* 130:1950–1961.

Frangiosa A, DeSanto LS, DeSanto NG. 2006. Acid-base balance in heart failure. *J Nephrol.* Suppl 9:S115–120, Review.

Krakauer J. 1997. *Into Thin Air.* New York: Villard.

Leon-Velarde F, Richalet JP. 2006. Respiratory control in residents at high altitude: physiology and pathophysiology. *High Alt Med Biol.* 7:125–137, Review.

McCarren B, Alison J, Herbert R. 2006. Manual vibration increases expiratory flow rate via increased intrapleural pressure in healthy adults: an experimental study. *Aust J Physiother.* 52:267–271.

Milic-Emili J. 2004. Does mechanical injury of the peripheral airways play a role in the genesis of COPD in smokers? *Chronic Obstru Pulonary Dis.* 1:85–92.

Moe OW, Fusten D. 2003. Clinical acid-base pathophysiology: disorders of plasma anion gap. *Best Pract Res Clin Endocrinol Metab.* 17:559–574, Review.

Pena F, Garcia O. 2006. Breathing generation and potential pharmacotherapeutic approaches to central respiratory disorders. *Curr Med Chem.* 13:2681–2693, Review.

West JB. 1985. *Everest, the Testing Place.* New York: McGraw-Hill.

Whitelaw W. 2006. Mechanisms of sleep apnea at altitude. *Adv Exp Med Biol.* 288:57–63, Review.

Wutzl A, Ploder O, Kermer C, Klug C. 2007. Mortality and causes of death after multimodality treatment for advanced oral and otopharyngeal cancer. *J Oral Maxillofac Surg.* 65:255–260.

CHAPTER 6 KIDNEYS AND RENAL PHYSIOLOGY

Note

1. Å; an angstrom is 0.1 of a nanometer, and it takes one billion nanometers to equal one meter. The radii of biologically important molecules are expressed in angstroms.

Suggested Reading

Agre P. 2006. The aquaporin water channels. *Proc Am Thorac Soc.* 3:5–13, Review.

Bouby N, Fernandes S. 2003. Mild dehydration, vasopressin, and the kidney: animal and human studies. *Eur J Clin Nutr.* Suppl 2:S39–S46, Review.

Broumand B. 2007. Diabetes: changing the fate of diabetics in the dialysis unit. *Blood Purif.* 25:39–47.

Cooper RG. 2006. Effect of tobacco smoking on renal function. *Indian J Med Res.* 124:261–268, Review.

Ichihara A, Kobori H, Nishiyama A, Navar LG. 2004. Renal renin-angiotensin system. *Contrib Nephrol.* 143:117–130, Review.

Jones AW. 2006. Urine as a biological specimen for forensic analysis of alcohol. *Toxicol Rev.* 25:15–35, Review.

Komlosi P, Fintha A, Bell PD. 2004. Current mechanisms of macula densa cell signaling. *Acta Physiol Scand.* 181:463–469, Review.

Persson PB. 2003. Renin: origin, secretion, and synthesis. *J Physiol.* 552:667–671, Review.

Peti-Peterdi J. 2005. Multiphoton imaging of renal tissues in vitro. *Am J Physiol Renal Physiol.* 288:F1079–F1083, Review.

Rahmoune H, Thompson PW, Ward JM, Smith CD, Hong G, Brown J. 2005. Glucose transporters in human renal proximal tubular cells isolated from the urine of patients with non-insulin-dependent diabetes. *Diabetes* 54:3427–3434.

Rehrer NJ. 2001. Fluid and electrolyte balance in ultra-endurance sport. *Sports Med.* 31:701–715, Review.

Schaan BD, Machado UF, Rogers S, Kelly DJ. 2006. Glucose transporters in animal models of diabetes and hypertension. *Am J Physiol Renal Physiol.* 29:F702–F703.

Schwartz GJ. 2002. Physiology and molecular biology of renal carbonic anhydrase. *J Nephrol. Suppl* 5:S61–S74, Review.

Schweda F, Kurtz A. 2004. Cellular mechanism of renin release. *Acta Physiol Scand.* 181:383–390, Review.

Senturk UK, Kuru O, Kocer G, Gunduz F. 2007. Biphasic pattern of exercise-induced proteinuria in sedentary and trained men. *Nephron Physiol.* 105:22–32.

CHAPTER 7 THE GASTROINTESTINAL SYSTEM

Suggested Reading

Azpiroz F. 2005. Intestinal perception: mechanisms and assessment. *Br J Nutr.* 93 Suppl 1:S7–S12, Review.

Bharucha AE. 2006. Update of tests of colon and rectal structure and function. *J Clin Gastroenterol.* 40:96–103, Review.

Buarner F. 2006. Enteric flora in health and disease. *Digestion* 73 Suppl 1:5–12, Review.

Gershon MD. 2005. Nerves, reflexes, and the enteric nervous system: pathogenesis of the irritable bowel syndrome. *J Clin Gastroenterol.* 39 (4 Suppl 3):S184–S193, Review.

Hou W, Schubert ML. 2006. Gastric secretion. *Curr Opin Gastroenterol.* 22:593–598, Review.

Karaki SI, Kuwahara A. 2004. Regulation of intestinal secretion involved in the interaction between neurotransmitters and prostaglandin E2. *Neurogastroenterol Motil.* 16 Suppl 1:96–99, Review.

Kunze WA, Furness JB. 1999. The enteric nervous system and regulation of intestinal motility. *Annu Rev Physiol.* 61:117–142, Review.

Layer P, Keller J. 1999. Pancreatic enzymes: secretion and luminal nutrient digestion in health and disease. *J Clin Gastroenterol.* 28:3–10, Review.

Lipman MJ, Reidenberg JS, Laitman JT. 2006. Anatomy of reflux: a growing health problem affecting structures of the head and neck. *Anat Rec B New Anat.* 289:261–270, Review.

Llena-Puy C. 2006. The role of saliva in maintaining oral health and as an aid to diagnosis. *Med Oral Patol Oral Cir Bucal.* 11:E449–E455, Review.

Sachs G, Weeks DL, Wen Y, Marcus EA, Scott DR. 2005. Acid acclimation by *Helicobacter pylori*. *Physiology* 20:429–438.

Schultze K. 2006. Imaging and modelling of digestion in the stomach and the duodenum. *Neurogastroenterol Motil.* 18:172–183, Review.

USDA. 2004. *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2005.* August.

Whole grains: the inside story. 2006. *Nutrition Action Health Letter* 33(4). May.

CHAPTER 8 THE REPRODUCTIVE SYSTEM

Suggested Reading

Budzynski E, Lee Y, Sakamoto K, Nishina PM. 2006. From vivarium to bedside: lessons learned from animal models. *Ophthalmic Genet.* 27:123–137, Review.

Cohen JE. 2003. Human population: the next half century. *Science* 302:1172–1175.

Dong C, Wilhelm D, Koopman P. 2004. Sex genes and cancer. *Cytogenet Genome Res.* 105:442–447, Review.

Gregersen PK, Behrens TW. 2006. Genetics of autoimmune diseases: disorders of immune homeostasis. *Nat Rev Genet.* 7:917–928, Review.

Ho GY, Bierman R, Beardsley L, Chang CK, Burk RD. 1998. Natural history of cervicovaginal papillomavirus infection in young women. *N Engl J Med.* 338:423–428.

Jorgensen N, Asklund C, Carlsen E, Skakkebaek NE. 2006. Coordinated European investigations of semen quality: results from studies of Scandinavian young men is a matter of concern. *Int J Androl.* 29:54–61; discussion 105–108, Review.

Kurjak A, Carrera JM. 2005. Declining fertility in the developed world. *J Perinat Med.* 33:95–99.

Merchant-Larios H, Moreno-Mendoza N. 2001. Onset of sex differentiation: dialog between genes and cells. *Arch Med Res.* 32:553–558, Review.

Ng K, Donat R, Chan L, et al. 2004. Sperm output of older men. *Human Repro.* 19:1811–1815.

Oatley JM, Avarbock MR, Telaranta AI, Fearon DT, Brinster RL. 2006. Identifying genes important for spermatogonial stem cell self-renewal and survival. *Proc Natl Acad Sci USA.* 103:9524–9529.

Pal L, Santoro N. 2003. Age-related decline in fertility. *Endocrin Metab Clin No Am.* 32:669–588, Review.

Rasmussen F, Gunneell D, Ekbom A. 2003. Birth weight, adult height, and testicular cancer: cohort study of 337,249 Swedish young men. *Cancer Causes Control* 14:595–598.

Roy A, Matzuk MM. 2006. Deconstructing mammalian reproduction: using knockouts to define fertility pathways. *Reproduction* 131:207–219, Review.

Thajeb P, Dai D, Chiang MF, Shyu WC. 2006. Genotype-phenotype correlation of maternally inherited disorders due to mutations in mitochondrial DNA. *Taiwan J Obstet Gynecol.* 45:201–207, Review.

Wilhelm D, Palmer S, Koopman P. 2007. Sex determination and gonadal development in mammals. *Physiol Rev.* 87(1):1–28.

CHAPTER 9 THE IMMUNE SYSTEM

Note

1. *Dorland's Illustrated Medical Dictionary*, 29th centennial edition (Philadelphia: WB Saunders, Elsevier Saunders, 2000), 878.

Suggested Reading

Broughton G, Janis JE, Attinger CE. 2006. The basic science of wound healing. *Plast Reconstr Surg.* 117 (7 Suppl):12S-34S, Review.

Cabrera CM, Cobo F, Nieto A, Concha A. 2006. Strategies for preventing immunologic rejection of transplanted human embryonic stem cells. *Cytotherapy* 8:517–518, Review.

Charest PG, Firtel RA. 2007. Big roles for small GTPases in the control of directed cell movement. *Biochem J.* 401:377–390, Review.

Christopher MJ, Link DC. 2007. Regulation of neutrophil homeostasis. *Curr Opin Hematol.* 14:3–8, Review.

Couzin J. 2006. Teams identify cardiac "stem cell." *Science* 314 (Nov. 24):1225.

Engelmann MG, Franz WM. 2006. Stem cell therapy after myocardial infarction: ready for clinical application? *Curr Opin Mol Ther.* 8:396–414, Review.

Giles GI. 2006. The redox regulation of thiol dependent signaling pathways in cancer. *Curr Pharm Des.* 12:4427–4443, Review.

Hamburger AE, Bjorkman PJ, Herr AB. 2006. Structural insights into antibody-mediated mucosal immunity. *Curr Top Microbiol Immunol.* 308:173–204, Review.

Hatakeyama M, Brzozowski T. 2006. Pathogenesis of *Helicobacter pylori* infection. *Helicobacter* Suppl 1:14–20, Review.

Lefkowitz RJ, Rajagopal K, Whalen EJ. 2006. New roles for beta-arrestins in cell signaling: not just for seven-transmembrane receptors. *Mol Cell.* 24:643–652, Review.

Marcenaro E, Dondero A, Moretta A. 2006. Multi-directional cross-regulation of NK cell function during innate immune responses.*Transpl Immunol.* 17:16–19, Review.

Michelsen KS, Arditi M. 2007. Toll-like receptors and innate immunity in gut homeostasis and pathology. *Curr Opin Hematol.* 14:48–54, Review.

Niess JH, Reinecker HC. 2006. Dendritic cells: the commanders-in-chief of mucosal immune defenses. *Curr Opin Gastroenterol.* 22:354–360, Review.

Reed JA, Patarca R. 2006. Regenerative dental medicine: stem cells and tissue engineering in dentistry. *J Environ Pathol Toxicol Oncol.* 25:537–69, Review.

Turina M, Fry DE, Polk HC. 2005. Acute hyperglycemia and the innate immune system: clinical, cellular, and molecular aspects. *Crit Care Med.* 33:1624–1633, Review.

Turksen K, Rao M. 2005. Issues in human embryonic stem cell biology. *Stem Cell Rev.* 1:79–81, Review.

CHAPTER 10 MUSCLE FUNCTION

Suggested Reading

Aufderheide TP, Lurie KG. 2006. Vital organ blood flow with the impedance threshold device. *Crit Care Med.* 34:S466–S473.

Behnia R, Moteni A, Igie R. 2003. Angiotensin-converting enzyme inhibitors: mechanisms of action and implications in anesthesia practice. *Curr Pharm Des.* 9:763–776, Review.

Brignole M. 2005. Neurally mediated syncope. *Ital Heart J.* 6:249–255, Review.

Doshi M, Murray PT. 2003. Approach to intradialytic hypotension in intensive care unit patients with acute renal failure. *Artif Organs* 27:772–780, Review.

Dubick MA, Bruttig SP, Wade CE. 2006. Issues of concern regarding the use of hypertonic/hyperoncotic fluid resuscitation of hemorrhagic hypotension. *Shock* 25:321–328, Review.

Fujita T. 2006. Spotlight on renin: the renin system, salt-sensitivity, and metabolic syndrome. *J Renin Angiotensin Aldosterone Syst.* 7:181–183, Review.

Guitierrez G, Reines HD, Wulf-Gutierrez ME. 2004. Clinical review: hemorrhagic shock. *Crit Care* 8:373–381, Review.

Karagiannis A, Mikhailidis DP, Athyros VG. 2007. The role of renin-angiotensin system inhibition in the treatment of hypertension in metabolic syndrome: are all the angiotensin receptor blockers equal? *Expert Opin Ther Targets* 11:191–205.

Loehe F, Bruns CJ, Nitsch SM, Angele MK. 2007. The role of L-arginine following trauma and blood loss. *Curr Opin Clin Nutr Metab Care* 10:80–87.

Sarafidis PA, Khosla N, Bakris GL. 2007. Antihypertensive therapy in the presence of proteinuria. *Am J Kidney Dis.* 49:12–26.

Taylor AA. 1994. Autonomic control of cardiovascular function: clinical evaluation in health and disease. *J Clin Pharmacol.* 34:363–374, Review.

Tox U, Stefen HM. 2006. Impact of inhibitors of the renin-angiotensin-aldosterone system on liver fibrosis and portal hypertension. *Curr Med Chem.* 13:3649–3661, Review.

Treschan TA, Peters J. 2006. The vasopressin system: physiology and clinical strategies. *Anesthesiology* 105:599–612, Review.

Worthley LI. 2000. Shock: a review of pathophysiology and management. Part I. *Crit Care Resusc.* 2:55–65.

CHAPTER 11 INTEGRATED PHYSIOLOGICAL RESPONSES

Note

1. The example of physiological gain in a cardiovascular feedback control system is taken from AC Guyton and JE Hall, *Textbook of Medical Physiology,* 11th ed. (Philadelphia: Elsevier Saunders, 2006), 280.

Suggested Reading

Arad M, Seidman EC, Seidman JG. 2007. AMP-activated protein kinase in the heart: role during health and disease. *Circ Res.* 100:474–488, Review.

Bers DM. 2006. Altered cardiac myocyte Ca regulation in heart failure. *Physiol.* 21:380–387, Review.

Cossu G, DeAngelis L, Borello U, Buffa V. 2000. Determination, diversification, and multipotency of mammalian myogenic cells. *Int J Dev Biol.* 44:699–706.

Dantzig JA, Liu TY, Goldman YE. 2006. Functional studies of individual myosin molecules. *Ann NY Acad Sci.* 1080:1–18, Review.

Dulhunty AF. 2006. Excitation-contraction coupling from the 1950s into the new millennium. *Clin Exper Pharmacol Physiol.* 33:763–772, Review.

Hughes SM, Salinas PC. 1999. Control of muscle fiber and motoneuron diversification. *Curr Opin Neurobiol.* 9:54–64.

Jurkat-Rott K, Lehmann-Horn F. 2007. Genotype-phenotype correlation and therapeutic rationale in hyperkalemic periodic paralysis. *Neurotherapeutics* 4: 216–224.

Merrill GF, Haddy FJ, Dabney JM. 1978. Adenosine, theophylline, and perfusate pH in the isolated perfused guinea pig heart. *Circ Res.* 42:225–229.

Nichols TR, Cope TC. 2004. Cross-bridge mechanisms underlying the history-dependent properties of muscle spindles and stretch reflexes. *Can J Physiol Pharmacol.* 82:569–576, Review.

Phillippou A, Maridaki M, Halapas A, Koutsilieris M. 2007. The role of insulin-like growth factor I (IGF-I) in skeletal muscle physiology. *In Vivo* 21:45–54, Review.

Poggesi C, Tesi C, Stehle R. 2005. Sarcomeric determinants of striated muscle relaxation kinetics. *Pflugers Arch.* 449:505–517, Review.

Robinson R, Carpenter D, Shaw MA, Halsall J, Hopkins P. 2006. Mutations in RYR1 in malignant hyperthermia and central core disease. *Hum Mutat.* 10:977–989, Review.

Rork TH, Hadzimichalis NM, Merrill GF. 2006. Acetaminophen attenuates peroxynitrite-activated matrix metalloproteinase-2-mediated troponin I cleavage in the isolated guinea pig myocardium. *J Mol Cell Cardiol.* 40:553–561.

Serysheva II, Chiu W, Ludtke SJ. 2007. Single-partial electron cryomicroscopy of the ion channels in the excitation-contraction junction. *Methods Cell Biol.* 79:407–435, Review.

Smith NP, Barclay CJ, Loiselle DS. 2005. The efficiency of muscle contraction. *Prog Biophys Mol Biol.* 88:1–58, Review.

Tyska MJ, Mooseker MS. 2003. Myosin-V motility: these levers were made for walking. *Trends Cell Biol.* 13:447–451, Review.

CHAPTER 12 FOR THE RECORD

Note

1. The numbers to the right of the ± symbols represent standard errors of the mean (s.e.m.) and denote the statistical variability in these mean or average values.

Suggested Reading

Arnold E, Heddle N, Lane S. 2005. Handheld computers and paper diaries for documenting the use of factor concentrates used in haemophilia home therapy: a qualitative study. *Haemophilia* 11:216–226.

Geison GL. 1987. *Physiology in the American Context, 1850–1940*. Bethesda, MD: American Physiological Society.

Merrill GF. 1986/87. The case for the use of animals in medicine. In *Advances in Animal Welfare Science*, ed. MW Fox and LD Mickley, 227–244. Washington, DC: United States Humane Society.

Merrill GF, Friedrichs GS, Wei HM. 1990. Circulating cholesterol and other lipids in college students performing self-selected exercise. *J Am Coll Nutr.* 9:256–261.

Rato Q, Cardosa AA. 2006. Retrospective review of the medical records of cardiology hospital outpatient clinics with the aim of determining the therapeutic approach to patients suffering from high blood pressure and/or dyslipidemia. *Rev Port Cardiol.* 25:889–905.

Staroselsky M, Volk LA, Tsurikova R. 2006. Improving electronic health record (EHR) accuracy and increasing compliance with health maintenance clinical guidelines through patient access and input. *Int J Med Inform.* 75:693–700.

USDA. 2004. *Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2005.* Washington, DC: USDA.

Whyte M. 2005. Computerized versus handwritten records. *Paediatr Nurs.* 17:15–18.

INDEX

ABOUT THE AUTHOR

GARY MERRILL earned BS, MS, and PhD degrees in physiology between 1967 and 1975. He spent one year as a postdoctoral fellow in physiology at the LSU Medical Center in New Orleans between 1975 and 1976. His entire academic career has been spent teaching and doing research in physiology at Rutgers University in New Brunswick/Piscataway, New Jersey. His teaching has been at the undergraduate and graduate levels. His research has focused on the mammalian cardiovascular system. Dr. Merrill's hobbies include hiking and camping with his family and friends, as well as yard work, car care, and other physical activity.