Douglas College Human Anatomy & Physiology II (2nd ed.)

Douglas College Human Anatomy & Physiology II (2nd ed.)

DOUGLAS COLLEGE BIOLOGY DEPARTMENT

OPENSTAX
HOUSTON, TX

How to Reuse & Attribute This Content

Contents

Preface

Welcome to the Douglas College Anatomy & Physiology open textbook!

This textbook is a project under development by our Biology faculty to ultimately provide students with all the factual information they need to succeed in the BIOL 1203 and BIOL 1209 courses at Douglas College in BC, Canada. Readers should be aware that the information herein is subject to change at any time as corrections, additions, or other important modifications are made. Current students at Douglas College should be aware that only the most recent version of this textbook will be considered by their instructors to be complete and correct. The most recent version of this second edition will remain accessible online at https://pressbooks.bccampus.ca/dcbiol120312092nded/, and the most recent version of the second edition of the companion textbook (developed for Douglas College's BIOL 1103 and BIOL 1109 courses) will also remain accessible online at https://pressbooks.bccampus.ca/dcbiol110311092nded/.

This textbook was developed initially as an adaptation of the OpenStax Anatomy & Physiology textbook, freely and perpetually available online at http://cnx.org/content/col11496/latest/. The original adaptations of that OpenStax textbook for Douglas College are accessible online at https://pressbooks.bccampus.ca/dcbiol111031109/ and https://pressbooks.bccampus.ca/dcbiol12031209/ In the first edition of the Douglas College adaptations the chapter and section numbers were left as they were in the version of the OpenStax A&P textbook, from which they were largely drawn. However, this second edition has been more extensively edited and rearranged to correspond with the curriculum used at Douglas College, so chapter and section numbers are no longer aligned specifically with the OpenStax A&P textbook.

About this Resource

Customization

Errata

Since this textbook is primarily web based, updates can be made 'live' when deemed pedagogically necessary. If you have a correction to suggest, please submit it by email for review to Dr. Jennifer Barker, whose current contact information can be obtained from the Biology faculty page of the Douglas College website: https://www.douglascollege.ca/programs-courses/faculties/science-technology/biology/faculty

About Anatomy and Physiology

Section 1: Hormonal Regulation

Building on what students learned in the companion textbook (Douglas College Human Anatomy & Physiology I), Unit 1 asks students students to explore the structure and functions of the endocrine system, to prepare them to understand how it is used to regulate other body systems that are discussed in subsequent units.

Unit 1 The Endocrine System

Section 2: Maintenance of the Body

In Units 2-9, students examine how the various compounds are transported into, around, and out of the body. Unit 3 also includes an introduction to how our body is defended against invading pathogens.

Unit 2 The Cardiovascular System
Unit 3 The Lymphatic System, Resistance & Immunity
Unit 4 The Respiratory System
Unit 5 The Digestive System and the Digestion and Absorption of Macromolecules
Unit 6 Nutrition
Unit 7 Cellular Respiration and Energy Metabolism
Unit 8 The Urinary System
Unit 9 Fluids and Electrolytes

Section 3: Reproduction

Units 10-11 introduce students to the reproductive system and to basic concepts in human genetics.

Unit 10 The Reproductive Systems
Unit 12 Human Genetics

Additional sections

A general introduction to the basics of human anatomy as well as the remaining systems of the human body are covered in the companion textbook to this one, designed for Douglas College's BIOL 1103 and BIOL 1109 courses.

About the Authors

Editors

Jennifer Barker, Douglas College
Casper De Villiers, Douglas College
Todd Harper, Douglas College

Contributing Authors

Jennifer Barker, Douglas College
Jessie Clasen, Douglas College
Reyniel Cruz-Aguado, Douglas College
Casper De Villiers, Douglas College
Luis R. Gonzalez-Torres, Douglas College
Leon J. Guppy, Douglas College
Todd Harper, Douglas College
Sarwat Jamil, Douglas College
Shamsa Jessa, Douglas College
Weissy Lee, Douglas College
Elinor Matheson, Douglas College
Rosemary Oh-McGinnis, Douglas College
Maxence Salomon, Douglas College
Lynette Sigola-Baretto, Douglas College
Mike Silvergieter, Douglas College
Liza Sutton, Douglas College
Ryan Viveiros, Douglas College
Shelley Weisser, Douglas College
Cheryl Tautorus, Douglas College

Special Thanks

The authors of this textbook wish to thank OpenStax for the initial creation of a college-level open Anatomy & Physiology textbook, without which it is unlikely this edition would have been produced.

We also wish to thank BCcampus for providing financial support for the development of vector-based images to accompany this textbook, for providing the instance of the Pressbooks platform on which this textbook is hosted, and for providing technical support to the authors.

Finally, we wish to thank the remaining faculty members of the Biology Department for their valuable input into the content and organization of this textbook, Sara McKinnon for creating the section on lever systems and its associated diagrams, and Zoir Amirdad for creating many of the scalable vector-based versions of the images found in this textbook and also available as separate auxiliary resources.

HORMONAL REGULATION

Unit 1: The Endocrine System

Unit Outline

Part 1: General properties of the endocrine system

- Endocrine vs Exocrine glands
- General functions of hormones
- Hormone secretion: regulation and stimuli
- Types of hormones
- Types of receptors
- Endocrine vs Nervous systems

Part 2: Major endocrine organs and their secretions

- Hypothalamus and pituitary glands
- Thyroid gland
- Parathyroid gland
- Adrenal cortex and adrenal medulla
- Pancreas
- Ovaries and testes
- Stomach and duodenum
- Thymus and pineal gland
- The special nature of prostaglandins

Learning Objectives

At the end of this unit, you should be able to:

I. Define “gland”.

II. Distinguish between endocrine glands and exocrine glands.

III. Describe the purpose and regulation of hormone secretion.

IV. Describe stimuli for hormone secretion.

V. Describe the main categories of hormones, and how this relates to their receptors and signaling pathways.

VI. Compare and contrast the nervous and endocrine systems.

VII. Identify on a diagram of the human body the locations of important endocrine glands.

VIII. Describe the hypothalamus and pituitary glands and their interrelationship.

IX. Describe the function and secretion of hormones released by the pituitary gland.

X. Describe the function and secretion of hormones released by the thyroid gland.

XI. Describe the function and secretion of hormones released by the parathyroid glands.

XII. Describe the function and secretion of hormones released by the adrenal gland.

XIII. Describe the function and secretion of hormones released by the pancreas.

XIV. Name the hormones produced by the following glands and describe their actions: ovaries, testes, stomach, duodenum, thymus and pineal gland.

XV. Describe prostaglandins, referring to their composition, where they are produced, where they generally have an effect, and four effects.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Define “gland”.

II. Distinguish between endocrine glands and exocrine glands.

1. Describe the different means by which exocrine and endocrine glands release their secretions.
2. Identify the general difference between the types of secretions that these two types of glands secrete and name two examples of secretions from each type of gland.
3. Name two organs in the body that have both exocrine and endocrine functions. Identify the exocrine and endocrine secretions of one of these organs.

III. Describe the purpose and regulation of hormone secretion.

1. Specify the fundamental function of the endocrine system (include the definition of homeostasis).
2. Describe six overall functions of hormones.
3. Describe an example of a hormonally based positive feedback mechanism, justifying why this example can be considered as one of positive feedback.
4. Name and describe the more common method of hormone regulation. Identify the end result of this method.

IV. Describe stimuli for hormone secretion.

1. Explain what is meant by humoral stimulation of hormonal release.
2. Describe three examples of hormones that have humoral stimuli.
    - Name the hormone, the organ that releases the hormone and the compound which is

controlled by the hormone.
    - Identify whether these are positive or negative feedback mechanisms.

3. Describe one example of a hormone that is controlled by levels of other hormone(s).

    - Name the hormone and the organ that releases it.
    - Name the organs that release hormones that control the release of this hormone
    - Describe how the levels of the first hormone and those of the controlling hormones are related to each other
    - Identify whether this is a positive or negative feedback mechanism.

V. Describe the main categories of hormones, and how this relates to their receptors and signaling pathways.

1. Explain the basis upon which hormones are divided into two major groups.
2. Name and describe the three types of hormones.

    - Identify from which compounds each type is derived.
    - Name two examples of each type and actions of each example.
    - Identify which type of hormone has the longest half-life and explain the reason for this difference.

3. Explain why, although they circulate throughout the body, hormones are able to target specific cells.
4. Name five responses that may occur when a hormone successfully interacts with a cell.
5. Distinguish between intracellular and extracellular receptors

    - For each type of receptor, identify the location (inside or on the cell membrane), and the type of hormone with which they interact (i.e., whether lipid or amino acid based).
    - Explain why the type of receptor used by a particular hormone is related to the hydrophilic nature of the hormone.
    - Give two examples of hormones that interact with each of the two types of receptor.
    - Distinguish between the general mechanism that occurs after hormones interact with each of the two types of receptors.

VI. Compare and contrast the nervous and endocrine systems.

1. Identify the type of intercellular communication each system uses.
2. Describe the anatomical relationship between the sending and receiving cells in each system.
3. Identify which system has the more rapid and specific method of message transmission and explain the reason for this.
4. Differentiate between the general purposes of the two systems (i.e., which type of body function is mainly governed by each type).

VII. Identify on a diagram of the human body the locations of each of the following glands (or parts of glands):

1. Pineal gland
2. Thymus

3. Hypothalamus
4. Adrenal glands
5. Adrenal cortex
6. Adrenal medulla
7. Anterior pituitary
8. Posterior pituitary
9. Pancreatic islets
10. Thyroid
11. Ovaries
12. Testes
13. Parathyroid glands

VIII. Describe the hypothalamus and pituitary glands and their interrelationship.

1. Justify the basis for labelling the hypothalamus-pituitary complex as the “command center” of the endocrine system.
2. Describe (or draw) the location of the hypothalamus and the anterior pituitary gland, and the anatomical connection between the two glands, including the nature of the vascular connection.
3. Describe how a signal is sent from the hypothalamus to the anterior pituitary, to either inhibit or stimulate the release of an anterior pituitary hormone.
4. Name and describe the functions of six hypothalamic hormones that control the secretions of the anterior pituitary.
5. Describe (or draw) the location of the hypothalamus and the posterior pituitary gland, and the anatomical connection between the two glands, including the nature of the neural connection.
6. Name the two hypothalamic hormones that are stored in and secreted from the posterior pituitary.

IX. Describe the function and secretion of hormones released by the pituitary gland.

1. Explain what is meant by four of the anterior pituitary hormones being referred to as “tropic” hormones.
2. Name and describe the functions of the two anterior pituitary hormones that do not control the secretion of other endocrine glands.
3. Describe how the levels of these hormones are controlled.
4. Name and describe the conditions caused by hypo- and hypersecretion of growth hormone in childhood and adulthood.
5. Name and describe the actions of the four tropic hormones released by the anterior pituitary.
6. Describe where the hormones that the posterior pituitary secretes are actually produced, and how they are transported to the posterior pituitary.
7. Name and describe the actions of the two hormones released by the posterior pituitary.
8. Describe how the levels of these hormones are controlled.
9. Name and describe the condition caused by hyposecretion of antidiuretic hormone.
10. Differentiate between the conditions of diabetes insipidus and diabetes mellitus.

X. Describe the function and secretion of hormones released by the thyroid gland.

1. Describe the location of the thyroid gland.

2. Name the two hormones released by the thyroid gland.
3. Describe the stimulus for, and control of, the release of thyroid hormone.
4. Name the two compounds that are grouped under the term “thyroid hormone”.
5. Name and define the bodily process that is increased by the release of thyroid hormone.
6. State four other processes for which thyroid hormone is required.
7. Name and describe the conditions caused by over and undersecretion of thyroid hormone in childhood and adulthood.
8. Describe a goiter and explain how this can develop from over- and understimulation of the thyroid gland.
9. Identify the cause for “simple goiter”.
10. Describe the stimulus for, control and action of, calcitonin. Identify the type of feedback mechanism involved.

XI. Describe the function and secretion of hormones released by the parathyroid glands.

1. Describe the location of the parathyroid glands.
2. Name and describe the actions of the hormone released by the parathyroid gland.
3. Describe the stimulus for and control of the release of this hormone. Identify the type of feedback mechanism involved.
4. Name and describe the conditions caused by over and undersecretion of parathyroid hormone.

XII. Describe the function and secretion of hormones released by the adrenal gland.

1. Describe the location and the two general divisions of the adrenal gland.
2. Name the three general classes of hormones produced by the adrenal cortex.
3. Identify the major mineralocorticoid and describe the stimuli for its release, and the effects of its action.
4. Identify the major glucocorticoid and describe the stimuli for its release, and the effects of its action.
5. Name and describe the actions and stimulus of the third group of hormones released by the adrenal cortex.
6. Name and describe the actions of the two hormones released by the adrenal medulla.
7. Name two physical and two psychological stressors.
8. Name and describe the three stages of the general adaptation syndrome. For each stage, identify the major hormone involved, its effects and either the purpose of the stage or its end result.
9. Describe Addison’s disease: a cause, the effect on cortisol secretion and some effects.
10. Describe Cushing’s syndrome: a cause, the effect on cortisol secretion and the resulting effect on fat distribution and blood glucose and sodium levels.

XIII. Describe the function and secretion of hormones released by the pancreas.

1. Describe the location of the pancreas.
2. Explain the exocrine and endocrine nature of the pancreas, including the name of the clusters of cells, and individual cell types that produce insulin and glucagon.
3. Define gluconeogenesis, glycogenolysis and glycogenesis.
4. Describe the stimulus for the release of, and three actions of, glucagon. Identify the type of

feedback mechanism involved.

5. Describe the stimulus for the release of, and five actions of, insulin. Identify the type of feedback mechanism involved.
6. Name and describe the condition caused by undersecretion or ineffectiveness of insulin.
7. Name another endocrine disorder with a very similar name that is caused by undersecretion of another hormone. Name that hormone.
8. Name and describe the actions of the counterregulatory hormones (include glucagon).

XIV. Name the hormones produced by the following glands and describe their actions: ovaries, testes, stomach, duodenum, thymus and pineal gland.

XV. Describe prostaglandins, referring to their composition, where they are produced, where they generally have an effect, and four effects.

Part 1. General Properties of the Endocrine System

Introduction: The two general categories of glands in the body

The term 'gland' refers to any organ that produces a secretion. These secretions are produced by specialized cells in the glands from various components in the blood. There are two general categories of glands in the body: exocrine glands and endocrine glands.

Exocrine glands are very diverse and include the salivary glands, mammary glands, sweat glands, pancreas, stomach, prostate, and several others. Their secretions are also varied – saliva, milk, sweat, digestive enzymes and fluids to accompany gametes – just from the glands mentioned above. These glands are called exocrine glands because they have tubes or ducts to carry their secretions from the gland to another part of the body. These ducts may be simple tubes or complex, tree-like groups of ducts. Because of these tubes, the exocrine glands are also known as the ducted glands.

On the other hand, endocrine glands do not have ducts. Their secretions, called hormones, are carried to various body tissues by the blood and lymph, where they bind to receptors on target cells, inducing a characteristic response. Endocrine glands are sometimes called the ductless glands, and they all produce substances similar in nature, in that they are all hormones.

Some organs in the body contain both endocrine tissue and exocrine tissue. These organs include the pancreas, stomach and small intestine, all of which produce both hormones and digestive enzymes. The exocrine function of the pancreas (i.e., secretion of digestive enzymes into the duodenum) will be studied during the section on digestion. The endocrine function of the pancreas (release of the hormones insulin and glucagon, both of which are important in the control of blood sugar levels) will be studied later in this chapter.

General functions of endocrine hormones

You may never have thought of it this way, but when you send a text message to two friends to meet you at local cafe at six, you're sending digital signals that (you hope) will affect their behavior—even though they are some distance away. Similarly, endocrine glands send chemical signals (hormones) to other cells in the body that influence their behavior. This long-distance intercellular communication, coordination, and control is critical for homeostasis, and it is the fundamental function of the endocrine system.

Although each has its own specific effects, hormones generally have the following functions:

- Some hormones stimulate exocrine glands to produce their secretions
- Some stimulate other endocrine glands to action
- Some affect the growth, development and personality of an individual

- Some regulate body chemistry such as the metabolism of cells
- Some regulate the contraction of muscle tissues and nervous stimulation
- Some control reproductive processes

Regulation of hormone secretion

Homeostasis is the condition in which the body's internal environment remains relatively constant within limits. One of the main functions of the endocrine system is to aid in the maintenance of homeostasis. To prevent abnormal hormone levels and a potential disease state, hormone levels must be tightly controlled. The body maintains this control by balancing hormone production and degradation. Feedback mechanisms govern the initiation and maintenance of most hormone secretion in response to various stimuli.

The concept of homeostasis and the mechanisms of feedback mechanisms were presented in Unit 8 of the Biology 1103/1109 textbook (for review, refer to: https://pressbooks.bccampus.ca/dcbiol110311092nded/chapter/unit-8-homeostasis/). Recall that there are two types of feedback mechanisms: positive and negative. Positive feedback mechanisms intensify a change in the body's physiological condition rather than reversing it and result in a definite end event. An example of a hormonally based positive feedback mechanism involves the release of oxytocin during childbirth. The initial release of oxytocin begins to signal the uterine muscles to contract, which pushes the fetus toward the cervix, causing it to stretch. This, in turn, signals the pituitary gland to release more oxytocin, causing labor contractions to intensify. This will bring about the final event of childbirth, after which the release of oxytocin decreases.

However, the more common method of hormone regulation is the negative feedback mechanism, which generally is involved in the continual maintenance of a characteristic within limits. Hormonally based negative feedback mechanisms are characterized by the inhibition of further secretion of a hormone in response to adequate levels of that hormone (as determined by the amount of the hormone in the blood, or by the extent of the effect that the hormone has had). This allows blood levels of the hormone to be regulated within a narrow range.

Stimuli for hormonal secretion

The stimulus for the levels of a particular hormone can be humoral, i.e., blood levels of non-hormone chemicals such as nutrients or ions. Changes in such levels can cause the release or inhibition of a hormone (under negative feedback control) to maintain homeostasis. For example, osmoreceptors in the hypothalamus detect changes in blood osmolarity (the concentration of solutes in the blood plasma) and will signal the hypothalamus to release greater or lesser amounts of antidiuretic hormone (ADH) to keep the levels of solutes in the blood within normal limits. The control of blood glucose levels by the pancreas is another example of such stimulation. Responding directly to the level of glucose in the blood, cells in the pancreas release appropriate amounts of the hormones insulin and glucagon to maintain normal blood glucose levels. A final example of response to the level of a nutrient or ion in the blood is the regulation of levels of calcium by the parathyroid gland, which responds to changes in calcium levels in the blood with the secretion of varying levels of parathyroid hormone. All these mechanisms will be covered in greater detail later in this chapter.

The stimulus for the secretion of a hormone may also be the presence of another hormone produced by a different endocrine gland. Such hormonal stimuli often involve the hypothalamus, which produces releasing and inhibiting hormones that control the secretion of a variety of pituitary hormones, that in turn, may affect other endocrine glands in the body. These secretions are also controlled through negative feedback mechanisms. An example of such a negative feedback mechanism is the release of glucocorticoid hormones from the adrenal glands, as directed by the hypothalamus and pituitary gland (this will also be covered in more detail later in this chapter). As the secretion of glucocorticoid from the adrenal glands cause concentrations of this hormone in the blood to rise, the hypothalamus and pituitary gland reduce their release of hormones that caused this secretion, thus signaling to the adrenal glands to decrease glucocorticoid secretion (Figure 1).

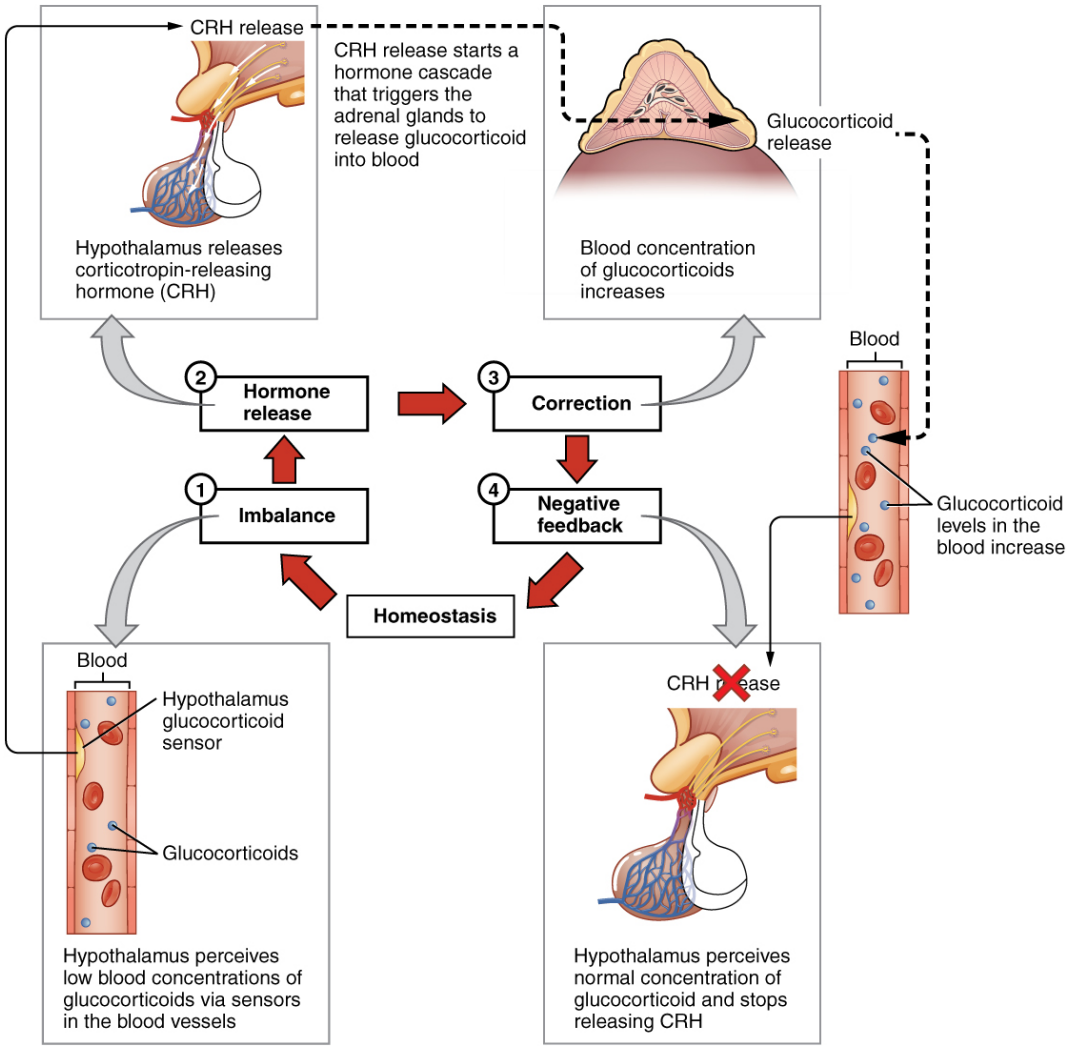


Figure 1. Negative Feedback Mechanism. *The release of adrenal glucocorticoids is stimulated by the release of hormones from the hypothalamus and pituitary gland. This signaling is inhibited when glucocorticoid levels become elevated by causing negative signals to the pituitary gland and hypothalamus.*

Types of Hormones

The hormones of the human body can be divided into two major groups on the basis of their chemical structure. Hormones derived from amino acids include amines, peptides, and proteins. Those derived from lipids include steroids (Table 1). These chemical groups affect a hormone's distribution, the type of receptors it binds to, and other aspects of its function.

Amine Hormones

Hormones derived from the modification of amino acids are referred to as amine hormones. Examples these include the metabolism-regulating thyroid hormones, as well epinephrine and norepinephrine, which play a role in the fight-or-flight response.

Peptide and Protein Hormones

Whereas the amine hormones are derived from a single amino acid, peptide and protein hormones consist of multiple amino acids that link to form an amino acid chain. Peptide hormones consist of short chains of amino acids, whereas protein hormones are longer polypeptides.

An example of a peptide hormone is antidiuretic hormone (ADH), a pituitary hormone important in fluid balance. Examples of protein hormones include growth hormone, which is produced by the pituitary gland, and follicle-stimulating hormone (FSH), which helps stimulate the maturation of eggs in the ovaries and sperm in the testes.

Steroid Hormones

The primary hormones derived from lipids are steroids. Steroid hormones are derived from the lipid cholesterol. For example, the reproductive hormones testosterone and the estrogens—which are produced by the gonads (testes and ovaries)—are steroid hormones. The adrenal glands produce the steroid hormone aldosterone, which is involved in osmoregulation, and cortisol, which plays a role in metabolism.

Like cholesterol, steroid hormones are not soluble in water (they are hydrophobic). Because blood is water-based, lipid-derived hormones must travel to their target cell bound to a transport protein. This more complex structure extends the half-life of steroid hormones to much longer than that of hormones derived from amino acids. A hormone's half-life is the time required for half the concentration of the hormone to be degraded. For example, the lipid-derived hormone cortisol has a half-life of approximately 60 to 90 minutes. In contrast, the amino acid–derived hormone epinephrine has a half-life of approximately one minute.

Pathways of Hormone Action

Although a given hormone may travel throughout the body in the bloodstream, it will affect the activity only of its target cells; that is, cells with receptors for that particular hormone. The message a hormone sends is received by a **hormone receptor**, a protein located either inside the cell or within the cell membrane. The receptor will process the message by initiating other signaling events or cellular mechanisms that result in the target cell's response. Hormone receptors recognize molecules with specific shapes and side groups, and respond only to those hormones that are recognized. The same type of receptor may be located on cells in different body tissues, and trigger somewhat different responses. Thus, the response triggered by a hormone depends not only on the hormone, but also on the target cell.

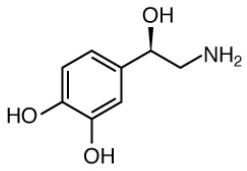
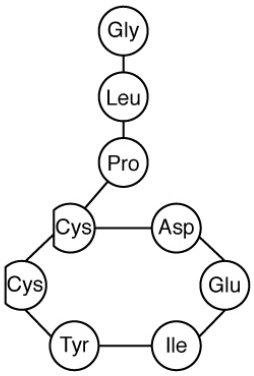
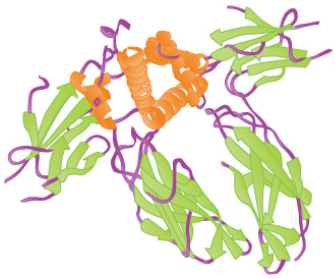
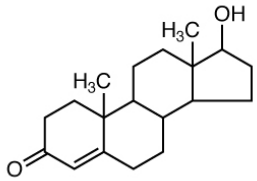
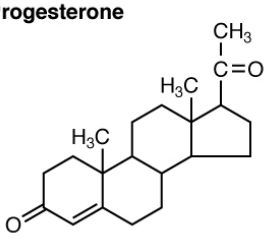
Hormone Class	Components	Example(s)
Amine Hormone	Amino acids with modified groups (e.g. norepinephrine's carboxyl group is replaced with a benzene ring)	**Norepinephrine** 
Peptide Hormone	Short chains of linked amino acids	**Oxytocin** 
Protein Hormone	Long chains of linked amino acids	**Human Growth Hormone** 
Steroid Hormones	Derived from the lipid cholesterol	**Testosterone**  **Progesterone** 

Table 1. Amine, Peptide, Protein, and Steroid Hormone Structure *(the structural formulae are not required as examinable material).*

Once the target cell receives the hormone signal, it can respond in a variety of ways. The response may include the stimulation of protein synthesis, activation or deactivation of enzymes, alteration in the permeability of the cell membrane, altered rates of mitosis and cell growth, and stimulation of the secretion of products. Moreover, a single hormone may be capable of inducing different responses in a given cell.

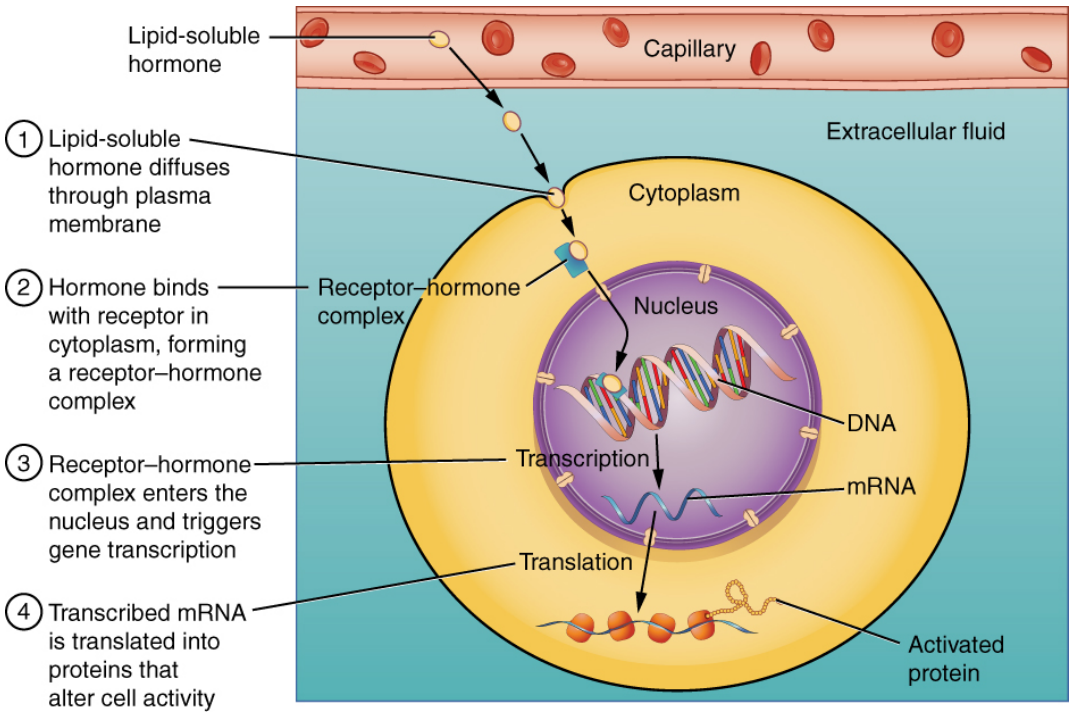


Figure 2. Binding of Lipid-Soluble Hormones. *A steroid hormone directly initiates the production of proteins within a target cell. Steroid hormones easily diffuse through the cell membrane. The hormone binds to its receptor in the cytosol, forming a receptor–hormone complex. The receptor–hormone complex then enters the nucleus and binds to the target gene on the DNA. Transcription of the gene creates a messenger RNA that is translated into the desired protein within the cytoplasm.*

Pathways Involving Intracellular Hormone Receptors

Intracellular hormone receptors are located inside the cell. Hormones that bind to this type of receptor must be able to cross the cell membrane. Steroid hormones are derived from cholesterol and therefore can readily diffuse through the lipid bilayer of the cell membrane to reach the intracellular receptor (Figure 2). Thyroid hormones are also lipid-soluble and can enter the cell. Both hormones bind to DNA within the nucleus and trigger protein synthesis. The particular proteins synthesized will exert an effect.

Pathways Involving Cell Membrane Hormone Receptors

Hydrophilic, or water-soluble, hormones are unable to diffuse through the lipid bilayer of the cell membrane and must therefore pass on their message to a receptor located at the surface of the cell (Figure 3). Except for thyroid hormones, which are lipid-soluble, all amino acid–derived hormones bind to cell membrane receptors that are located, at least in part, on the extracellular surface of the cell membrane. Therefore, they do not directly affect the production of proteins, but instead initiate a signaling cascade (a series of sequential activation of enzymes within the cell) that can trigger a wide variety of effects, from nutrient metabolism to the synthesis of different hormones and other products. The effects vary according to the type of target cell and which signaling cascade is activated inside the cell. Examples of hormones that use this mechanism include calcitonin, which is important for bone construction and regulating blood calcium levels, and glucagon, which affects blood glucose levels.

Overall, such signaling cascades significantly increase the efficiency, speed, and specificity of the hormonal response, as thousands of signaling events can be initiated simultaneously in response to a very low concentration of hormone in the bloodstream, so the action of the hormone can be rapid and substantial. Additionally, the duration of this type of hormone signal is short.

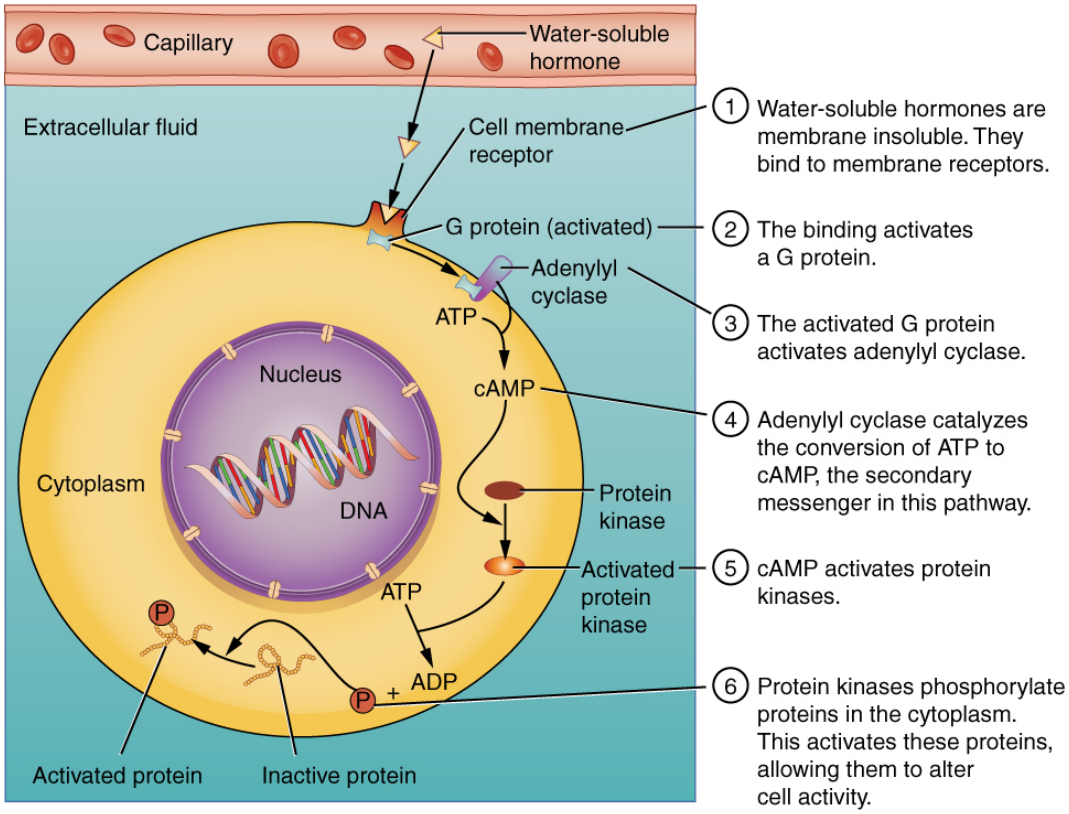


Figure 3. Binding of Water-Soluble Hormones. *Water-soluble hormones cannot diffuse through the cell membrane. These hormones must bind to a surface cell-membrane receptor. The receptor then initiates a cell-signaling pathway within the cell involving G proteins, adenylyl cyclase, the secondary messenger cyclic AMP (cAMP), and protein kinases. In the final step, these protein kinases phosphorylate proteins in the cytoplasm. This activates proteins in the cell that carry out the changes specified by the hormone. (The specific steps of the cell-signaling pathway are not required as examinable material).*

Comparison of the endocrine and nervous systems

Communication is a process in which a sender transmits signals to one or more receivers to control and coordinate actions. The part that the endocrine system plays in this has been stated, however in the human body, another major organ system participates in relatively "long distance" communication: the nervous system. Together, these two systems are primarily responsible for maintaining homeostasis in the body. Although both systems function to allow communication within the body, there are some significant differences in the anatomy and physiology and thus how the function is carried out between the endocrine and nervous systems.

The **nervous system** uses two types of intercellular communication—electrical and chemical signaling—either by the direct action of an electrical potential, or in the latter case, through the action of chemical neurotransmitters such as serotonin or norepinephrine. Neurotransmitters act locally and rapidly. When an electrical signal in the form of an action potential arrives at a synaptic terminal, it results in the release of neurotransmitters, which diffuse across the synaptic cleft (the gap between a sending neuron and a receiving neuron or muscle cell). Once the neurotransmitters interact (bind) with receptors on the receiving (post-synaptic) cell, the receptor stimulation is transduced into a response such as continued electrical signaling or modification of cellular response. The target cell responds within milliseconds of receiving the chemical "message"; this response then ceases very quickly once the neural signaling ends. In this way, neural communication enables body functions that involve quick, brief actions, such as movement, sensation, and cognition.

In contrast, the **endocrine system** uses just one method of communication: chemical signaling through **hormones**, which are secreted into the extracellular fluid. As previously stated, hormones are transported (primarily) via the bloodstream throughout the body, where they bind to receptors on target cells, inducing a characteristic response. As a result, endocrine signaling requires more time than neural signaling to prompt a

response in target cells, though the precise amount of time varies with different hormones. For example, the hormones released when you are confronted with a dangerous or frightening situation, called the fight-or-flight response, occur by the release of adrenal hormones—epinephrine and norepinephrine—within seconds. In contrast, it may take up to 48 hours for target cells to respond to certain reproductive hormones. Similarly, due to the mechanism of transmission of hormonal signals, the effect tends to last longer than a nervous stimulation.

Visit this link to watch an animation of the events that occur when a hormone binds to a cell membrane receptor. Direct link: http://openstaxcollege.org/l/hormonebind

In addition, endocrine signaling is typically less specific than neural signaling. The same hormone may play a role in a variety of different physiological processes depending on the target cells involved. For example, the hormone oxytocin promotes uterine contractions in women in labor. It is also important in breastfeeding and may be involved in the sexual response and in feelings of emotional attachment in both males and females.

In general, the nervous system involves quick responses to rapid changes in the external environment, and the endocrine system is usually slower acting—taking care of the internal environment of the body, maintaining homeostasis, and controlling reproduction (Table 2).

Watch this Crash Course video for an overview of the endocrine system! Direct link: https://youtu.be/eWHH9je2zG4

So how does the fight-or-flight response that was mentioned earlier happen so quickly if hormones are usually slower acting? It is because the two systems are connected. It is the fast action of the nervous system in response to the danger in the environment that stimulates the adrenal glands to secrete their hormones. As a result, the nervous system can cause rapid endocrine responses to keep up with sudden changes in both the external and internal environments when necessary.

Table 2. Major characteristics of endocrine and nervous systems

	Endocrine system	Nervous system
Signaling mechanism(s)	Chemical	Chemical / electrical
Primary chemical signal	Hormones	Neurotransmitters
Distance travelled by chemical signal	Long or short	Always short
Response time	Fast or slow	Always fast
Duration of response	Longer than nervous	Very short
Environment targeted	Internal	Internal and external

Part 2. Major endocrine organs and their secretions

The major endocrine glands are shown in Figure 4, and are listed along with their associated hormones and their effects in Table 3. Note that hypothalamic hormones are not listed in Table 3; these consist of releasing/inhibiting hormones for the anterior pituitary hormones, and two hormones that are made by the hypothalamus, but actually stored and then secreted by the posterior pituitary. They are discussed in more detail in the next section of this chapter.

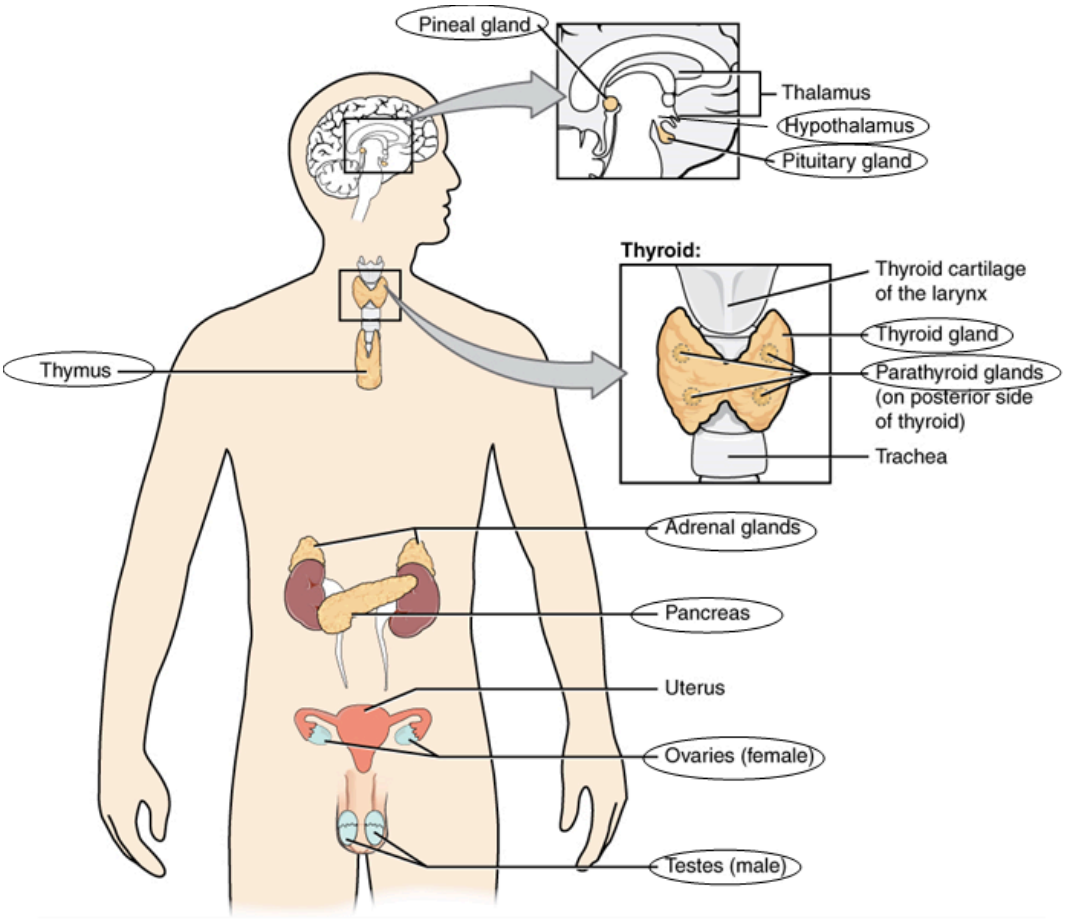


Figure 4. Endocrine System. *Endocrine glands and cells are located throughout the body and play an important role in homeostasis.*

Table 3: Endocrine Glands* and Their Major Hormones (*excluding the hypothalamus)

Endocrine gland	Associated hormones	Chemical class	Ultimate effect
Anterior pituitary	Growth hormone (GH)	Protein	Promotes growth of body tissues
Anterior pituitary	Prolactin (PRL)	Peptide	Promotes milk production
Anterior pituitary	Thyroid-stimulating hormone (TSH)	Glycoprotein	Stimulates thyroid hormone release
Anterior pituitary	Adrenocorticotropic hormone (ACTH)	Peptide	Stimulates hormone release by adrenal cortex
Anterior pituitary	Follicle-stimulating hormone (FSH)	Glycoprotein	Stimulates gamete production
Anterior pituitary	Luteinizing hormone (LH)	Glycoprotein	Stimulates androgen production by gonads
Posterior pituitary	Antidiuretic hormone (ADH)	Peptide	Stimulates water reabsorption by kidneys
Posterior pituitary	Oxytocin	Peptide	Stimulates uterine contractions during childbirth; may increase sperm motility and testosterone production; modulates social behaviour and social bonding
Thyroid	Thyroxine (T_4), triiodothyronine (T_3)	Amine	Generally increase basal metabolic rate
Thyroid	Calcitonin	Peptide	Reduces blood Ca^{2+} levels
Parathyroid	Parathyroid hormone (PTH)	Peptide	Increased blood Ca^{2+} levels
Adrenal cortex	Aldosterone	Steroid	Increases blood Na^{+} levels
Adrenal cortex	Cortisol, corticosterone, cortisone (CORT)	Steroid	Increase blood glucose levels
Adrenal medulla	Epinephrine (adrenaline), norepinephrine (noradrenaline)	Amine	Stimulate fight-or-flight response
Pineal	Melatonin	Amine	regulates sleep cycles
Pancreas	Insulin	Protein	Reduces blood glucose levels
Pancreas	Glucagon	Protein	Increases blood glucose levels
Testes	Testosterone (T)	Steroid	Stimulates development of male secondary sex characteristics and sperm production
Ovaries	Estrogens, progesterone	Steroid	Stimulate development of female secondary sex characteristics and prepare the body for childbirth

Hypothalamus

The Hypothalamus and Pituitary Gland

The hypothalamus–pituitary complex can be thought of as the "command center" of the endocrine system for basically two reasons. Besides secreting several hormones that directly produce responses in target tissues,

it secretes hormones that regulate the synthesis and secretion of hormones of other endocrine glands. In addition, the hypothalamus–pituitary complex coordinates the messages of the endocrine and nervous systems. In many cases, a stimulus received by the nervous system must pass through the hypothalamus–pituitary complex to be translated into hormones that can initiate a response.

The **hypothalamus** is a structure of the diencephalon of the brain located anterior and inferior to the thalamus (Figure 5). The hypothalamus is anatomically and functionally related to the **pituitary gland** (or hypophysis), a bean-sized organ suspended from it by a stem called the **infundibulum** (or pituitary stalk).

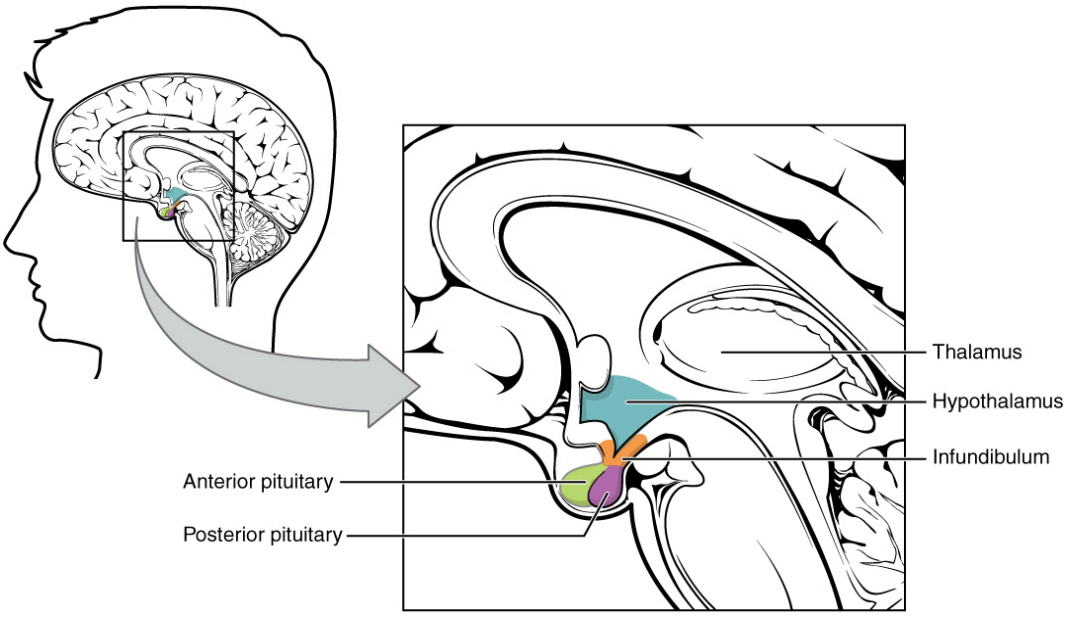


Figure 5. Hypothalamus–Pituitary Complex. *The hypothalamus region lies inferior and anterior to the thalamus. It connects to the pituitary gland by the stalk-like infundibulum. The pituitary gland consists of an anterior and posterior lobe, with each lobe secreting different hormones in response to signals from the hypothalamus.*

The **pituitary gland** is cradled within a cup-like hollow in the sphenoid bone of the skull. It consists of two lobes that arise from different parts of embryonic tissue: the anterior pituitary (also known as the adenohypophysis) is glandular tissue that develops from the primitive digestive tract, whereas the posterior pituitary (neurohypophysis) is neural tissue that is essentially an extension of the hypothalamus.

Hormonal secretion by the hypothalamus

All of the hormones that the hypothalamus produces either are directly secreted by the hypothalamus and control the release of hormones by the anterior pituitary (six of these will be discussed below) or are stored in and released by the posterior pituitary (there are two, as presently discussed).

Hypothalamic control of anterior pituitary gland secretion

The secretion of all hormones from the anterior pituitary is regulated by two classes of hormones secreted by the hypothalamus: releasing hormones that stimulate the secretion of hormones from the anterior pituitary, and inhibiting hormones that inhibit secretion (i.e., the anterior pituitary never increases or decreases the release of one of its hormones, without being signaled to do so by the hypothalamus).

Hypothalamic hormones that control the anterior pituitary are secreted by neurons in the hypothalamus and enter the anterior pituitary through blood vessels (Figure 6). Within the infundibulum (the connecting tissue between the hypothalamus and the pituitary) is a bridge of capillaries that connects the hypothalamus to the anterior pituitary. This network, called the **hypophyseal portal system**, allows hypothalamic hormones to be transported to the anterior pituitary without first entering the systemic circulation. Hormones produced by the anterior pituitary in response to these releasing or inhibiting hormones sent by the hypothalamus then enter a secondary capillary plexus, and from there drain into the general circulation.

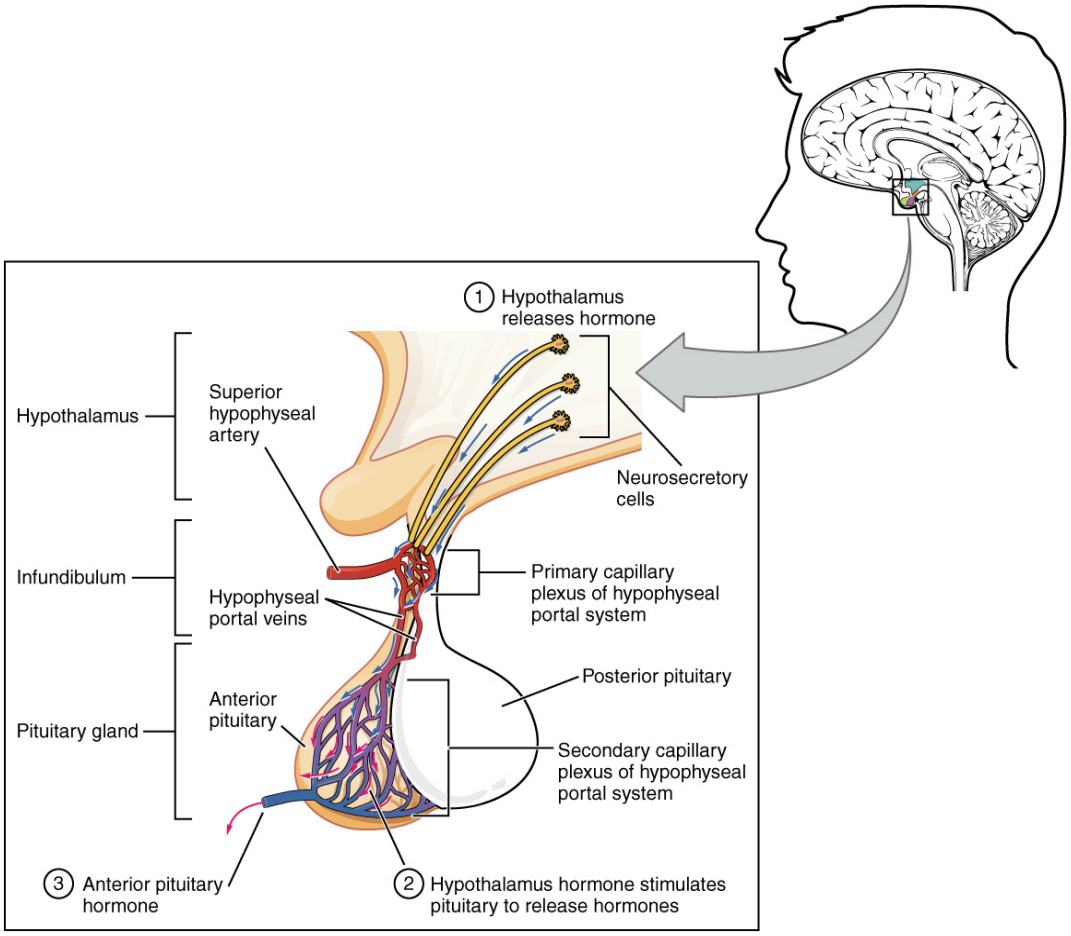


Figure 6. Anterior Pituitary. *The hypothalamus produces separate hormones that stimulate or inhibit hormone production in the anterior pituitary. Hormones from the hypothalamus reach the anterior pituitary via the hypophyseal portal system.*

Four of the hormones the hypothalamus produces act as releasing factors which stimulate the secretion of five separate hormones from the anterior pituitary gland. These four releasing hormones are named after the pituitary hormones whose secretions they stimulate:

- Adrenocorticotropic hormone releasing hormone (ACTHRH, or CRH for corticotropin releasing hormone)
- Thyroid stimulating hormone releasing hormone (TSHRH, or TRH for thyrotropin releasing hormone)
- Growth hormone releasing hormone (GHRH)
- Gonadotropin releasing hormone (GnRH), which stimulates the release of hormones known as gonadotropins: follicle stimulating hormone (FSH) and luteinizing hormone (LH)

The hypothalamus also produces inhibiting factors, including *growth hormone inhibiting hormone* (GHIH) and *prolactin inhibiting hormone* (PIH).

Cells of the hypothalamus also produce hormones that are stored in and secreted by the posterior pituitary, rather than being secreted from the hypothalamus itself. The hypothalamus produces the hormones oxytocin and antidiuretic hormone, both of which are transported, stored in and then released from the posterior pituitary gland (discussed in the section describing secretions of the posterior pituitary) (Fig. 7).

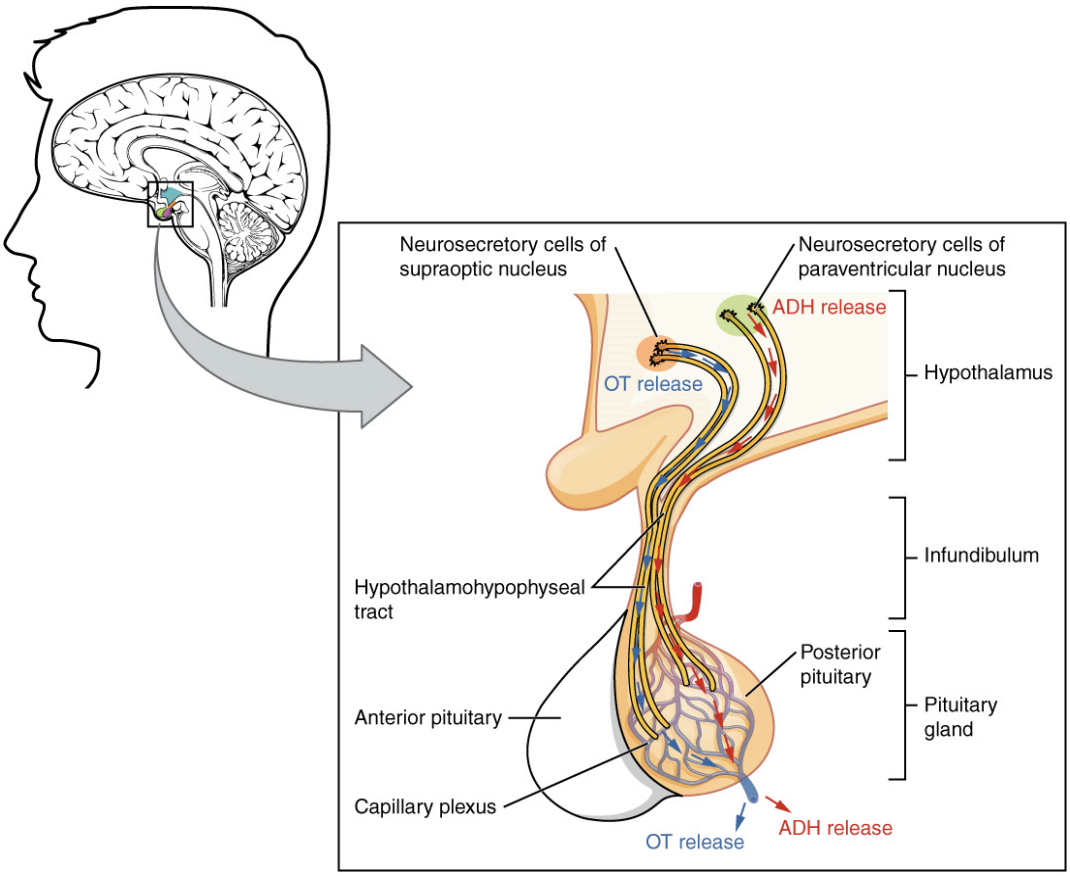


Figure 7. Posterior Pituitary. *Neurosecretory cells in the hypothalamus release oxytocin (OT) or ADH into the posterior lobe of the pituitary gland. These hormones are stored or released into the blood via the capillary plexus.*

Hormonal secretion by the anterior pituitary (Table 4)

The anterior pituitary produces several hormones, including growth hormone (GH) and prolactin, neither of which affect other endocrine glands, and thyroid-stimulating hormone (TSH), adrenocorticotropic hormone (ACTH), follicle-stimulating hormone (FSH), and luteinizing hormone (LH). Of the hormones of the anterior pituitary, these last four (TSH, ACTH, FSH, and LH) are collectively referred to as tropic hormones (the suffix "tropic" = "turning towards / having an influence on") because they travel to and affect the function of other endocrine glands.

Growth Hormone

The endocrine system regulates the growth of the human body, protein synthesis, and cellular replication. A major hormone involved in this process is growth hormone (GH), also called somatotropin ("soma" means body; "tropin" means going towards/having an effect on) —a protein hormone produced and secreted by the anterior pituitary gland. Its primary function is anabolic; it promotes protein synthesis and tissue building, increases catabolism of fats and generally slows down the catabolism of carbohydrates (thus helping to maintain blood glucose levels) (Figure 8). GH levels are controlled by the release of growth hormone-releasing hormone (GHRH) and growth hormone-inhibiting hormone (GHIH, also known as somatostatin) from the hypothalamus.

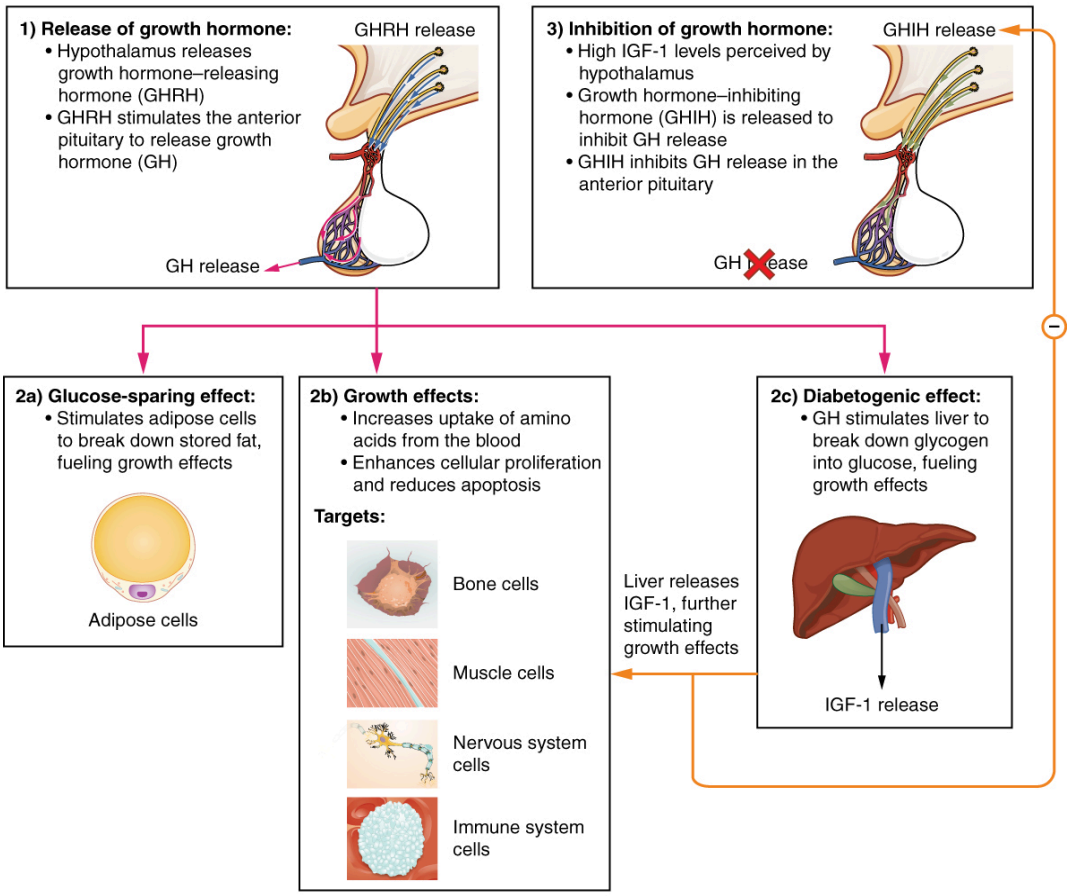


Figure 8. Hormonal Regulation of Growth. *Growth hormone (GH) directly accelerates the rate of protein synthesis in skeletal muscle and bones. Insulin-like growth factor 1 (IGF-1) is activated by growth hormone and indirectly supports the formation of new proteins in muscle cells and bone.*

Disorders of Growth Hormone Secretion

Any abnormal levels of secretion of GH would affect all the above listed processes. In general, any underproduction of a hormone is called hyposecretion, and an overproduction is called hypersecretion.

Hyposecretion of GH during childhood (that is, during the growing years), results in a condition known as **pituitary dwarfism**. Under these circumstances, the epiphyseal plates in the long bones close before the person has grown to normal height. The other body organs also do not grow normally, so a person with pituitary dwarfism has many childlike characteristics. The situation is readily treated by administering growth hormone to the individual before the epiphyseal plates close.

Growth hormone is produced throughout life, and is necessary for normal metabolism and maintenance of body tissues in the adult. A **hyposecretion of GH in the adult** results in a condition known as **Simmonds' disease**. This is characterized by extreme weakness, a wasting of body tissues and a loss of weight.

If there is a **hypersecretion of GH during childhood**, the individual will be abnormally large. This condition is known as **gigantism**.

In an adult, GH can no longer lengthen long bones, but excessive amounts can cause very slow growth of various parts of the body. **Hypersecretion of GH** results in a condition known as **acromegaly,** which is characterized by the very slow continual growth and thickening of bones (especially those of the hands, feet, cheeks and jaws) and also growth/thickening of organs such as the heart, liver and kidneys, as well as other tissues such as eyelids, forehead, lips, tongue and nose.

Prolactin

As its name implies, **prolactin** (**PRL**) promotes lactation (milk production) in women. During pregnancy, it contributes to the development of the mammary glands, and after birth, it stimulates the mammary glands to

produce breast milk. (As will be noted later, the let-down (release) of milk from the breasts occurs in response to stimulation from oxytocin.)

In a non-pregnant woman, prolactin secretion is inhibited by prolactin-inhibiting hormone (PIH), which is actually the neurotransmitter dopamine, and is released from neurons in the hypothalamus. Only during pregnancy do prolactin levels rise in response to prolactin-releasing hormone (PRH) from the hypothalamus.

The Tropic Hormones

Thyroid-Stimulating Hormone

The activity of the thyroid gland is regulated by **thyroid-stimulating hormone** (**TSH**), also called thyrotropin, released from the anterior pituitary. In turn, TSH is released from the anterior pituitary in response to thyrotropin-releasing hormone (TRH or TSHRH) from the hypothalamus. As discussed shortly, TSH triggers the secretion of thyroid hormones by the thyroid gland. In a classic negative feedback mechanism, elevated levels of thyroid hormones in the bloodstream then trigger a decrease in production of TRH and subsequently TSH.

Adrenocorticotropic Hormone

The **adrenocorticotropic hormone** (**ACTH**), also called corticotropin, stimulates the adrenal cortex (the outer layer of the adrenal glands) to secrete corticosteroid hormones such as cortisol.

The release of ACTH is regulated by the corticotropin-releasing hormone (CRH) from the hypothalamus in response to normal physiologic rhythms (The release of ACTH typically peaks in the morning, and reaches its lowest levels in late evening). A variety of stressors can also influence its release, and the role of ACTH in the stress response is discussed later in this unit.

Follicle-Stimulating Hormone and Luteinizing Hormone

Several endocrine glands secrete a variety of hormones that control the development and regulation of the reproductive system (these glands include the anterior pituitary, the adrenal cortex, and the gonads – the testes in males and the ovaries in females). Much of the development of the reproductive system occurs during puberty and is marked by the development of sex-specific characteristics in both male and female adolescents. Puberty is initiated by **gonadotropin-releasing hormone** (**GnRH**), a hormone produced and secreted by the hypothalamus. GnRH stimulates the anterior pituitary to secrete **gonadotropins**—hormones that regulate the function of the gonads. The levels of GnRH are regulated through a negative feedback mechanism; high levels of reproductive hormones inhibit the release of GnRH. Throughout life, gonadotropins regulate reproductive function and, in the case of women, the onset and cessation of reproductive capacity.

The gonadotropins include two hormones: 1) **Follicle-stimulating hormone** (**FSH**) which stimulates the production and maturation of sex cells, or gametes (ova in females and sperm in males). FSH also promotes follicular growth; these follicles then release estrogens in the ovaries. 2) **Luteinizing hormone** (**LH**) triggers ovulation, as well as the production of estrogens and progesterone by the ovaries. LH stimulates production of testosterone by the testes.

Hormonal secretion by the posterior pituitary (Table 5)

The posterior pituitary is actually an extension of neurons that originate in two specific nuclei (clusters of neuronal cell bodies) in the hypothalamus. While the cell bodies of these neurons rest in the hypothalamus, their axons descend as the hypothalamic–hypophyseal tract within the infundibulum, and end in axon terminals that comprise the posterior pituitary (Figure 7).

The posterior pituitary gland does not produce hormones, but rather stores and secretes hormones produced by the hypothalamus. Neuronal cell bodies of one group of cells in the hypothalamus produces the hormone oxytocin, whereas neuronal cell bodies of another group of cells produces antidiuretic hormone (ADH). These hormones then travel along the axons belonging to the respective neurons into storage sites in the axon terminals of the posterior pituitary. In response to later signals from the hypothalamus, the hormones are then released from the axon terminals into the bloodstream.

Oxytocin

When fetal development is complete, the peptide-derived hormone oxytocin (tocia- = “childbirth”) stimulates

uterine contractions and dilation of the cervix. Oxytocin is continually released throughout childbirth through a positive feedback mechanism that continues until birth.

Although the mother's high blood levels of oxytocin begin to decrease immediately following birth, oxytocin continues to play a role in maternal and newborn health. First, oxytocin is necessary for the milk ejection reflex (commonly referred to as "let-down") in breastfeeding women. Secondly, in both males and females, oxytocin is thought to contribute to parent–newborn bonding, known as attachment. In general, oxytocin is also thought to be involved in feelings of love and closeness, as well as in the sexual response.

Antidiuretic Hormone (ADH)

ADH is an important hormone of the urinary system. The solute concentration of the blood, or blood osmolarity, may change in response to the consumption of certain foods and fluids, as well as in response to disease, injury, medications, or other factors. Blood osmolarity is constantly monitored by **osmoreceptors**—specialized cells within the hypothalamus that are particularly sensitive to the concentration of sodium ions and other solutes.

In response to high blood osmolarity, which can occur during dehydration or following a very salty meal, the osmoreceptors in the hypothalamus signal the posterior pituitary to release **antidiuretic hormone** (**ADH**). The target cells of ADH are located in the tubular cells of the kidneys. Its effect is to increase epithelial permeability to water, allowing increased water reabsorption. A greater concentration of water in the blood results in a reduced concentration of solutes. ADH is also known as vasopressin because, in very high concentrations, it causes constriction of blood vessels, which increases blood pressure by increasing peripheral resistance. The release of ADH is controlled by a negative feedback mechanism. As blood osmolarity decreases, the hypothalamic osmoreceptors sense the change and prompt a corresponding decrease in the secretion of ADH. As a result, less water is reabsorbed from the urine filtrate.

Disorder of Antidiuretic Hormone Secretion

Interestingly, drugs can affect the secretion of ADH. For example, alcohol consumption inhibits the release of ADH, resulting in increased urine production that can eventually lead to dehydration and a hangover. A disease called **diabetes insipidus** is characterized by chronic **underproduction of ADH** that causes chronic dehydration. Because little ADH is produced and secreted, not enough water is reabsorbed by the kidneys. Although patients feel thirsty, and increase their fluid consumption, this doesn't effectively decrease the solute concentration in their blood because ADH levels are not high enough to trigger water reabsorption in the kidneys. Electrolyte imbalances can occur in severe cases of diabetes insipidus.

Important note: The term "diabetes" simply means overflow, referring to the copious amounts of urine produced due to an excessive loss of water. In the case of diabetes *insipidus*, the excessive loss of water into the urine is caused by the underproduction of ADH, leading to pale (insipid) urine. In the case of diabetes *mellitus*, the overflow is due to the presence of glucose in the urine, caused by the lack of sufficient or effective insulin, which will be covered later in this chapter. These are two separate diseases caused by the insufficiency of two different hormones.

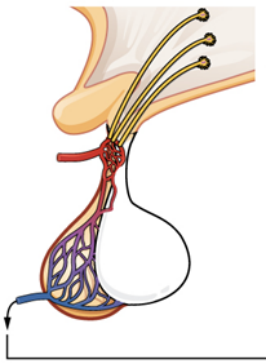


Table 4. Major anterior pituitary hormones, their hypothalamic releasing/inhibiting hormones, targets and some effects.

Pituitary hormone	Hypothalamic releasing (+) or inhibiting (-) hormone	Target	Effects Include
Tropic hormones:			
Lutenizing hormone (LH)	Gonadotropin-releasing hormone (GnRH) (+)	Reproductive system	Stimulates production of sex hormones by the gonads
Follicle-stimulating hormone (FSH)	GnRH (+)	Reproductive system	Stimulates production of sperm and eggs
Thyroid-stimulating hormone (TSH)	Thyroid releasing hormone (TRH) (+)	Thyroid gland	Stimulates the release of thyroid hormone (TH). TH regulates metabolism.
Adrenocorticotropic hormone (ACTH)	Corticotropin-releasing hormone (CRH) (+)	Adrenal cortex	Induces targets to produce glucocorticoids, which regulate metabolism and then stress response
Non-tropic hormones:			
Prolactin (PRL)	Prolactin-releasing hormone (PRH) (+) or prolactin-inhibiting hormone (PIH) (-)	Mammary glands	Promotes milk production
Growth hormone (GH)	Growth hormone-releasing hormone (GHRH) (+) or growth hormone-inhibiting hormone (GHIH) (-)	Liver, bone, muscles	Induces targets to produce compounds that stimulate body growth and a higher metabolic rate.

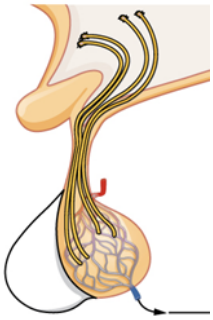


Table 5. Posterior pituitary hormones (made in hypothalamus), targets, and some effects

Released from posterior pituitary	Targets include	Effects include
Oxytocin	Female reproductive system	Triggers uterine contractions during childbirth
Antidiuretic hormone (ADH)	Kidneys, sweat glands, circulatory system	Stimulates water reabsorption in the kidneys

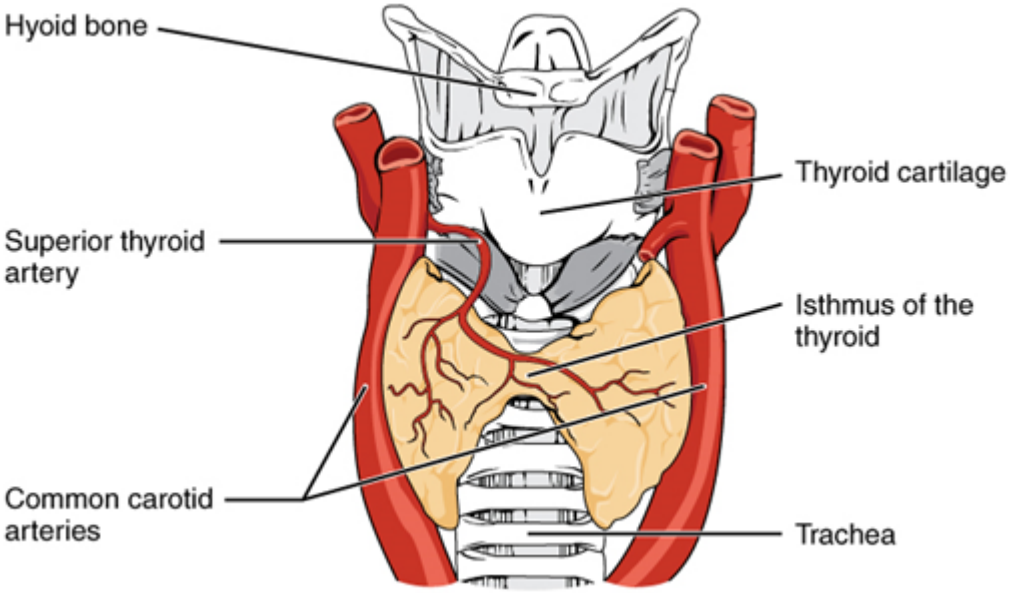


Figure 9. Anterior view of thyroid gland.

The Thyroid Gland

A butterfly-shaped organ, the **thyroid gland** is located anterior to the trachea, just inferior to the larynx (Figure 9). The medial region, called the isthmus, is flanked by wing-shaped left and right lobes. Each of the thyroid lobes are embedded with parathyroid glands, primarily on their posterior surfaces. The thyroid gland produces and secretes two hormones: thyroid hormone and calcitonin.

Thyroid Hormone

As with many other hormones, the release of thyroid hormone is under negative feedback control, as a result of stimulation by TSH (thyroid stimulating hormone) released by the anterior pituitary. Recall that TSH is stimulated in turn by the release of TSHRH (thyroid stimulating hormone releasing hormone) released by the hypothalamus. Thyroid hormone actually consists of two slightly different compounds: T_3 (triiodothyronine) and T_4 (thyroxine). These compounds are often referred to as metabolic hormones because their levels influence the body's basal metabolic rate, which is the amount of energy used by the body to make ATP at rest. When T_3 and T_4 bind to intracellular receptors, they cause an increase in nutrient breakdown (thus causing a breakdown of fats and carbohydrates), and the increased use of oxygen to produce ATP. In addition, T_3 and T_4 initiate the transcription of genes involved in glucose oxidation. The process is inefficient, and an increased amount of heat is released as a byproduct of the increased rate of cellular respiration. This so-called calorigenic effect (calor- = "heat") raises body temperature.

Adequate levels of thyroid hormones are also required for protein synthesis and for fetal and childhood tissue development and growth. They are especially critical for normal development of the nervous system both *in utero* and in early childhood, and they continue to support neurological function in adults.

These thyroid hormones also have a complex interrelationship with reproductive hormones, and deficiencies can influence libido, fertility, and other aspects of reproductive function. Finally, thyroid hormones increase the body's sensitivity to catecholamines (epinephrine and norepinephrine) from the adrenal medulla by upregulation of receptors in the blood vessels and the heart. When levels of T_3 and T_4 hormones are excessive, this effect accelerates the heart rate, strengthens the heartbeat, and increases blood pressure. Because thyroid hormones regulate metabolism, heat production, protein synthesis, and many other body functions, thyroid disorders can have severe and widespread consequences.

Disorders of Thyroid Hormone Secretion

Hypersecretion of the thyroid hormones may have numerous causes, including an excess production of TSH (thyroid stimulating hormone) by the anterior pituitary gland, the presence of a tumor in the thyroid gland itself, or an immune disorder.

Hypersecretion of thyroid hormone caused by an immune disorder is known as **Graves' disease.** Also known as exophthalmic goiter, the effects are similar in children and adults. One symptom is that the thyroid gland becomes enlarged – a condition called **goiter** (explained below). There is also a very high metabolic rate, which causes the person to have a high pulse rate, lose weight and become nervous and irritable. Inflammation and swelling occurs in the tissue behind the eyes, causing the eyes to protrude. Treatment involves either surgically removing part of the thyroid gland, or suppressing thyroid secretions with drugs.

Hyposecretion of thyroid hormones in childhood leads to **cretinism.** In this condition, the skeleton fails to grow and mature, and the brain does not develop fully. Other symptoms include general lethargy, a slow heart rate and a low body temperature. In addition, there is a tendency to easily gain weight.

Hyposecretion of thyroid hormone in adulthood is called **myxedema.** This is characterized by swelling and puffiness of the facial tissue (caused by an increased accumulation of hydrophilic sugar protein complexes in the skin) and lowered body metabolism. There is also a tendency to easily gain weight. Because the brain is already fully developed, there is no mental retardation.

A symptom of many thyroid disorders is a **goiter**, which is an increase in the overall size of the thyroid gland (Figure 10). Interestingly, a goiter can arise whether the thyroid gland is synthesizing too much or too little of thyroid hormone.

Consistent overstimulation of the thyroid gland that then produces more than normal amounts of thyroid hormone can occur in diseases such as Grave's disease. Such stimulation may result in an increase in the size of the thyroid gland (= goiter).

On the other hand, underactivity of the thyroid gland would result in lower amounts of thyroid hormone in the blood. This would decrease the negative feedback effect that thyroid hormone has on the production of TRH from the hypothalamus and TSH from the anterior pituitary. The levels of these compounds in the blood will then rise and stay elevated as long as the level of thyroid hormone remains low. This will continuously stimulate the thyroid gland, causing it to increase in size (= goiter). One of the causes for the inability of the thyroid gland to produce thyroid hormone in the first place is a condition known as **simple goiter**, which occurs when the body's intake of iodine (which is required for the production of thyroid hormone) is not sufficient for its needs.

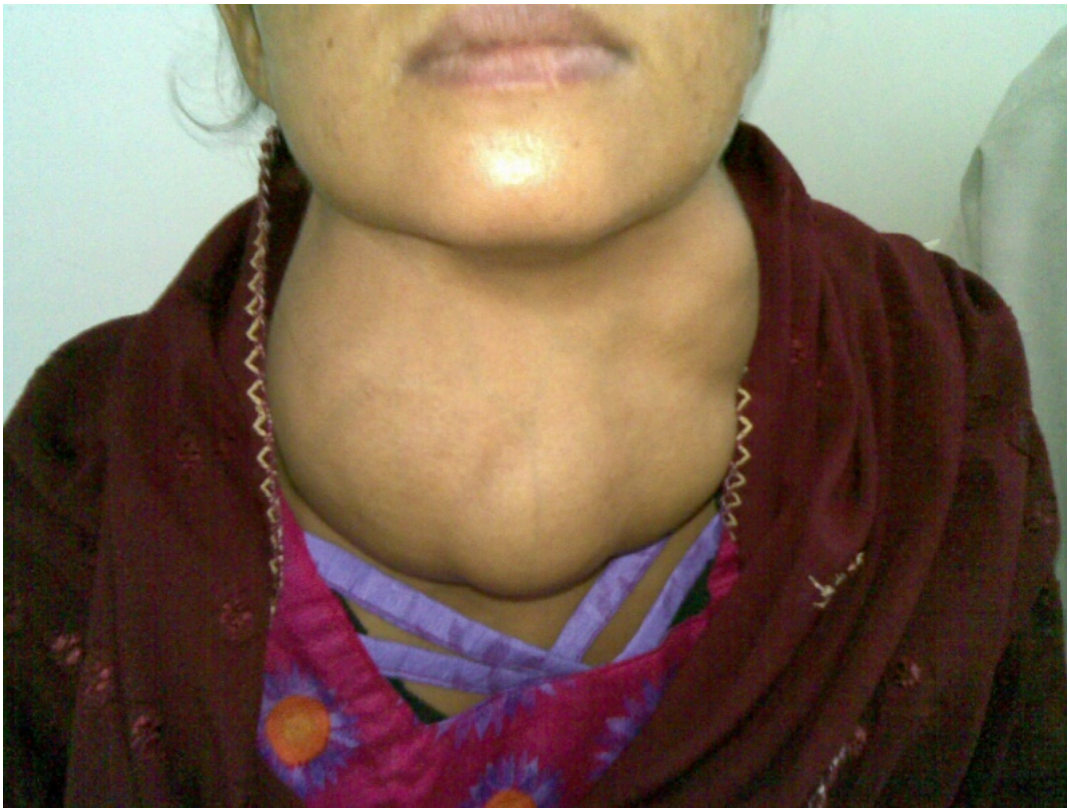


Figure 10. Goiter. *(credit: "Almazi"/Wikimedia Commons)*

Calcitonin

The thyroid gland also secretes a hormone called calcitonin that is released in response to a rise in blood calcium levels. It appears to have a function in decreasing blood calcium concentrations by:

- Inhibiting the activity of osteoclasts, bone cells that release calcium into the circulation by degrading bone matrix
- Increasing osteoblastic activity (thereby increasing deposition of calcium into bones)
- Decreasing calcium absorption in the intestines
- Increasing calcium loss in the urine

However, these functions are usually not significant in maintaining calcium homeostasis (people with long term increased or decreased calcitonin secretion due to disease usually do not show abnormal blood calcium levels), so the importance of calcitonin is not entirely understood. The production of calcitonin does however respond to levels of calcium in the blood (Figure 12), and pharmaceutical preparations of calcitonin are sometimes prescribed to reduce osteoclast activity in people with osteoporosis and to reduce the degradation of cartilage in people with osteoarthritis.

Calcium is critical for many biological processes, acting as a second messenger in many signaling pathways, and essential for muscle contraction, nerve impulse transmission, and blood clotting. The necessary tight regulation of blood calcium levels is mainly carried out by the parathyroid glands.

The Parathyroid Glands

The **parathyroid glands** are tiny, round structures usually found embedded in the posterior surface of the thyroid gland (Figure 11). A thick connective tissue capsule separates the glands from the thyroid tissue. Most people have four parathyroid glands. The primary functional cells of the parathyroid glands cells produce and

secrete the **parathyroid hormone** (**PTH**) (also known as parathormone), the major hormone involved in the regulation of blood calcium levels.

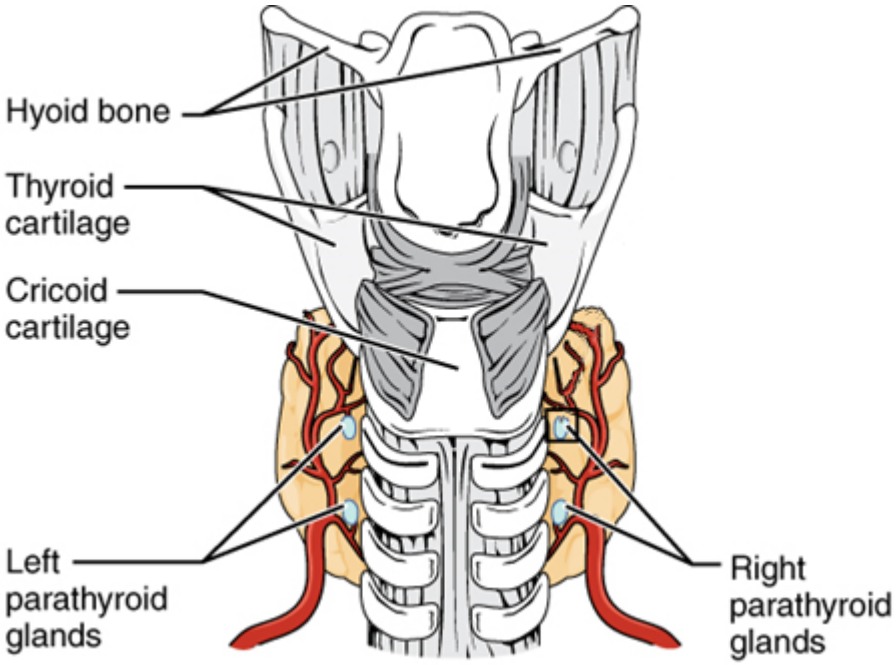


Figure 11. Parathyroid Glands. *The small parathyroid glands are embedded in the posterior surface of the thyroid gland.*

The parathyroid glands produce and secrete PTH, a peptide hormone, in response to low blood calcium levels (Figure 12). PTH secretion causes the increase in blood calcium via the following mechanisms:

- Stimulating the activity of osteoclasts, bone cells that release calcium into the circulation by degrading bone matrix
- Inhibiting osteoblastic activity (thereby decreasing deposition of calcium into bones)
- Increasing calcium absorption in the intestines by initiating the production of the steroid hormone calcitriol (also known as 1,25-dihydroxyvitamin D), which is the active form of vitamin D3, in the kidneys. Calcitriol then stimulates increased absorption of dietary calcium by the intestines.
- Decreasing calcium loss in the urine by causing increased reabsorption of calcium (and magnesium) in the kidney tubules from the urine filtrate

A negative feedback mechanism regulates the levels of PTH, with rising blood calcium levels inhibiting further release of PTH.

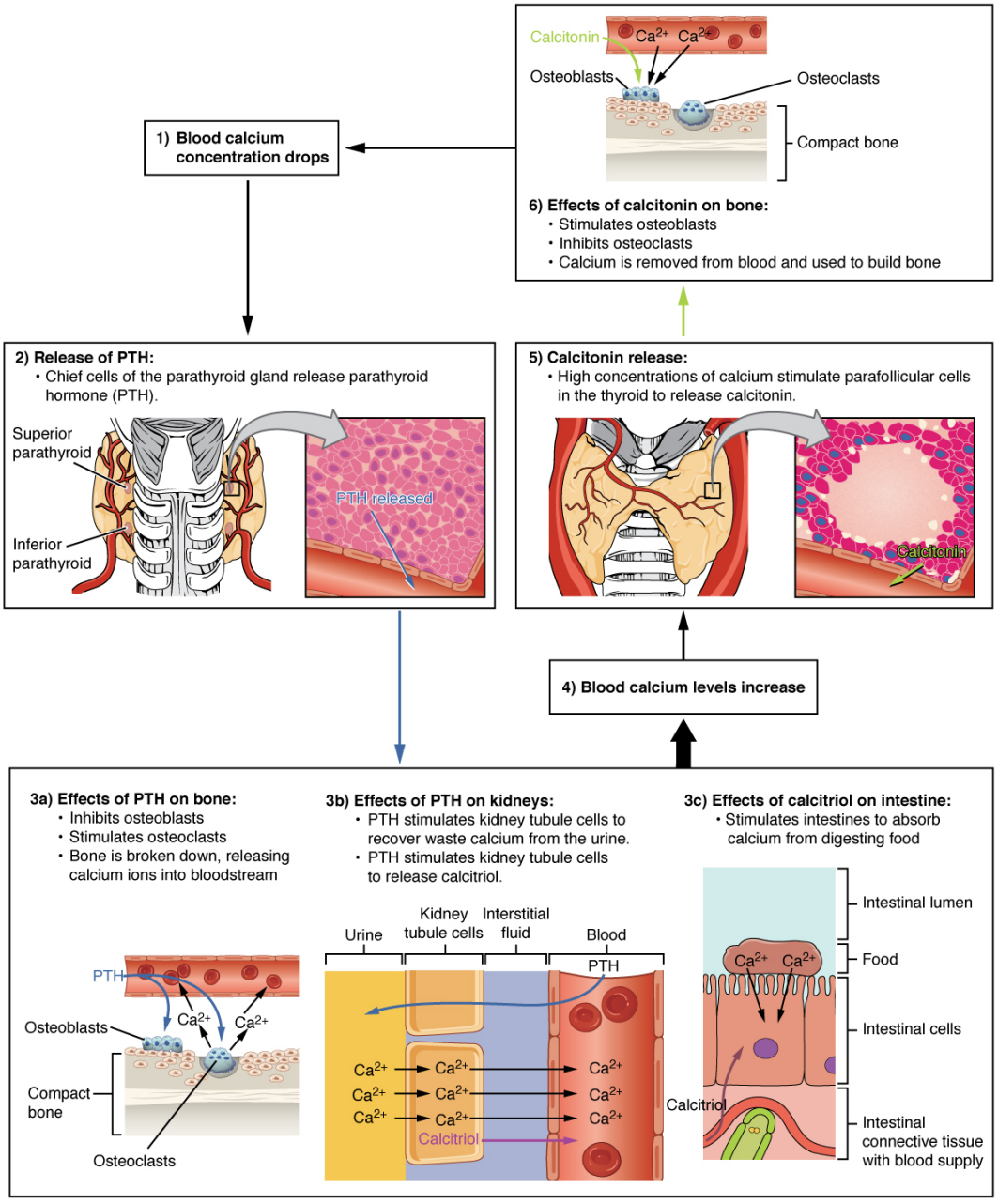


Figure 12. Hormones Involved in Maintaining Blood Calcium Homeostasis. *When blood calcium levels drop below the normal range, the release of parathyroid hormone increases and acts to increase calcium levels. When blood calcium levels rise beyond normal levels, the release of calcitonin from the thyroid gland increases and acts to decrease calcium levels.*

Disorders of parathyroid hormone secretion

Abnormally high activity of the parathyroid gland can cause **hyperparathyroidism**, a disorder caused by an overproduction of PTH that results in excessive calcium reabsorption from bone. Hyperparathyroidism can significantly decrease bone density, leading to spontaneous fractures or deformities. As blood calcium levels rise, cell membrane permeability to sodium is decreased (the exact mechanism has not yet been clarified), and the responsiveness of the nervous system is reduced. At the same time, calcium deposits may collect in the body's tissues and organs, impairing their functioning.

In contrast, abnormally low blood calcium levels may be caused by parathyroid hormone deficiency, called **hypoparathyroidism**, which may develop following injury or surgery involving the thyroid gland. Low blood calcium increases membrane permeability to sodium, resulting in muscle twitching, cramping, spasms, or

convulsions, a condition known as tetany. Severe deficits can paralyze muscles, including those involved in breathing, and can be fatal.

The Adrenal Glands

The **adrenal glands** are wedges of glandular and neuroendocrine tissue adhering to the top of the kidneys by a fibrous capsule (Figure 13). The adrenal glands have a rich blood supply and experience one of the highest rates of blood flow in the body. The adrenal gland consists of an outer cortex of glandular tissue and an inner medulla of nervous tissue.

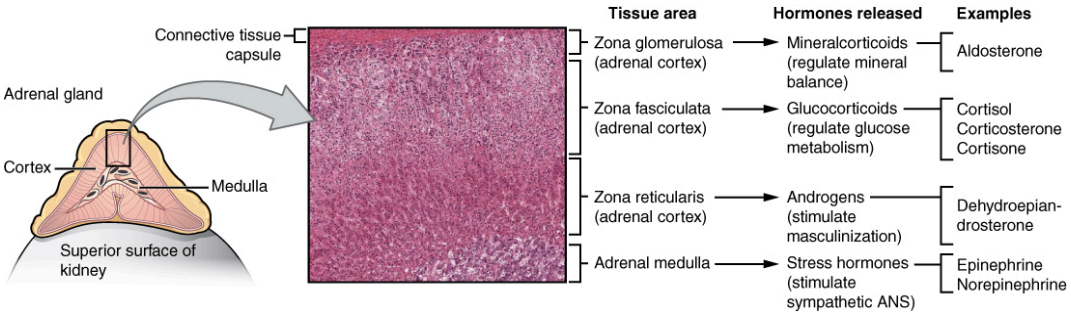


Figure 13. Adrenal Glands. *Both adrenal glands sit atop the kidneys and are composed of an outer cortex and an inner medulla, all surrounded by a connective tissue capsule. The cortex can be subdivided into additional zones, all of which produce different types of hormones. LM × 204. (Micrograph provided by the Regents of University of Michigan Medical School © 2012). (Knowledge of the zones' names and responsibility for particular hormone production is not required as examinable material.)*

Hormones of the adrenal cortex

Mineralocorticoids

The most superficial region of the adrenal cortex (zona glomerulosa) produces a group of hormones collectively referred to as **mineralocorticoids** because of their effect on body minerals, especially sodium and potassium. These hormones are essential for fluid and electrolyte balance.

Aldosterone is the major mineralocorticoid. It is important in the regulation of the concentration of sodium and potassium ions in urine, sweat, and saliva. For example, it is released in response to elevated blood K^+, low blood Na^+, low blood pressure, low blood volume and activation of the renin-angiotensin-aldosterone system (RAAS) (this hormone will be discussed again during the unit dealing with the renal system). In response, aldosterone increases the excretion of K^+ and the retention of Na^+, which in turn increases blood volume and blood pressure.

Glucocorticoids

The cells of the middle layer (zona fasciculata) produce hormones called **glucocorticoids** because of their role in glucose metabolism. The most important of these is **cortisol**. In response to long-term stressors, the hypothalamus secretes CRH, which in turn triggers the release of ACTH by the anterior pituitary. ACTH triggers the release of cortisol. The overall effect is to inhibit tissue building while stimulating the breakdown of stored nutrients to maintain adequate fuel supplies. In conditions of long-term stress, for example, cortisol promotes the catabolism of glycogen to glucose, the catabolism of stored triglycerides into fatty acids and glycerol, and the catabolism of muscle proteins into amino acids. Cortisol increases the body's resistance to stress by

increasing muscle metabolism, maintaining the excitability of nerves and increasing the amount of sugars in the body by promoting **gluconeogenesis**, the conversion of fats and proteins into sugars.

Androgens

The innermost layer of the cortex (zona reticularis) produces small amounts of sex hormones called **androgens,** which are converted into **testosterone** and **estrogen** in the tissues. The adrenal cortex serves as the source of sex hormones in an individual until the gonads mature at puberty. The sex hormones of the adrenal cortex play a role in the development of secondary sex characteristics in both males and females. In males, they further increase muscle development. It is interesting to note that the production of the cortical sex hormones is under the influence of adrenocorticotropic hormone from the anterior pituitary gland, and not follicle stimulating hormone or luteinizing hormone.

Hormones of the Adrenal Medulla

Epinephrine and norepinephrine (also sometimes called adrenalin and noradrenalin): These two hormones are both involved in the response to fear, excitement and danger. They both increase blood pressure, and the rate and depth of breathing. Epinephrine, however, increases heart rate and blood sugar levels, whereas norepinephrine reduces the blood flow to the gut and skin.

The Stress Response

One of the major functions of the adrenal gland is to respond to stress. Stress can be defined as some occurrence that disrupts homeostasis, and can be either physical or psychological or both. Physical stresses include exposing the body to injury, walking outside in cold and wet conditions without a coat on, or malnutrition. Psychological stresses include the perception of a physical threat, a fight with a loved one, or just a bad day at school.

The body responds in different ways to short-term stress and long-term stress following a pattern known as the **general adaptation syndrome**. The first stage of the general adaptation syndrome is called the **alarm reaction**. This "fight-or-flight" response, the result of a short-term stressor, is mediated by the hormones epinephrine and norepinephrine from the adrenal medulla, as stimulated by the sympathetic nervous system. Their function is to prepare the body for extreme physical exertion by increasing the heart rate, dilating the airways, and other related responses. Once this stressor is relieved, the body quickly returns to normal.

If the stressor is not soon relieved, the body attempts to adapt to the stressor in the second stage called the **stage of resistance**. The primary stress hormone at this stage is cortisol, secreted by the adrenal cortex as a result of signals sent by the hypothalamus and pituitary gland. Cortisol's effects are widespread. They include the maintenance of blood glucose through synthesizing glucose from compounds such as proteins (gluconeogenesis), and lipolysis so fatty acids can be used for energy by the body, thus preserving glucose. Additionally, the activity of the immune system and inflammation are reduced, as the resources of the body are directed towards dealing with the stress. The physiological adaptations during this stage allow the body to resist the most immediate negative effects of a longer-term stressor.

However, if the stressor continues for a longer term, the final stage of the stress response may occur. This is known as the **stage of exhaustion**. At this point, the resources of the body have become depleted and individuals may begin to suffer depression, severe fatigue, or even a fatal heart attack. The continued release of cortisol and other hormones associated with long term stress can cause damage to a variety of organ systems, and this condition has been linked to many diseases such as rheumatoid arthritis, hypertension and gastrointestinal diseases.

Disorders of Glucocorticoid (Cortisol) Secretion

Addison's disease can be caused by an autoimmune attack on the adrenal cortex, resulting in a hyposecretion of adrenal cortex hormones, including glucocorticoids. Since one of the actions of glucocorticoids is maintenance of blood glucose levels, hyposecretion of this hormone is characterized by hypoglycemia. This leads to muscular weakness, mental lethargy and weight loss.

Abnormal hypersecretion of glucocorticoids from the adrenal cortex can be directly caused by a tumor of the adrenal cortex, (or indirectly caused through an overstimulation from other areas of the body; for example, an

increased secretion of ACTH from the anterior pituitary gland). A hypersecretion of glucocorticoids due to an adrenal cortex tumour leads to **Cushing's syndrome**. Associated with this syndrome is a redistribution of body fats, so the legs become spindly, the face becomes moon-shaped, a buffalo hump appears on the back and the abdomen becomes pendulous. The cause of this redistribution of body fat is poorly understood but is thought to be at least partially due to a difference in the distribution of glucocorticoid receptors between the various body tissues.

Another sign of Cushing's syndrome is excess blood glucose, to be expected considering cortisol's effect of increasing blood glucose. Not so expected is the increase in blood sodium that can be observed with this disease (which does not cause an increase in the levels of aldosterone). Cortisol does in fact have a very minor mineralocorticoid effect, but this is negligible at normal levels of cortisol. However, the very large values for blood cortisol that result in Cushing's syndrome can exert a very noticeable effect, resulting in elevated levels of blood sodium.

The Endocrine Pancreas

The **pancreas** is a long, slender organ, most of which is located posterior to the bottom half of the stomach (Figure 14). Although it is primarily an exocrine gland, secreting a variety of digestive enzymes, the pancreas has an endocrine function. Its **pancreatic islets** – clusters of cells formerly known as the islets of Langerhans – secrete the hormones glucagon and insulin.

Cells and Secretions of the Pancreatic Islets

Cells residing in the pancreatic islets include the following two types of cells. The **alpha cell** produces the hormone glucagon and makes up approximately 20 percent of each islet. Glucagon plays an important role in blood glucose regulation; low blood glucose levels stimulate its release. The **beta cell** produces the hormone insulin and makes up approximately 75 percent of each islet. Elevated blood glucose levels stimulate the release of insulin.

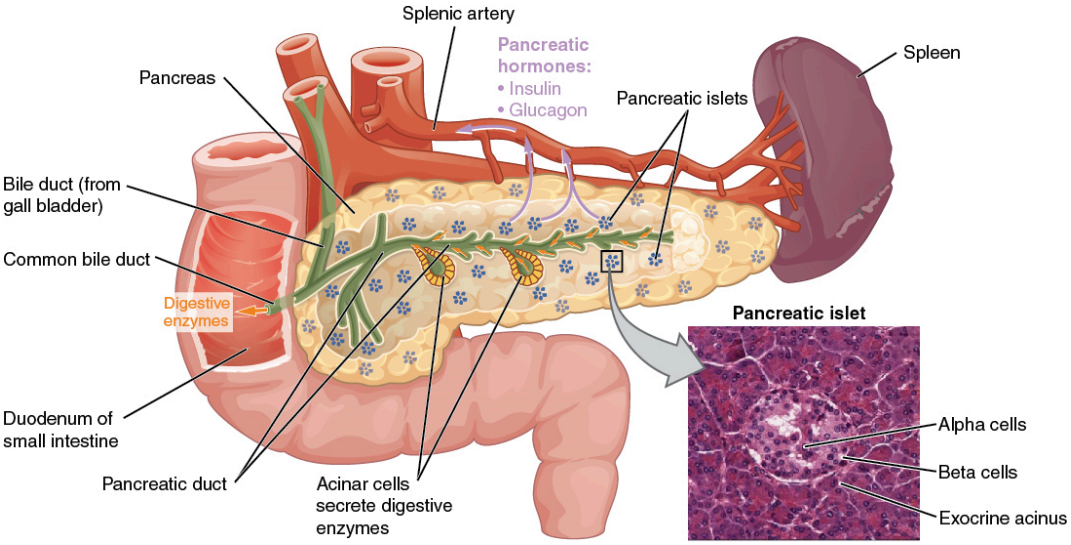


Figure 14. Pancreas. *The pancreatic exocrine function involves the acinar cells secreting digestive enzymes that are transported into the small intestine by the pancreatic duct. Its endocrine function involves the secretion of insulin (produced by beta cells) and glucagon (produced by alpha cells) within the pancreatic islets. These two hormones regulate the rate of glucose metabolism in the body. The micrograph reveals pancreatic islets. LM × 760. (Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

Regulation of Blood Glucose Levels by Insulin and Glucagon

Glucose is required for cellular respiration and is the preferred fuel for all body cells. The body derives glucose

from the breakdown of the carbohydrate-containing foods and drinks we consume. Glucose not immediately taken up by cells for fuel can be stored by the liver and muscles as glycogen or converted to triglycerides and stored in the adipose tissue. Hormones regulate both the storage and the utilization of glucose as required. Receptors located in the pancreas sense blood glucose levels, and subsequently the pancreatic cells secrete glucagon or insulin to maintain normal levels.

Glucagon

Receptors in the pancreas can sense the decline in blood glucose levels, such as during periods of fasting or during prolonged labor or exercise (Figure 15). In response, the alpha cells of the pancreas secrete the hormone **glucagon**, which has several effects:

- It stimulates the liver to perform **glycogenolysis**, the breaking down of glycogen into its component glucose building blocks. The resulting glucose is then released into the circulation for use by body cells.
- It stimulates gluconeogenesis in the liver, converting amino acids from body proteins into glucose.
- It stimulates lipolysis, the breakdown of stored triglycerides into free fatty acids and glycerol. Some of the free glycerol released into the bloodstream travels to the liver, which converts it into glucose. This is also a form of gluconeogenesis.

Taken together, these actions increase blood glucose levels. The activity of glucagon is regulated through a negative feedback mechanism; rising blood glucose levels inhibit further glucagon production and secretion.

Insulin

The primary function of **insulin** is to facilitate the uptake of glucose into body cells. Red blood cells, as well as cells of the brain, kidneys, and the lining of the small intestine, do not have insulin receptors on their cell membranes and do not require insulin for glucose uptake. Although all other body cells do require insulin if they are to take glucose from the bloodstream, skeletal muscle cells and adipose cells are the primary targets of insulin.

Insulin also reduces blood glucose levels by stimulating glycolysis, the metabolism of glucose for generation of ATP. Moreover, it stimulates the liver to perform **glycogenesis**, converting excess glucose into glycogen for storage, and it inhibits enzymes involved in glycogenolysis and gluconeogenesis. Finally, insulin promotes triglyceride and protein synthesis. The secretion of insulin is regulated through a negative feedback mechanism. As blood glucose levels decrease, further insulin release is inhibited.

Disorders of Insulin Secretion / Sensitivity

Dysfunction of insulin production and secretion, as well as the target cells' responsiveness to insulin, can lead to a condition called **diabetes mellitus,** traditionally characterized by the "three P's": polyphagia (excessive eating), polydipsia (excessive thirst) and polyuria (excessive urine production). Note that this is a different disease than diabetes insipidus, which is a decrease in the secretion of ADH, also causing increased urine production, but no glucose in the urine.

The hormonal control of blood glucose is more complex than just the interaction between insulin and glucagon. Along with glucagon, the following hormones are called "counterregulatory" hormones, because their effects on glucose levels are opposite to that of insulin, i.e., they all act to increase the level of glucose in the blood:

- **Epinephrine** stimulates the breakdown of glycogen in the liver and muscle (glycogenolysis)
- **Growth hormone** stimulates the mobilization and breakdown of fats, and decreases the uptake of glucose by fat cells
- **Cortisol** stimulates the breakdown of proteins and their use in the production of glucose (gluconeogenesis)

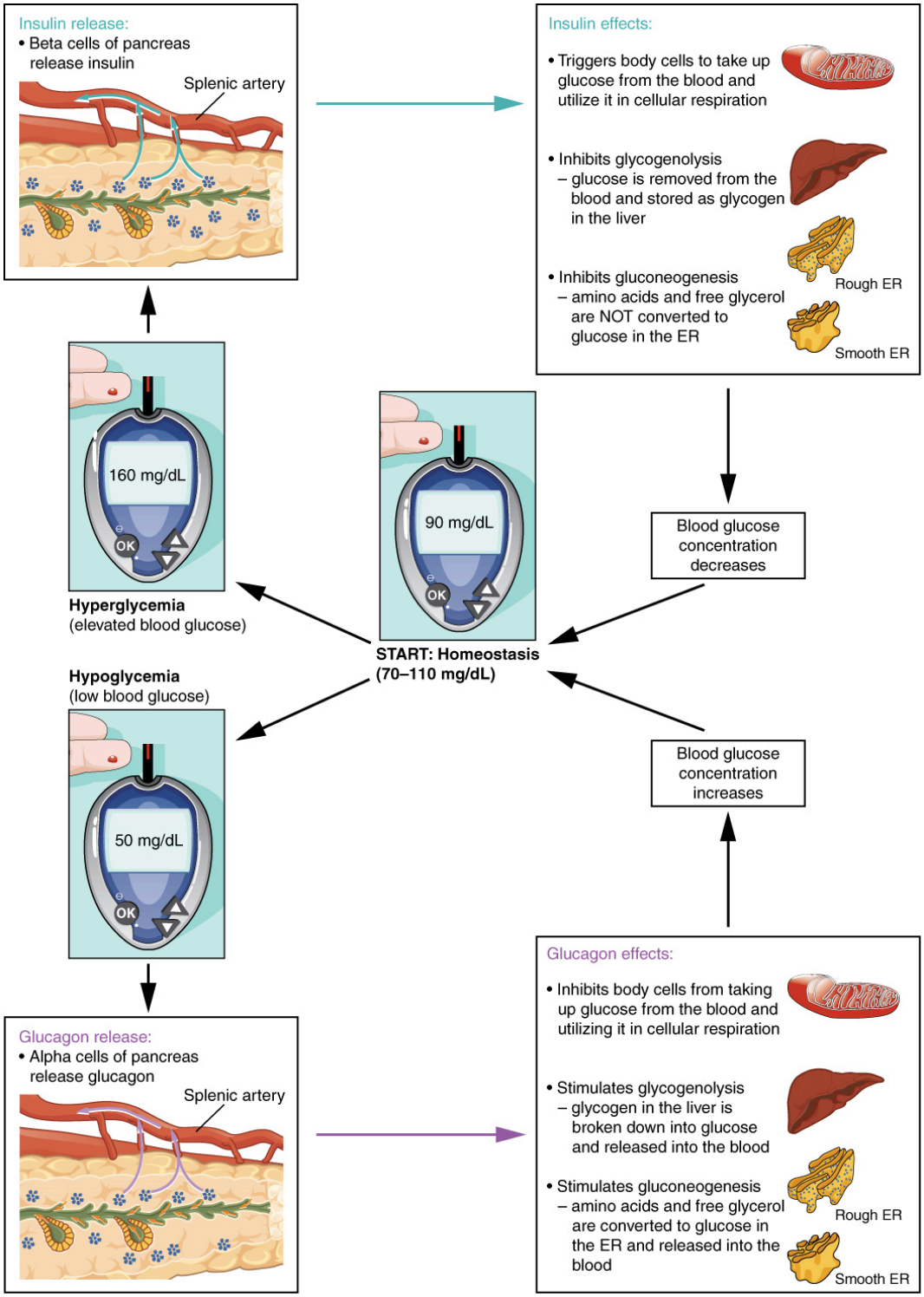


Figure 15. Homeostatic Regulation of Blood Glucose Levels. *Blood glucose concentration is tightly maintained between 70 mg/dL and 110 mg/dL. If blood glucose concentration rises above this range, insulin is released, which stimulates body cells to remove glucose from the blood. If blood glucose concentration drops below this range, glucagon is released, which stimulates body cells to release glucose into the blood.*

The endocrine functions of the ovaries and testes

The **ovaries** produce two hormones: estrogen and progesterone. **Estrogen** is produced by ovarian follicles. Estrogen stimulates the growth of both primary and secondary sex characteristics. Primary sex characteristics

in the female include the growth of the uterus and vagina, and the secondary characteristics include the development of body hair, enlarged breasts and a wider pelvis.

Progesterone is produced by the corpus luteum. It stimulates the development of milk-secretory tissue in the breasts, prepares the uterine lining for the implantation of a fertilized oocyte, and helps to maintain pregnancy.

The **testes** in the male produce the hormone **testosterone**. This hormone stimulates the development of primary sex characteristics, such as the accessory glands and the penis, and secondary sex characteristics such as body hair and a deepening of the voice.

The endocrine functions of the stomach and the duodenum

The principal function of the **stomach** is, of course, digestion. We are most familiar with the capacity of the stomach to store food, and its role in the mechanical and chemical digestion of food. However, there is also some endocrine tissue in the stomach. The secretion of hydrochloric acid in the stomach, and some enzymes, is under the control of a hormone called **gastrin**, which is produced by glandular tissue in the wall of the stomach. This hormone is produced by the stomach, travels through the bloodstream, and stimulates the exocrine tissue of the stomach.

The **duodenum** also has glandular tissue in its walls. One of the hormones it produces is called **secretin.** Secretin travels through the blood and stimulates the pancreas to produce pancreatic juice, which then enters the duodenum and aids in digestion.

The thymus and the pineal gland

The **thymus gland** was once thought to be a vestigial organ. However, it has been determined that it is a central gland of the lymphatic system, which is involved in the body's immune system. As part of this activity, the gland produces a hormone called **thymosin**, which is involved in the maturation and development of the immune system. This gland is larger in infants and decreases in size through adulthood. As a result of the change in size of the gland, the amounts of thymosin produced similarly decrease throughout adulthood.

The **pineal gland**, situated in the brain, produces the hormone **melatonin**. This hormone acts on the hypothalamus inhibiting the release of luteinizing hormone, thus affecting the activity of the gonads. The pineal gland is an interesting structure, because it is responsive to light and may be involved in the seasonal behaviour changes by some animals in response to changes in day length.

The special nature of the prostaglandins

Prostaglandins are lipids, much like some hormones. However, they are not produced in special organs or glands, but rather by many cell types from lipids in their own plasma membranes. They usually do not travel long distances within the body, but typically have effects on the tissue where they are produced.

The effects of prostaglandins are numerous, including the regulation of blood pressure, regulation of stomach secretions, stimulation and inhibition of uterine contractions, and the transmission of nerve impulses. At least 15 different prostaglandins have been discovered so far. They are vital for the normal functioning of the body. Prostaglandins were originally discovered in the secretions of the prostate gland, from which they got their name.

MAINTENANCE OF THE BODY

Unit 2: The Cardiovascular System

Unit outline

Blood

Part 1: An Overview of Blood

- Functions of blood
- Composition of blood
- Characteristics of blood
- Blood plasma

Part 2: Production of the Formed Elements

- Sites of hemopoiesis
- Differentiation of formed elements from stem cells
- Hemopoietic growth factors

Part 3: Erythrocytes

- Hemoglobin

Part 4: Leukocytes and Platelets

- Characteristics of leukocytes
- Platelets
- Disorders of platelets

Part 5: Hemostasis

- Vascular Spasm
- Formation of the platelet plug
- Coagulation
- Clotting factors involved in coagulation
- Extrinsic pathway
- Intrinsic pathway
- Common pathway
- Fibrinolysis
- Plasma anticoagulants
- Disorders of clotting

The Heart

Part 1: Heart Anatomy

- Location of the heart

- Chambers and circulation through the heart
- Membranes, surface features, and layers
- Internal structure of the heart
- Heart valve structure and function
- Coronary circulation

Part 2: Cardiac Muscle and Electrical Activity

- Conduction system of the heart
- Electrocardiogram

Part 3: Cardiac Cycle

- Pressure and flow
- Phases of the cardiac cycle
- Heart sounds
- Cardiac output

Part 4: Cardiac Physiology

- Heart rates
- Correlation between heart rates and cardiac output
- Cardiovascular centres
- Other factors influencing heart rate

Blood Vessels and Circulation

Part 1: Structure and function of blood vessels

- Shared structures
- Arteries
- Arterioles
- Capillaries
- Venules
- Veins

Part 2: Blood flow, blood pressure, and resistance

- Components of arterial blood pressure
- Pulse
- Measurement of blood pressure
- Variables affecting blood flow and blood pressure
- Cardiac output
- Compliance
- Blood volume
- Blood viscosity
- Vessel length and diameter

- Venous system
- Skeletal muscle pump
- Respiratory pump

Part 3: Capillary Exchange

Part 4: Hemostatic Regulation of the Vascular System

- Neural regulation
- The cardiovascular centres in the brain
- Baroreceptor reflexes
- Endocrine regulation
- Autoregulation of perfusion

Part 5: Circulatory Pathways

- Pulmonary circulation
- Overview of systemic arteries
- The aorta
- Coronary circulation
- Aortic arch branches
- Thoracic aorta and major branches
- Abdominal aorta and major branches
- Arteries serving the upper and lower limbs
- Overview of systemic veins
- The superior and inferior vena cavae
- Veins draining the lower limbs

Learning Objectives

At the end of this unit, you should be able to:

I. Describe the general nature and functions of blood, specify the main components of blood and describe the importance of each.

II. Describe the production of the formed elements of blood.

III. Describe the major factors that stimulate the body to produce more erythrocytes.

IV. Specify the types of leukocytes (white blood cells), their origins and relative quantities in normal blood.

V. Describe the procedure, what information is provided by, and the normal range for the following tests: hemoglobin (Hb), hematocrit (Hct).

VI. Describe the structure and function of platelets.

VII. Specify the two main components of blood that give blood its viscosity, and describe the importance of each to the blood.

VIII. Define hemostasis and describe the mechanisms involved in achieving hemostasis: vascular spasm, platelet plug formation, blood clotting.

IX. Describe the following disorders of hemostasis: thrombus, embolus, hemophilia.

X. Describe how the process of blood clotting is regulated, particularly with respect to prevention of blood clotting when it is not required, rapid initiation and progression of blood clotting when damage occurs, localization of blood clotting to the damaged region, and the dissolution of blood clots (fibrinolysis).

XI. Describe how each of the following affects blood clotting: vitamin K, anticoagulant drugs, thrombolytic agents.

XII. Describe the anatomy of the human heart with respect to the following: location, size, and shape.

XIII. Define and describe the location of the following: pericardium, epicardium, myocardium, endocardium.

XIV. Describe the anatomy and relationship to each other of the four chambers of the heart including the location and general makeup of all valves.

XV. Describe the double circulation and blood flow through the heart and explain the role of the four valves in controlling the direction of blood flow.

XVI. Briefly describe the major components of the coronary circulation and parts of the heart that they feed.

XVII. Specify the components of the conduction system of the heart and describe their functions in the normal conduction of an electrical impulse through the heart and explain the events which constitute and complete the heart beat (i.e. cardiac cycle).

XVIII. Describe the major components of the human electrocardiogram (ECG) and relate these to the electrical and mechanical events of the heart.

XIX. Describe the following major mechanisms that control heart rate: autonomic system, hormones, ionic composition of the blood, and body temperature.

XX. Define the terms systole and diastole in relation to contraction of the chambers of the heart

XXI. Describe relationships between the following components of the cardiovascular system and explain their functions: blood, artery, vein, capillary, atria, and ventricles.

XXII. Compare the structure and function of arteries, veins, and capillaries.

XXIII. Describe what is meant by blood pressure and specify the following: five factors which affect blood pressure, the major mechanisms that control blood pressure, and the average blood pressure of a young adult.

XXIV. Describe what is felt when a pulse is located, and specify four points where an arterial pulse may be felt.

XXV. Describe the following components of the cardiovascular system: the main arteries leaving the

heart, and those serving the trunk, appendages, and heart; the main veins entering the heart, and those draining the trunk, appendages, and heart.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Describe the general nature and functions of blood, specify the main components of blood and describe the importance of each.

1. What are the formed elements of blood?

II. Describe the production of the formed elements of blood.

1. Specify the origin and function of each of the formed elements.
2. The site(s) of production of formed elements.

III. Describe the major factors that stimulate the body to produce more erythrocytes.

1. For the hormone erythropoietin, state:
    - Its site of production and release.
    - The stimuli for its release.
    - Its physiological effects.
2. What would the effect on blood pressure be of increasing erythropoietin release or concentration? Explain your reasoning.

IV. Specify the types of leukocytes (white blood cells), their origins and relative quantities in normal blood.

V. Describe the procedure, what information is provided by, and the normal range for the following tests: hemoglobin (Hb), hematocrit (Hct).

1. Briefly explain how to determine the hematocrit of a blood sample, then explain:
    - What specific information does it give you about an individual's blood?
    - What can that information be used to determine?
2. Briefly explain how to determine the hemoglobin content of a blood sample, then explain:
    - What specific information does it give you about an individual's blood?
    - What can that information be used to determine?
3. Define "anemia".

VI. Describe the structure and function of platelets.

1. What cellular components do platelets possess, and which cellular components do platelets not possess?

VII. Specify the two main components of blood that give blood its viscosity, and describe the importance of each to the blood.

1. What factors contribute to the viscosity of blood?
2. Describe in general terms how each factor is normally regulated by the body.

VIII. Define hemostasis and describe the mechanisms involved in achieving hemostasis: vascular spasm, platelet plug formation, blood clotting.

1. Define “hemostasis” and describe why hemostasis is vital to maintaining homeostasis in the human body.
2. What is the specific chemical stimulus that causes the smooth muscle of blood vessels walls to contract when they are damaged, and what is the functional purpose of this contraction?
3. Describe in detail the formation of a platelet plug. Include in your description references to the specific stimulus that initially activates platelets, a definition of ‘platelet activation’, and a description of how activated platelets recruit additional platelets to a damaged site.
4. Compare and contrast the stimuli, events, and end result of the intrinsic and extrinsic pathways of blood clotting.
5. Is it possible to stimulate either the intrinsic or extrinsic pathway of blood clotting, without stimulating the other one? Explain your reasoning.
6. Draw a flow chart to describe in detail the intrinsic, extrinsic, and common pathways of blood clotting.

IX. Describe the following disorders of hemostasis: thrombus, embolus, hemophilia.

1. What is a ‘thrombus’? How is a thrombus produced, and what is the danger of thrombus production in the human body?
2. What is an ‘embolus’? How is an embolus produced, and what is the danger of embolus production in the human body?
3. Define ‘hemophilia’. What is the most common cause of hemophilia? Add an annotation to the detailed flow chart you drew of blood clotting to indicate this information.

X. Describe how the process of blood clotting is regulated, particularly with respect to prevention of blood clotting when it is not required, rapid initiation and progression of blood clotting when damage occurs, localization of blood clotting to the damaged region, and the dissolution of blood clots (fibrinolysis).

1. Describe the mechanisms in place that both allow rapid production of a blood clot when needed, and prevention of blood clot formation when there is no damage to a blood vessel.
2. How are blood clots normally dissolve when they are no longer needed?

XI. Describe how each of the following affects blood clotting: vitamin K, anticoagulant drugs, thrombolytic agents.

1. Describe the normal function of vitamin K in the human body. Refer to your detailed flow chart of blood clotting to explain the consequences of a vitamin K deficiency.

2. List two examples of anticoagulant drugs and two examples of thrombolytic agents. For each, briefly describe their mechanism of action and contrast their effects on blood clots and/or blood clot formation.

XII. Describe the anatomy of the human heart with respect to the following: location, size, and shape.

1. Use correct anatomical terms and complete sentences to describe the position of the human heart in relation to the lungs, diaphragm, vertebral column and thoracic cavity.

XIII. Define and describe the location of the following: pericardium, epicardium, myocardium, endocardium.

1. Draw a simple diagram of the heart wall showing all the following structures and on your diagram, wherever possible identify the *specific tissue type* each layer is composed of:
    - Pericardium
    - Epicardium
    - Myocardium
    - Endocardium

XIV. Describe the anatomy and relationship to each other of the four chambers of the heart including the location and general makeup of all valves.

1. Distinguish between the upper and lower chambers of the heart
2. Describe the partitioning of the heart into left and right chambers.
3. Name and describe the structure and location of the valves of the heart

XV. Describe the double circulation and blood flow through the heart and explain the role of the four valves in controlling the direction of blood flow.

1. State the names of the two parts of the circulation and their general function in terms of where blood flow is conducted.
2. Clearly state the function of each of the four valves found in the human heart. Your description of their function should make reference to the specific location where blood is moving from and to as it passes through each valve.
3. Describe the location and function of each of the two main arteries that carry blood out of the heart, and the main veins that carry blood into the human heart.
4. List, in order, **all** the structures through which blood passes as it moves through the heart until it exits, starting from:
    - Its entrance into the heart from the venae cavae and coronary sinus
    - Its entrance into the heart from the pulmonary veins
5. Draw a simplified diagram of the human heart showing the vessels connected directly to its chambers. Show and label **all** of the following components:
    - Pericardium
    - Epicardium
    - Myocardium

- Endocardium
- Right atrium
- Left atrium
- Right ventricle
- Left ventricle
- Tricuspid valve
- Bicuspid valve
- Coronary sinus
- Aorta
- Superior vena cava
- Inferior vena cava
- Chordae tendinae
- Aortic semilunar valve
- Pulmonary semilunar valve
- Pulmonary trunk
- Right pulmonary arteries
- Left pulmonary arteries
- Right pulmonary veins
- Left pulmonary veins

XVI. Briefly describe the major components of the coronary circulation and parts of the heart that they feed.

1. How are the cardiac muscle fibers of the heart supplied with nutrients?
2. How is waste removed from the cardiac muscle fibers of the heart?

XVII. Specify the components of the conduction system of the heart and describe their functions in the normal conduction of an electrical impulse through the heart and explain the events which constitute and complete the heart beat (i.e. cardiac cycle).

1. Describe, briefly, the general properties of the sinoatrial node (SA Node), atrioventricular node (AV Node), atrioventricular bundle (the bundle of His), right and left bundle branches, and the Purkinje fibers, and the role of each of these in the conduction of a cardiac impulse.
2. Describe one heartbeat in detail. Include in your description all the events of the conduction system and the heart muscle, and all the structures the blood passes through (in order!) as it moves through the heart.
3. Describe what would happen if the sequence described above is not followed.

XVIII. Describe the major components of the human electrocardiogram (ECG) and relate these to the electrical and mechanical events of the heart.

1. Draw and label the major components (P wave, QRS wave, T wave, and the appropriate gaps between them) of a normal ECG tracing. Then describe:
    - The electrical events in the heart that underlie each wave.
    - The relative times where the ‘lub’ and ‘dub’ sounds of a heartbeat could be heard.
    - The physical events in the heart that underlie the ‘lub’ and ‘dub’ sounds of a heartbeat.

XIX. Describe the following major mechanisms that control heart rate: autonomic system, hormones, ionic composition of the blood, and body temperature.

1. Describe the influence of proprioceptors, chemoreceptors, and baroreceptors on the cardiovascular centres in the medulla.
2. Describe how hormones modify heart rate.
3. Describe how the ionic composition of the blood influences heart rate.
4. Describe how body temperature affects heart rate.

XX. Define the terms systole and diastole in relation to contraction of the chambers of the heart

XXI. Describe relationships between the following components of the cardiovascular system and explain their functions: blood, artery, vein, capillary, atria, and ventricles.

XXII. Compare the structure and function of arteries, veins, and capillaries.

XXIII. Describe what is meant by blood pressure and specify the following: five factors which affect blood pressure, the major mechanisms that control blood pressure, and the average blood pressure of a young adult.

1. Define the term “blood pressure”.
2. Describe how blood pressure is measured, and what is considered a “normal” blood pressure.
3. Define cardiac output and describe how each of the following physiological factors affect blood pressure:
    - Heart rate
    - Contractility (strength of contraction) of the heart
    - Blood volume
    - Peripheral resistance
    - Blood viscosity
4. Describe how blood pressure is regulated by:
    - The nervous system
    - The endocrine system
    - Autoregulation

XXIV. Describe what is felt when a pulse is located, and specify four points where an arterial pulse may be felt.

1. When you manually “take someone’s pulse”, what is causing the pulsing pressure waves you feel?
2. List four locations on the human body where a pulse can be taken manually and explain why an arterial pulse can be felt at specific locations rather than just anywhere on the human body.

XXV. Describe the following components of the cardiovascular system: the main arteries leaving the heart, and those serving the trunk, appendages, and heart; the main veins entering the heart, and those draining the trunk, appendages, and heart.

1. Draw a flow chart showing the components of the cardiovascular system. Start with the three main components (heart, blood vessels, and blood), and continue by specifying all the constituent parts of each.

2. Compare and contrast (clearly!) the anatomical structure and function of arteries, veins, and blood capillaries.
3. Draw a simple diagram of the human cardiovascular system that shows both circuits, indicating the vessels blood is moved through as it is passed to and from the head, arms, organs of the abdomen, and lungs. Your diagram should include:
    - The main arteries leaving the heart
    - The main arteries serving the trunk, appendages and the heart
    - The main veins entering the heart
    - The main veins draining the trunk, appendages and the heart

Blood

Single-celled organisms do not need blood. They obtain nutrients directly from and excrete wastes directly into their environment. The human organism cannot do that. Our large, complex bodies need blood to deliver nutrients to and remove wastes from our trillions of cells. The heart pumps blood throughout the body in a network of blood vessels. Together, these three components—blood, heart, and vessels—makes up the cardiovascular system.

Part 1: An Overview of Blood

Recall that **blood** is a connective tissue. Like all connective tissues, it is made up of cellular elements and an extracellular matrix. The cellular elements—referred to as the **formed elements**—include **erythrocytes** (red blood cells, or RBCs), **leukocytes** (white blood cells, or WBCs), and cell fragments called **platelets**. The extracellular matrix, called **plasma**, makes blood unique among connective tissues because it is fluid. This fluid, which is mostly water, perpetually suspends the formed elements and enables them to circulate throughout the body within the cardiovascular system.

Functions of Blood: The primary function of blood is to deliver oxygen and nutrients to and remove wastes from body cells, but that is only the beginning of the story. The specific functions of blood also include defense and maintenance of homeostasis.

Transportation: Nutrients from the foods you eat are absorbed in the digestive tract. Most of these travel in the bloodstream directly to the liver, where they are processed and released back into the bloodstream for delivery to body cells. Oxygen from the air you breathe diffuses into the blood, which moves from the lungs to the heart, which then pumps it out to the rest of the body. Moreover, endocrine glands scattered throughout the body release their products, called hormones, into the bloodstream, which carries them to distant target cells. Blood also picks up cellular wastes and by products, and transports them to various organs for removal. For instance, blood moves carbon dioxide to the lungs for exhalation from the body, and various waste products are transported to the kidneys and liver for excretion from the body in the form of urine or bile.

Defense: Many types of leukocytes protect the body from external threats, such as disease-causing bacteria that have entered the bloodstream in a wound. Other leukocytes seek out and destroy internal threats, such as cells with mutated DNA that could multiply to become cancerous, or body cells infected with viruses.

When damage to the vessels results in bleeding, blood platelets and certain proteins dissolved in the plasma, the fluid portion of the blood, interact to block the ruptured areas of the blood vessels involved. This protects the body from further blood loss.

Maintenance of Homeostasis: Recall that body temperature is regulated via a classic negative-feedback loop. If you were exercising on a warm day, your rising core body temperature would trigger several homeostatic mechanisms, including increased transport of blood from your core to your body periphery, which is typically

cooler. As blood passes through the vessels of the skin, heat would be dissipated to the environment, and the blood returning to your body core would be cooler. In contrast, on a cold day, blood is diverted away from the skin to maintain a warmer body core. In extreme cases, this may result in frostbite.

Blood also helps to maintain the chemical balance of the body. Proteins and other compounds in blood act as buffers, which thereby help to regulate the pH of body tissues. Blood also helps to regulate the water content of body cells.

Composition of Blood: You have probably had blood drawn from a superficial vein in your arm, which was then sent to a lab for analysis. Some of the most common blood tests—for instance, those measuring lipid or glucose levels in plasma—determine which substances are present within blood and in what quantities. Other blood tests check for the composition of the blood itself, including the quantities and types of formed elements.

One such test, called a **hematocrit**, measures the percentage of red blood cells, clinically known as erythrocytes, in a blood sample. It is performed by spinning the blood sample in a specialized centrifuge, a process that causes the heavier elements suspended within the blood sample to separate from the lightweight, liquid plasma (Figure 1). Because the heaviest elements in blood are the erythrocytes, these settle at the bottom of the hematocrit tube. Located above the erythrocytes is a pale, thin layer composed of the remaining formed elements of blood.

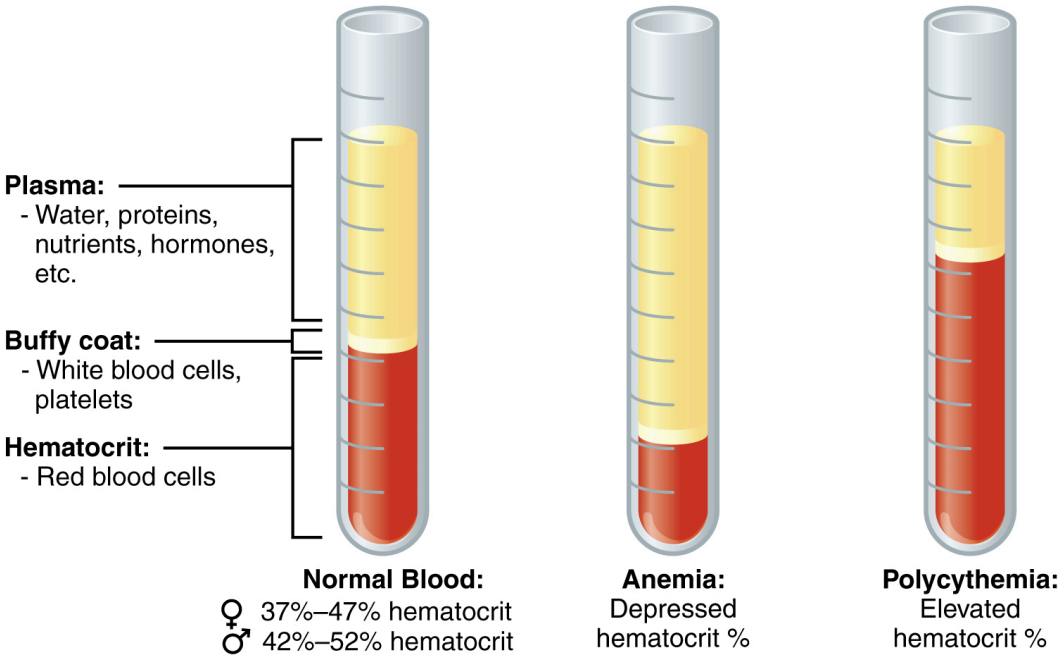


Figure 1. Composition of Blood. *The cellular elements of blood include a vast number of erythrocytes and comparatively fewer leukocytes and platelets. Plasma is the fluid in which the formed elements are suspended. A sample of blood spun in a centrifuge reveals that plasma is the lightest component. It floats at the top of the tube separated from the heaviest elements, the erythrocytes, by a buffy coat of leukocytes and platelets. Hematocrit is the percentage of the total sample that is comprised of erythrocytes. Depressed and elevated hematocrit levels are shown for comparison.*

This pale, thin layer of centrifuged blood sample consists of the white blood cells, clinically known as leukocytes, and the platelets, cell fragments also called thrombocytes. This layer is referred to as the **buffy coat** because of its colour; it normally constitutes less than 1% of a blood sample. Above the buffy coat is the blood plasma, normally a pale, straw-coloured fluid, which constitutes the remainder of the sample.

The volume of erythrocytes after centrifugation is also commonly referred to as **packed cell volume (PCV)**. In normal blood, about 45% of a sample is erythrocytes. The hematocrit of any one sample can vary significantly, and may be 36-50%, depending on sex and other factors. Normal hematocrit values for females range from 37 to 47%, with a mean value of 41%; for males, hematocrit ranges from 42 to 52%, with a mean of 47%. The

percentage of other formed elements, the leukocytes and platelets, is extremely small so it is not normally considered with the hematocrit. The **mean plasma percentage** is the percent of blood that is *not* erythrocytes: for females, it is approximately 59% (or 100 minus 41), and for males, it is approximately 53% (or 100 minus 47).

Characteristics of Blood: When you think about blood, the first characteristic that probably comes to mind is its colour. Blood that has just taken up oxygen in the lungs is bright red, and blood that has released oxygen in the tissues is a more dusky red. This is because hemoglobin is a pigment that changes colour, depending upon the degree of oxygen saturation.

Blood is viscous and somewhat sticky to the touch. It has a viscosity approximately five times greater than water. Viscosity is a measure of a fluid's thickness or resistance to flow, and is influenced by plasma proteins and formed elements (usually albumin concentration and the number of erythrocytes) within the blood. The viscosity of blood has a dramatic impact on blood pressure and flow. Consider the difference in flow between water and honey. The more viscous honey would demonstrate a greater resistance to flow than the less viscous water. The same principle applies to blood.

The normal temperature of blood is slightly higher than normal body temperature—about 38 °C (or 100.4 F), compared to 37 °C (or 98.6 F) for an internal body temperature reading, although daily variations of 0.5 °C are normal. Although the surface of blood vessels is relatively smooth, as blood flows through them, it experiences some friction and resistance, especially as vessels age and lose their elasticity, thereby producing heat. This accounts for its slightly higher temperature.

The pH of blood averages about 7.4, but can range from 7.35 to 7.45 in a healthy person. Blood is therefore somewhat more basic (alkaline) on a chemical scale than pure water, which has a pH of 7.0. Blood contains numerous buffers that help to regulate pH.

Blood constitutes approximately 8% of adult body weight. Adult males typically average about 5-6 liters of blood; adult females average 4-5 liters.

Blood Plasma: Like other fluids in the body, plasma is composed primarily of water, and is about 92% water. Dissolved or suspended within this water is a mixture of substances, most of which are proteins. There are literally hundreds of substances dissolved or suspended in the plasma, although many of them are found only in very small quantities.

Plasma Proteins: About 7% of the volume of plasma – nearly all that is not water – is made of proteins. These include several plasma proteins (proteins that are unique to the plasma), plus a much smaller number of regulatory proteins, including enzymes and some hormones (Table 1).

- **Albumin** is the most abundant of the plasma proteins. Manufactured by the liver, albumin molecules serve as binding proteins—transport vehicles for fatty acids and steroid hormones. Recall that lipids are hydrophobic; however, their binding to albumin enables their transport in the watery plasma. Albumin is also the most significant contributor to the osmotic pressure of blood; that is, its presence holds water inside the blood vessels and draws water from the tissues, across blood vessel walls, and into the bloodstream. This in turn helps to maintain both blood volume and blood pressure. Albumin normally accounts for approximately 54% of the total plasma protein content, in clinical levels of 3.5–5.0 g/dL blood.
- The second most common plasma proteins are the **globulins**. A heterogeneous group, there are three main subgroups known as alpha, beta, and gamma globulins. The alpha and beta globulins transport iron, lipids, and the fat-soluble vitamins A, D, E, and K to the cells; like albumin, they also contribute to osmotic pressure. The gamma globulins are proteins involved in immunity and are better known as antibodies or immunoglobulins. Although other plasma proteins are produced by the liver, immunoglobulins are produced by specialized leukocytes known as plasma cells. Globulins make up approximately 38% of the total plasma protein volume, in clinical levels of 1.0–1.5 g/dL blood.
- The least abundant plasma protein is **fibrinogen**. Like albumin and the alpha and beta globulins, fibrinogen is produced by the liver. It is essential for blood clotting, a process described later in this chapter. Fibrinogen accounts for about 7% of the total plasma protein volume, in clinical levels of 0.2–0.45 g/dL blood.

Other Plasma Solutes: In addition to proteins, plasma contains a wide variety of other substances. These include various **electrolytes**, such as sodium, potassium, and calcium ions; **dissolved gases,** such as oxygen, carbon dioxide, and nitrogen; various **organic nutrients,** such as vitamins, lipids, glucose, and amino acids; and **metabolic wastes.** All of these non-protein solutes combined contribute approximately 1% to the total volume of plasma.

Check out this CrashCourse video to learn more about the components of blood! Direct link: https://youtu.be/HQWIcSp9Sls

Table 1. Major blood components

Component and % of blood	Subcomponent and % of component	Type and % (where appropriate)	Site of production	Major function(s)
Plasma 46-63%	Water 92%	Fluid	Absorbed by intestinal tract or produced by metabolism	Transport medium
	Plasma proteins 7%	Albumin 54-60%	Liver	Maintain osmotic concentration, transport lipid molecules
		Globulins 35-38%	Alpha globulins: liver	Transport, maintain osmotic concentration
			Beta globulins: liver	Transport, maintain osmotic concentration
			Gamma globulins (immunoglobulins): plasma cells	Immune responses
		Fibrinogen 4-7%	Liver	Blood clotting in hemostasis
	Regulatory proteins <1%	Hormones and enzymes	Various sources	Regulate various body functions
	Other solutes 1%	Nutrients, gases, and wastes	Absorbed by intestinal tract, exchanged in respiratory system, or produced by cells	Numerous and varied
Formed elements 37-54%	Erythrocytes 99%	Erythrocytes	Red bone marrow	Transport gases (primarily O_2, some CO_2)
	Leukocytes <1%	Granular leukocytes: neutrophils, eosinophils, basophils	Red bone marrow	Nonspecific immunity
		Agranular leukocytes: lymphocytes, monocytes	Lymphocytes: red bone marrow and lymphatic tissue	Lymphocytes: specific immunity
			Monocytes: red bone marrow	Monocytes: nonspecific immunity
	Platelets <1%		Megakaryocytes in red bone marrow	Hemostasis

Part 2: Production of the Formed Elements

The lifespan of the formed elements is very brief. Although one type of leukocyte called memory cells can survive for years, most erythrocytes, leukocytes, and platelets normally live only a few hours to a few weeks. Thus,

the body must form new blood cells and platelets quickly and continuously. When you donate a unit of blood during a blood drive (approximately 475 mL, or about 1 pint), your body typically replaces the donated plasma within 24 hours, but it takes about 4 to 6 weeks to replace the blood cells. This restricts the frequency with which donors can contribute their blood. The process by which this replacement occurs is called **hemopoiesis**, or **hematopoiesis** (from the Greek root haima- = "blood"; -poiesis = "production").

Sites of Hemopoiesis: Prior to birth, hemopoiesis occurs in a number of tissues, beginning with the yolk sac of the developing embryo, and continuing in the foetal liver, spleen, lymphatic tissue, and eventually the red bone marrow. Following birth, most hemopoiesis occurs in the red marrow, a connective tissue within the spaces of spongy (cancellous) bone tissue. In children, hemopoiesis can occur in the medullary cavity of long bones; in adults, the process is largely restricted to the cranial and pelvic bones, the vertebrae, the sternum, and the proximal epiphyses of the femur and humerus.

Differentiation of Formed Elements from Stem Cells: All formed elements arise from stem cells of the red bone marrow. Recall that stem cells undergo mitosis plus cytokinesis (cellular division) to give rise to new daughter cells: One of these remains a stem cell and the other differentiates into one of any number of diverse cell types (Figure 2).

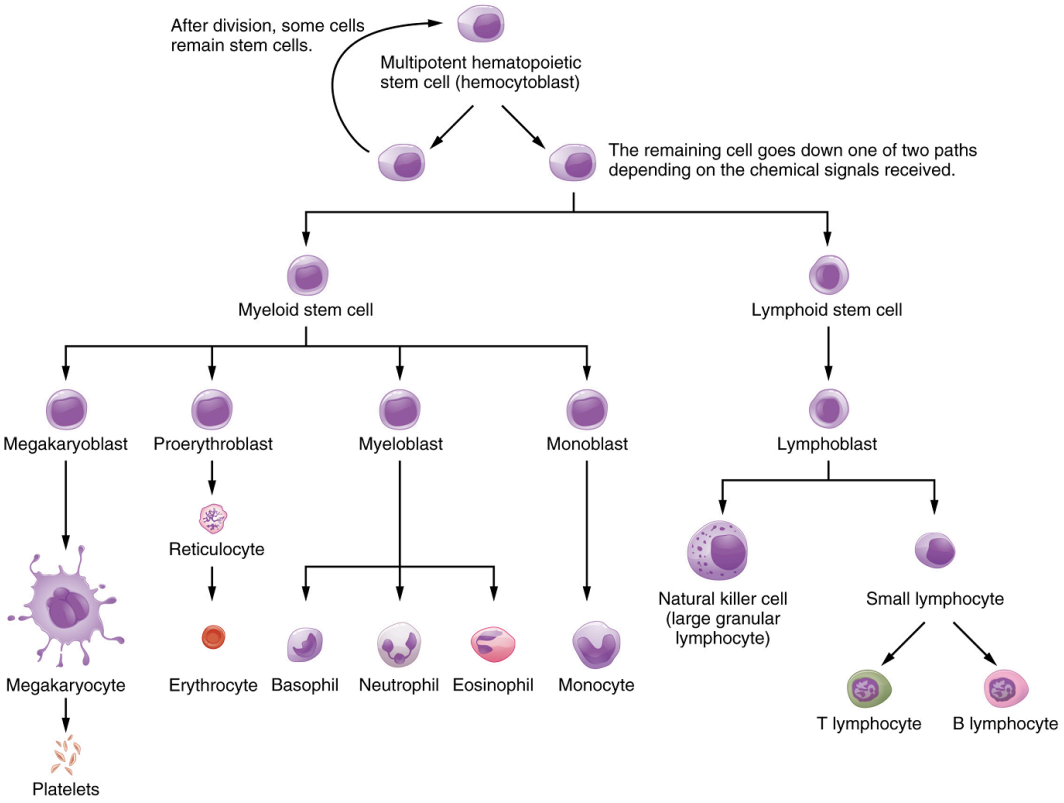


Figure 2. Hematopoietic System of Bone Marrow. *Hemopoiesis is the proliferation and differentiation of the formed elements of blood.*

Hemopoietic Growth Factors: Development from stem cells to precursor cells to mature cells is again initiated by hemopoietic growth factors. The growth factor responsible for the production of erythrocytes is **erythropoietin (EPO).** Erythropoietin is a hormone secreted by the kidneys in response to low oxygen levels. Some athletes use synthetic EPO as a performance-enhancing drug (called blood doping) to increase RBC counts and subsequently increase oxygen delivery to tissues throughout the body. EPO is a banned substance in most organized sports, but it is also used medically in the treatment of certain anemia, specifically those triggered by certain types of cancer, and other disorders in which increased erythrocyte counts and oxygen levels are desirable.

Part 3: Erythrocytes

The **erythrocyte**, commonly known as a red blood cell (or RBC), is by far the most common formed element: A single drop of blood contains millions of erythrocytes and just thousands of leukocytes. Specifically, males have about 5.4 million erythrocytes per microliter (μL) of blood, and females have approximately 4.8 million per μL. In fact, erythrocytes are estimated to make up about 25% of all cells in the body. As you can imagine, they are quite small cells, with a mean diameter of only about 7–8 micrometers (μm) (Table 2). The primary functions of erythrocytes are to pick up inhaled oxygen from the lungs and transport it to the body's tissues, and to pick up some (about 24%) of the carbon dioxide waste produced at the tissues and transport it to the lungs for exhalation. Erythrocytes remain within the vascular network. Although leukocytes typically leave the blood vessels to perform their defensive functions, movement of erythrocytes from the blood vessels is abnormal. Their unique structure enables them to change their shape to squeeze through capillaries.

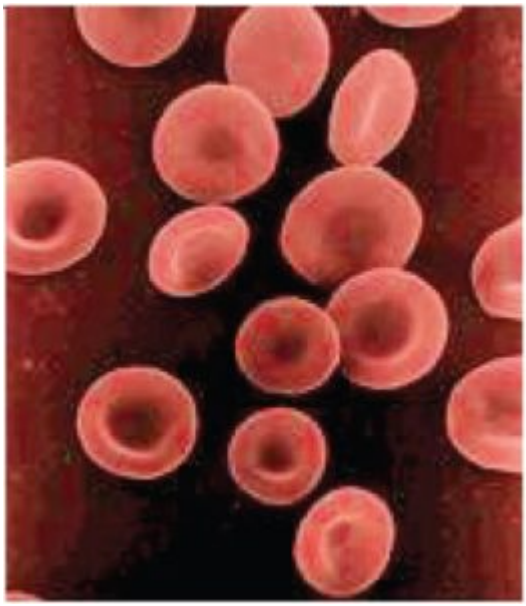


Figure 3. Shape of Red Blood Cells. *Erythrocytes are biconcave discs with shallow centres. This shape optimizes the ratio of surface area to volume, facilitating gas exchange. It also enables them to fold up as they move through narrow blood vessels.*

Erythrocytes are biconcave disks; that is, they are plump at their periphery and very thin in the centre (Figure 3). Since they lack most organelles, there is more interior space for the presence of the hemoglobin molecules that transport gases. The biconcave shape also provides a greater surface area across which gas exchange can occur, relative to its volume; a sphere of a similar diameter would have a lower surface area-to-volume ratio. In the capillaries, the oxygen carried by the erythrocytes can diffuse into the plasma and then through the capillary walls to reach the cells, whereas some of the carbon dioxide produced by the cells as a waste product diffuses into the capillaries to be picked up by the erythrocytes. Capillary beds are extremely narrow, slowing the passage of the erythrocytes and providing an extended opportunity for gas exchange to occur. However, the space within capillaries can be so minute that, despite their own small size, erythrocytes may have to fold in on themselves if they are to make their way through.

Formed element	Major subtypes	Numbers present per microliter (μL) and mean (range)	Appearance in a standard blood smear	Summary of functions	Comments
Erythrocytes (red blood cells)		5.2 million (4.4–6.0 million)	Flattened biconcave disk; no nucleus; pale red color	Transport oxygen and some carbon dioxide between tissues and lungs	Lifespan of approximately 120 days
Leukocytes (white blood cells)		7000 (5000–10,000)	Obvious dark-staining nucleus	All function in body defenses	Exit capillaries and move into tissues; lifespan of usually a few hours or days
	Granulocytes including neutrophils, eosinophils, and basophils	4360 (1800–9950)	Abundant granules in cytoplasm; nucleus normally lobed	Nonspecific (innate) resistance to disease	Classified according to membrane-bound granules in cytoplasm
	Neutrophils 	4150 (1800–7300)	Nuclear lobes increase with age; pale lilac granules	Phagocytic; particularly effective against bacteria. Release cytotoxic chemicals from granules	Most common leukocyte; lifespan of minutes to days
	Eosinophils 	165 (0–700)	Nucleus generally two-lobed; bright red-orange granules	Phagocytic cells; particularly effective with antigen- antibody complexes. Release antihistamines. Increase in allergies and parasitic infections	Lifespan of minutes to days
	Basophils 	44 (0–150)	Nucleus generally two-lobed but difficult to see due to presence of heavy, dense, dark purple granules	Promotes inflammation	Least common leukocyte; lifespan unknown
	Agranulocytes including lymphocytes and monocytes	2640 (1700–4950)	Lack abundant granules in cytoplasm; have a simple-shaped nucleus that may be indented	Body defenses	Group consists of two major cell types from different lineages
	Lymphocytes 	2185 (1500–4000)	Spherical cells with a single often large nucleus occupying much of the cell's volume; stains purple; seen in large (natural killer cells) and small (B and T cells) variants	Primarily specific (adaptive) immunity: T cells directly attack other cells (cellular immunity); B cells release antibodies (humoral immunity); natural killer cells are similar to T cells but nonspecific	Initial cells originate in bone marrow, but secondary production occurs in lymphatic tissue; several distinct subtypes; memory cells form after exposure to a pathogen and rapidly increase responses to subsequent exposure; lifespan of many years
	Monocytes 	455 (200–950)	Largest leukocyte with an indented or horseshoe-shaped nucleus	Very effective phagocytic cells engulfing pathogens or worn out cells; also serve as antigen-presenting cells (APCs) for other components of the immune system	Produced in red bone marrow; referred to as macrophages after leaving circulation
Platelets 		350,000 (150,000–500,000)	Cellular fragments surrounded by a plasma membrane and containing granules; purple stain	Hemostasis plus release growth factors for repair and healing of tissue	Formed from megakaryocytes that remain in the red bone marrow and shed platelets into circulation

Table 2. Summary of Formed Elements in Blood

Hemoglobin: Hemoglobin is a large molecule made up of proteins and iron. It consists of four folded chains of a protein called globin, designated alpha 1 and 2, and beta 1 and 2 (Figure 4a). Each of these globin molecules is bound to a red pigment molecule called heme, which contains an ion of iron (Fe^{2+}) (Figure 4b).

Each iron ion in the heme can bind to one oxygen molecule; therefore, each hemoglobin molecule can transport four oxygen molecules. An individual erythrocyte may contain about 300 million hemoglobin molecules, and therefore can bind to and transport up to 1.2 billion oxygen molecules. These oxygen molecules come from the air we breathe; they diffuse across the respiratory membrane in the lungs, then into erythrocytes where they can bind to hemoglobin and be carried back to the heart and then to the rest of the body.

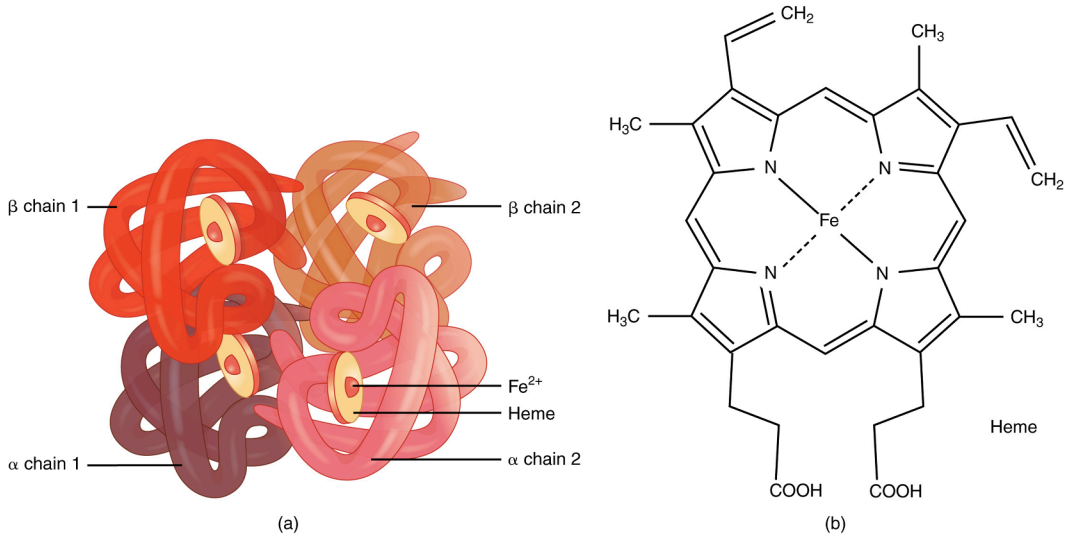


***Figure 4. Hemoglobin**. (a) A molecule of hemoglobin contains four globin proteins, each of which is bound to one molecule of the iron-containing pigment heme. A single erythrocyte can contain 300 million hemoglobin molecules, and thus more than 1 billion oxygen molecules. (b) The chemical structure of one heme group includes an iron atom to which an oxygen molecule can bind.*

Carbon dioxide enters the bloodstream at the tissue level, and among other transport mechanisms can bind to one end of a subunit of hemoglobin. From the capillaries, the carbon dioxide is carried back to the lungs, where it is released.

Changes in the levels of erythrocytes can have significant effects on the body's ability to effectively deliver oxygen to the tissues. Ineffective hematopoiesis results in insufficient numbers of erythrocytes and results in one of several forms of **anemia**. An overproduction of erythrocytes produces a condition called **polycythemia**. The primary drawback with polycythemia is not a failure to directly deliver enough oxygen to the tissues, but rather the increased viscosity of the blood, which makes it more difficult for the heart to circulate the blood.

In patients with insufficient hemoglobin, the tissues may not receive sufficient oxygen, resulting in another form of anemia.

In contrast to anemia, an elevated erythrocyte count is called **polycythemia** and is detected in a patient's elevated hematocrit. It can occur transiently in a person who is dehydrated; when water intake is inadequate or water losses are excessive, the plasma volume falls. As a result, the hematocrit rises. A mild form of polycythemia is chronic but normal in people living at high altitudes; the decreased oxygen availability at high altitudes released in erythropoietin release (discussed earlier in this chapter), resulting in increased erythrocyte production. Some elite athletes train at high elevations specifically to induce this phenomenon. Finally, a type of bone marrow disease called polycythemia vera (from the Greek vera = "true") causes an excessive production of immature erythrocytes. Polycythemia vera can dangerously elevate the viscosity of blood, raising blood pressure and making it more difficult for the heart to pump blood throughout the body. It is a relatively rare disease that occurs more often in men than women and is more likely to be present in elderly patients those over 60 years of age.

Part 4: Leukocytes and Platelets

The leukocyte, commonly known as a white blood cell (or WBC), is a major component of the body's defenses against disease. Leukocytes protect the body against invading microorganisms and body cells with mutated

DNA, and they clean up debris. Platelets are essential for the repair of blood vessels when damage to them has occurred; they also provide growth factors for healing and repair.

Characteristics of Leukocytes: Although leukocytes and erythrocytes both originate from hematopoietic stem cells in the bone marrow, they differ from each other in many significant ways. The types of leukocytes will be discussed in a succeeding unit (Immunity).

Platelets: You may occasionally see platelets referred to as thrombocytes, but because this name suggests they are a type of cell, it is not accurate. A platelet is not a cell but rather a fragment of the cytoplasm of a cell called a **megakaryocyte** that is surrounded by a plasma membrane. Megakaryocytes are descended from myeloid stem cells and are large, typically 50–100 µm in diameter, and contain an enlarged, lobed nucleus. Thrombopoietin, a glycoprotein secreted by the kidneys and liver, stimulates the proliferation of megakaryoblasts, which mature into megakaryocytes. These remain within bone marrow tissue (Figure 5) and ultimately form platelet-precursor extensions that extend through the walls of bone marrow capillaries to release into the circulation thousands of cytoplasmic fragments, each enclosed by a bit of plasma membrane. These enclosed fragments are platelets. Each megakarocyte releases 2000–3000 platelets during its lifespan. Following platelet release, megakaryocyte remnants, which are little more than a cell nucleus, are consumed by macrophages (macrophages are discussed further in the Immunity chapter).

Platelets are relatively small, 2–4 µm in diameter, but numerous, with typically 150,000–160,000 per µL of blood. After entering the circulation, approximately one-third migrate to the spleen for storage for later release in response to any rupture in a blood vessel. They then become activated to perform their primary function, which is to limit blood loss. Platelets remain only about 10 days, then are phagocytized by macrophages found in the spleen and liver. Platelets are critical to hemostasis, the stoppage of blood flow following damage to a vessel. They also secrete a variety of growth factors essential for growth and repair of tissue, particularly connective tissue. Infusions of concentrated platelets are now being used in some therapies to stimulate healing.

Disorders of Platelets: Thrombocytosis is a condition in which there are too many platelets. This may trigger formation of unwanted blood clots (thrombosis), a potentially fatal disorder. If there is an insufficient number of platelets, called **thrombocytopenia**, blood may not clot properly, and excessive bleeding may result.

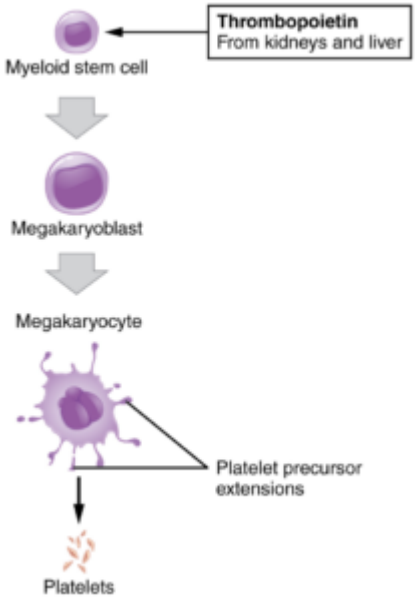


Figure 5. Platelets. *Platelets are derived from cells called megakaryocytes.*

Part 5: Hemostasis

Platelets are key players in **hemostasis**, the process by which the body seals a ruptured blood vessel and

prevents further loss of blood. Although rupture of larger vessels usually requires medical intervention, hemostasis is quite effective in dealing with small, simple wounds. There are three steps to the process: vascular spasm, the formation of a platelet plug, and coagulation (blood clotting) (Figure 6). Failure of any of these steps will result in hemorrhage – excessive bleeding.

Vascular Spasm: When a vessel is severed or punctured, or when the wall of a vessel is damaged, vascular spasm occurs. In **vascular spasm**, the smooth muscle in the walls of the vessel contracts dramatically. Small blood vessels have smooth muscle arranged in circular layers; larger vessels also have longitudinal layers of smooth muscle. The circular layers tend to constrict the flow of blood, whereas the longitudinal layers, when present, draw the vessel back into the surrounding tissue, often making it more difficult for a surgeon to locate, clamp, and tie off a severed vessel. The vascular spasm response is believed to be triggered by several chemicals called endothelins that are released by vessel-lining cells and by pain receptors in response to vessel injury. This phenomenon typically lasts for up to 30 minutes, although it can last for hours.

Formation of the Platelet Plug: In the second step, platelets, which normally float free in the plasma, encounter the area of vessel rupture with the exposed underlying connective tissue and collagenous fibers. The platelets begin to clump together, become spiked and sticky, and bind to the exposed collagen and endothelial lining. This process is assisted by a glycoprotein in the blood plasma called von Willebrand factor, which helps stabilize the growing **platelet plug**. As platelets collect, they simultaneously release chemicals from their granules into the plasma that further contribute to hemostasis. Among the substances released by the platelets are:

- Adenosine diphosphate (ADP), which helps additional platelets to adhere to the injury site, reinforcing and expanding the platelet plug
- Serotonin, which maintains vasoconstriction
- Prostaglandins and phospholipids, which also maintain vasoconstriction and help to activate further clotting chemicals

A platelet plug can temporarily seal a small opening in a blood vessel. Plug formation, in essence, buys the body time while more sophisticated and durable repairs are being made. In a similar manner, even modern naval warships still carry an assortment of wooden plugs to temporarily repair small breaches in their hulls until permanent repairs can be made.

Coagulation: Those more sophisticated and more durable repairs are collectively called **coagulation**, the formation of a blood clot. The process is sometimes characterized as a cascade, because one event prompts the next as in a multi-level waterfall. The result is the production of a gelatinous but robust clot made up of a mesh of **fibrin –** an insoluble filamentous protein derived from the blood plasma protein fibrinogen (Table 1) – in which platelets and blood cells are trapped.

Clotting Factors Involved in Coagulation: In the coagulation cascade, chemicals called **clotting factors** (or coagulation factors) prompt reactions that activate still more coagulation factors (Figure 6b). The process is complex, but is initiated along two basic pathways: the extrinsic pathway which normally is triggered by tissue damage, and the intrinsic pathway which begins in the bloodstream and is triggered by damage to the wall of the vessel.

Both of these merge into a third pathway, referred to as the common pathway (Figure 6b). All three pathways are dependent upon the 12 known clotting factors, including Ca^{2+} and vitamin K (Table 3). Clotting factors are secreted primarily by the liver and the platelets. The liver requires the fat-soluble vitamin K to produce many of them. Vitamin K (along with biotin and folate) is somewhat unusual among vitamins in that it is not only consumed in the diet but is also synthesized by bacteria residing in the large intestine. The calcium ion, also considered as factor IV, is derived from the diet and from the breakdown of bone. Some recent evidence indicates that activation of various clotting factors occurs on specific receptor sites on the surfaces of platelets.

The 12 clotting factors are numbered I through XIII according to the order of their discovery. Factor VI was once

believed to be a distinct clotting factor, but is now thought to be identical to factor V. Rather than renumber the other factors, factor VI was allowed to remain as a placeholder and also a reminder that knowledge changes over time.

Extrinsic Pathway: The quicker responding and more direct **extrinsic pathway** (also known as the **tissue factor** pathway) begins when damage occurs to the surrounding tissues, such as in a traumatic injury. Upon contact with blood plasma, the damaged extravascular cells, which are extrinsic to the bloodstream, release factor III (thromboplastin). Ca^{2+} and factor VII (proconvertin), activated by factor III forms an enzyme complex. This enzyme complex leads to activation of factor X (Stuart–Prower factor), which activates the common pathway discussed below. The events in the extrinsic pathway are completed in a matter of seconds.

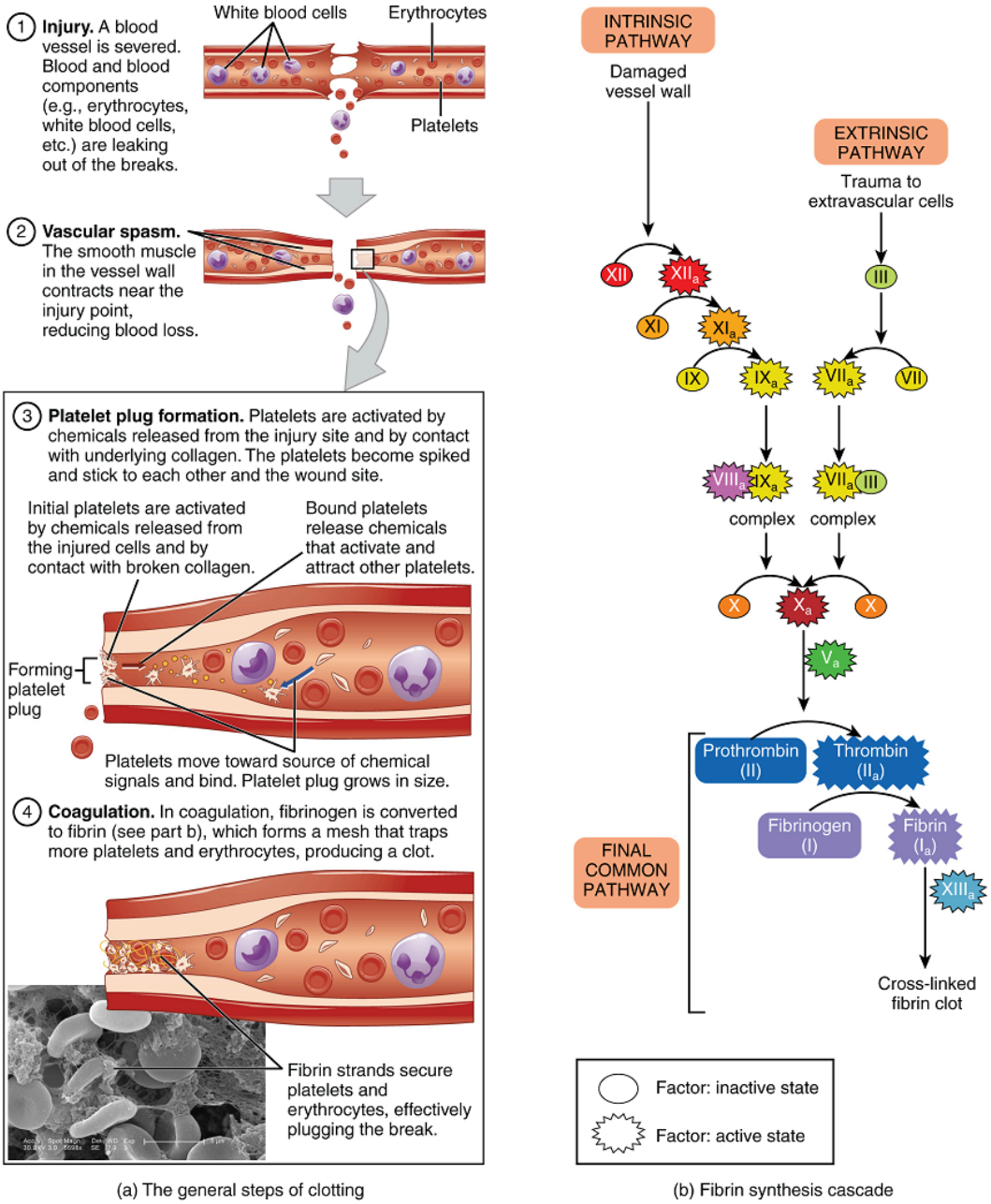


Figure 6. Hemostasis. *(a) An injury to a blood vessel initiates the process of hemostasis. Sealing a damaged blood vessel involved three main processes. First, vascular spasm constricts the flow of blood. Next, a platelet plug forms to temporarily seal small openings in the vessel. (b) Coagulation enables the repair of the vessel wall once the leakage of blood has stopped. The synthesis of fibrin in blood clots involves either an intrinsic pathway or an extrinsic pathway, both of which lead to a common pathway. (credit a: Kevin MacKenzie)*

Intrinsic Pathway: The intrinsic pathway (also known as the contact activation pathway) is longer and more

complex. In this case, the clotting factors involved are all intrinsic to (present within) the bloodstream. This pathway is prompted by damage to the walls of blood vessels that exposes the initiating clotting factor (clotting factor XII) to collagen. Within the body, factor XII is typically activated when it encounters negatively charged molecules, such as inorganic polymers and phosphate produced earlier in the series of intrinsic pathway reactions. Factor XII sets off a series of reactions to form an enzyme complex that activates factor X (Stuart-Prower factor or thrombokinase), leading to the common pathway. The events in the intrinsic pathway are completed in a few minutes.

Table 3: Clotting factors
(*vitamin K required)

Factor	Name	Type of molecule	Source	Pathway(s)
I	Fibrinogen	Plasma protein	Liver	Common; converted into fibrin
II	Prothrombin	Plasma protein	Liver*	Common; converted into fibrin
III	Tissue thromboplastin or tissue factor (TF)	Lipoprotein mixture	Damaged cells and platelets	Extrinsic
IV	Calcium ions (Ca^{2+})	Inorganic ions in plasma	Diet, platelets, bone matrix	Entire process
V	Proaccelerin	Plasma protein	Liver, platelets	Extrinsic and intrinsic
VI	Not used (historical use; identical to factor V)	Not used	Not used	Not used
VII	Proconvertin	Plasma protein	Liver*	Extrinsic
VIII	Antihemolytic factor A	Plasma protein factor	Platelets and endothelial cells	Intrinsic; deficiency results in hemophilia A
IX	Antihemolytic factor B (plasma thromboplastin component)	Plasma protein	Liver*	Intrinsic; deficiency results in hemophilia B
X	Thrombokinase (Stuart-Prower factor)	Protein	Liver*	Extrinsic and intrinsic
XI	Antihemolytic factor C (plasma thromboplastin antecedent)	Plasma protein	Liver	Intrinsic; deficiency results in hemophilia C
XII	Hageman factor	Plasma protein	Liver	Intrinsic; initiates clotting in vitro; activates plasmin
XIII	Fibrin-stabilizing factor	Plasma protein	Liver, platelets	Stabilizes fibrin; slows fibrinolysis

Common Pathway: Both the intrinsic and extrinsic pathways lead to the **common pathway**, in which fibrin is produced to seal off the vessel. Once factor X has been activated by either the intrinsic or extrinsic pathway, the enzyme prothrombinase converts factor II, the inactive enzyme prothrombin, into the active enzyme **thrombin**. (Note that if the enzyme thrombin were not normally in an inactive form, clots would form spontaneously, a condition not consistent with life.) Then, thrombin converts factor I, the soluble fibrinogen, into insoluble fibrin protein strands. Factor XIII then stabilizes the fibrin clot.

The stabilized clot is acted upon by contractile proteins within the platelets. As these proteins contract, they pull on the fibrin threads, bringing the edges of the clot more tightly together, somewhat as we do when tightening loose shoelaces. This process also wrings out of the clot a small amount of fluid called **serum**, which is blood plasma without its clotting factors.

Fibrinolysis: To restore normal blood flow as the vessel heals, the clot must eventually be removed. **Fibrinolysis** is the gradual degradation of the clot. Again, there is a fairly complicated series of reactions that involves factor XII and protein-catabolizing enzymes. During this process, the inactive protein plasminogen is converted into the active plasmin, which gradually breaks down the fibrin of the clot. Additionally, bradykinin, a vasodilator, is released, reversing the effects of the serotonin and prostaglandins from the platelets. This allows the smooth muscle in the walls of the vessels to relax and helps to restore the circulation.

Plasma Anticoagulants: An anticoagulant is any substance that opposes coagulation. Several circulating plasma anticoagulants play a role in limiting the coagulation process to the region of injury and restoring a normal, clot-free condition of blood. For instance, a cluster of proteins collectively referred to as the protein C system inactivates clotting factors involved in the intrinsic pathway. Tissue factor pathway inhibitor (TFPI) inhibits the conversion of the inactive factor VII to the active form in the extrinsic pathway. **Antithrombin** inactivates factor X and opposes the conversion of prothrombin (factor II) to thrombin in the common pathway. Basophils release **heparin**, a short-acting anticoagulant that also opposes prothrombin. Heparin is also found on the surfaces of cells lining the blood vessels. A pharmaceutical form of heparin is often administered therapeutically, for example, in surgical patients at risk for blood clots.

Among the many known biochemical activities of aspirin is its role as an anticoagulant. Aspirin (acetylsalicylic acid) is very effective at inhibiting the aggregation of platelets. It is routinely administered during a heart attack or stroke to reduce the adverse effects. Physicians sometimes recommend that patients at risk for cardiovascular disease take a low dose of aspirin on a daily basis as a preventive measure. However, aspirin can also lead to serious side effects, including increasing the risk of ulcers. A patient is well advised to consult a physician before beginning any aspirin regimen.

Disorders of Clotting: Either an insufficient or an excessive production of platelets can lead to severe disease or death. As discussed earlier, an insufficient number of platelets, called thrombocytopenia, typically results in the inability of blood to form clots. This can lead to excessive bleeding, even from minor wounds.

Another reason for failure of the blood to clot is the inadequate production of functional amounts of one or more clotting factors. This is the case in the genetic disorder **hemophilia**, which is actually a group of related disorders, the most common of which is hemophilia A, accounting for approximately 80% of cases. This disorder results in the inability to synthesize sufficient quantities of factor VIII. Regular infusions of clotting factors isolated from healthy donors can help prevent bleeding in hemophiliac patients. At some point, genetic therapy may become a viable option.

A **thrombus** (plural = thrombi) is an aggregation of platelets, erythrocytes, and even WBCs typically trapped within a mass of fibrin strands. While the formation of a clot is normal following the hemostatic mechanism just described, thrombi can form within an intact or only slightly damaged blood vessel. A thrombus can seriously impede blood flow to or from a region and will cause a local increase in blood pressure. If flow is to be maintained, the heart will need to generate a greater pressure to overcome the resistance.

When a portion of a thrombus breaks free from the vessel wall and enters the circulation, it is referred to as an **embolus**. An embolus that is carried through the bloodstream can be large enough to block a vessel critical to a major organ. When it becomes trapped, an embolus is called an embolism. In the heart, brain, or lungs, an embolism may accordingly cause a heart attack, a stroke, or a pulmonary embolism. These are medical emergencies.

A class of drugs collectively known as thrombolytic agents can help speed up the degradation of an abnormal clot. If a thrombolytic agent is administered to a patient within 3 hours following a thrombotic stroke, the patient's prognosis improves significantly. However, some strokes are not caused by thrombi, but by hemorrhage. Thus, the cause must be determined before treatment begins. Tissue plasminogen activator is an enzyme that catalyzes the conversion of plasminogen to plasmin, the primary enzyme that breaks down clots. It is released naturally by endothelial cells but is also used in clinical medicine as a thrombolytic agent. New research is progressing using compounds isolated from the venom of some species of snakes, particularly vipers and cobras, which may also have therapeutic value as thrombolytic agents.

The Heart

In this section, you will explore the remarkable pump that propels the blood into the vessels. There is no single better word to describe the function of the heart other than "pump," since its contraction develops the pressure that ejects blood into the major vessels: the aorta and pulmonary trunk. From these vessels, the blood is distributed to the remainder of the body. Although the connotation of the term "pump" suggests a

mechanical device made of steel and plastic, the anatomical structure is a living, sophisticated muscle. As you read this chapter, try to keep these twin concepts in mind: pump and muscle.

Although the term “heart” is an English word, cardiac (heart-related) terminology can be traced back to the Latin term, “kardia.” Cardiology is the study of the heart, and cardiologists are the physicians who deal primarily with the heart.

Part 1: Heart Anatomy

The vital importance of the heart is obvious. If one assumes an average rate of contraction of 75 contractions per minute, a human heart would contract approximately 108,000 times in one day, more than 39 million times in one year, and nearly 3 billion times during a 75-year lifespan. Each of the major pumping chambers of the heart ejects approximately 70 mL blood per contraction in a resting adult. This would be equal to 5.25 liters of fluid per minute and approximately 14,000 liters per day. Over one year, that would equal 10,000,000 liters (2.6 million gallons) of blood sent through roughly 96,000 km (60,000 miles) of vessels. In order to understand how that happens, it is necessary to understand the anatomy and physiology of the heart.

Location of the Heart: The human heart is located within the thoracic cavity, medially between the lungs in the space known as the mediastinum (Figure 7). Within the mediastinum, the heart is separated from the other mediastinal structures by a tough membrane known as the pericardium, or pericardial sac, and sits in its own space called the **pericardial cavity**. The dorsal surface of the heart lies near the bodies of the vertebrae, and its anterior surface sits deep to the sternum and costal cartilages. The great veins, the superior and inferior venae cavae, and the great arteries, the aorta and pulmonary trunk, are attached to the superior surface of the heart, called the base. The base of the heart is located at the level of the third costal cartilage (Figure 7). The inferior tip of the heart, the apex, lies just to the left of the sternum between the junction of the fourth and fifth ribs near their articulation with the costal cartilages. The right side of the heart is deflected anteriorly, and the left side is deflected posteriorly. It is important to remember the position and orientation of the heart when placing a stethoscope on the chest of a patient and listening for heart sounds, and also when looking at images taken from a midsagittal perspective. The slight deviation of the apex to the left is reflected in a depression in the medial surface of the inferior lobe of the left lung, called the cardiac notch.

Shape and Size of the Heart: The shape of the heart is similar to a pinecone, rather broad at the superior surface and tapering to the apex (Figure 7). A typical heart is approximately the size of your fist: 12 cm (5 in) in length, 8 cm (3.5 in) wide, and 6 cm (2.5 in) in thickness. Given the size difference between most members of the sexes, the weight of a female heart is approximately 250–300 grams (9 to 11 ounces), and the weight of a male heart is approximately 300–350 grams (11 to 12 ounces). The heart of a well-trained athlete, especially one specializing in aerobic sports, can be considerably larger than this. Cardiac muscle responds to exercise in a manner similar to that of skeletal muscle. That is, exercise results in the addition of protein myofilaments that increase the size of the individual cells without increasing their numbers, a concept called hypertrophy. Hearts of athletes can pump blood more effectively at lower rates than those of nonathletes.

Chambers and Circulation through the Heart: The human heart consists of four chambers: The left side and the right side each have one **atrium** and one **ventricle**. Each of the upper chambers, the right atrium (plural = atria) and the left atrium, acts as a receiving chamber and contracts to push blood into the lower chambers, the right ventricle and the left ventricle, respectively. The ventricles serve as the primary pumping chambers of the heart, propelling blood to the lungs or to the rest of the body.

There are two distinct but linked circuits in the human circulation called the pulmonary and systemic circuits. Although both circuits transport blood and everything it carries, we can initially view the circuits from the point of view of gases. The **pulmonary circuit** transports blood to and from the lungs, where it picks up oxygen and delivers carbon dioxide for exhalation. The **systemic circuit** transports oxygenated blood to virtually all of the tissues of the body and returns relatively deoxygenated blood and carbon dioxide to the heart to be sent back to the pulmonary circulation.

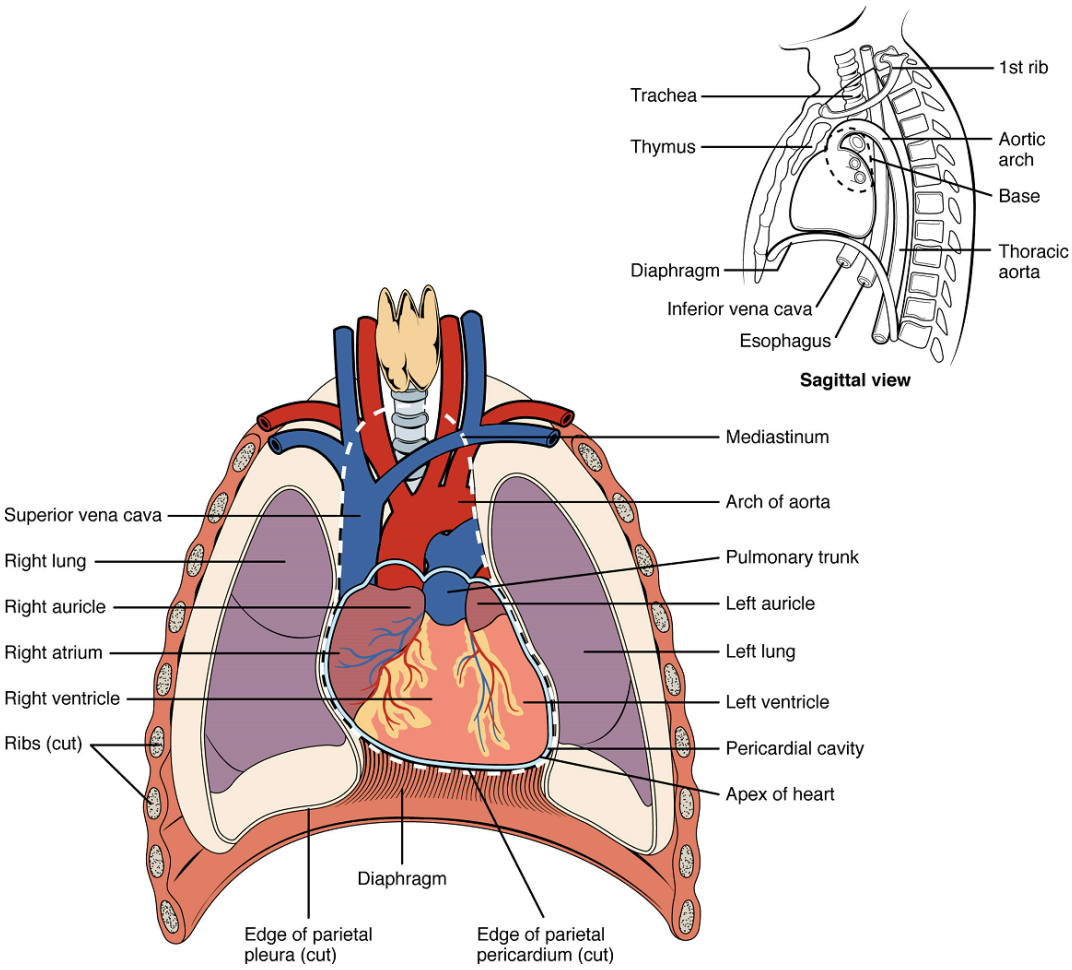


Figure 7. Position of the Heart in the Thorax. *The heart is located within the thoracic cavity, medially between the lungs in the mediastinum. It is about the size of a fist, is broad at the top, and tapers toward the base.*

The right ventricle pumps deoxygenated blood into the **pulmonary trunk**, which leads toward the lungs and bifurcates into the left and right **pulmonary arteries**. These vessels in turn branch many times before reaching the pulmonary capillaries, where gas exchange occurs: Carbon dioxide exits the blood and oxygen enters. The pulmonary trunk arteries and their branches are the only arteries in the post-natal body that carry relatively deoxygenated blood. Highly oxygenated blood returning from the pulmonary capillaries in the lungs passes through a series of vessels that join together to form the **pulmonary veins**—the only post-natal veins in the body that carry highly oxygenated blood. The pulmonary veins conduct blood into the left atrium, which pumps the blood into the left ventricle, which in turn pumps oxygenated blood into the aorta and on to the many branches of the systemic circuit. Eventually, these vessels will lead to the systemic capillaries, where exchange with the tissue fluid and cells of the body occurs. In this case, oxygen and nutrients exit the systemic capillaries to be used by the cells in their metabolic processes, and carbon dioxide and waste products will enter the blood.

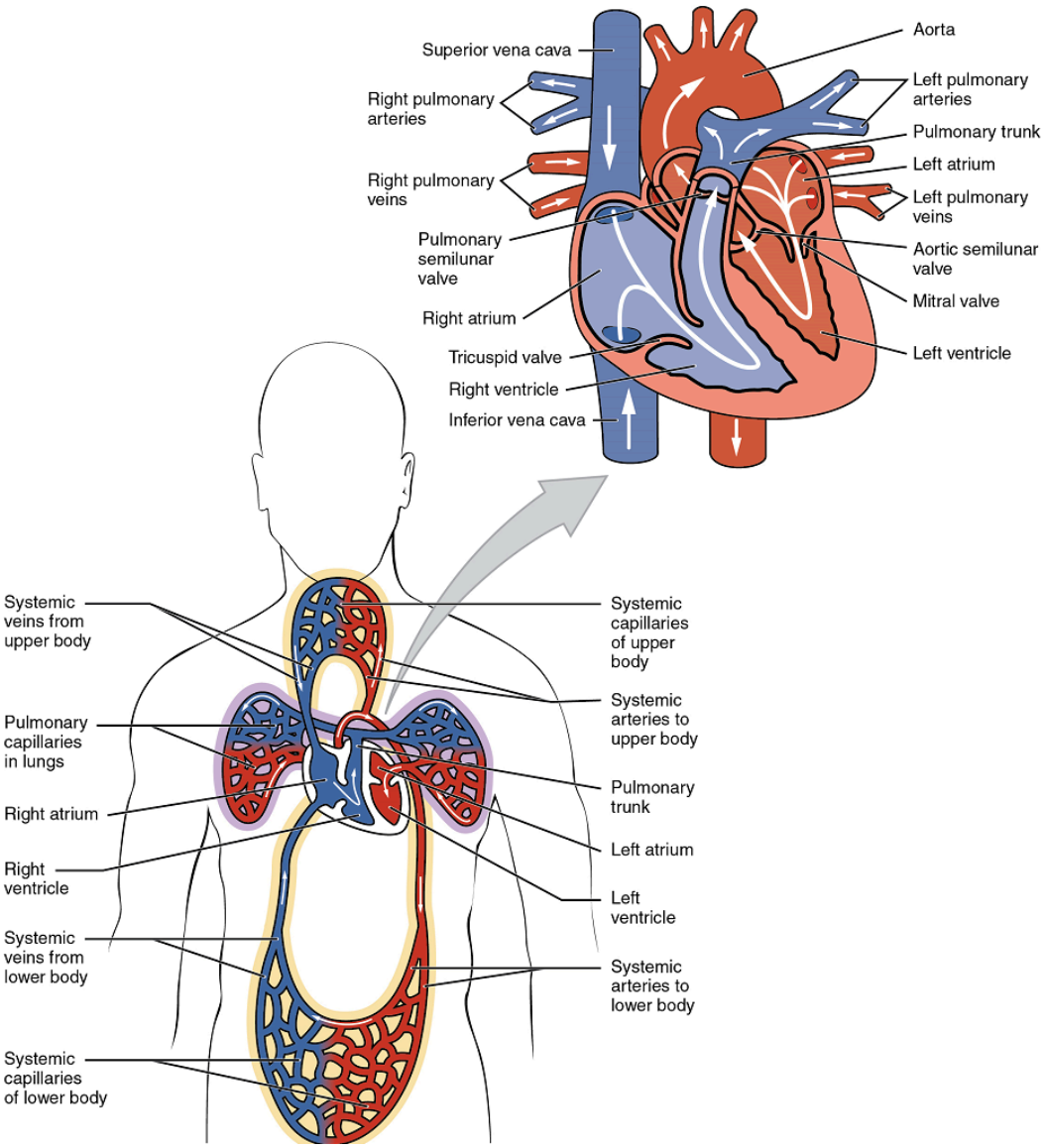


Figure 8. Dual System of the Human Blood Circulation. *Blood flows from the right atrium to the right ventricle, where it is pumped into the pulmonary circuit. The blood in the pulmonary artery branches is low in oxygen but relatively high in carbon dioxide. Gas exchange occurs in the pulmonary capillaries (oxygen into the blood, carbon dioxide out), and blood high in oxygen and low in carbon dioxide is returned to the left atrium. From here, blood enters the left ventricle, which pumps it into the systemic circuit. Following exchange in the systemic capillaries (oxygen and nutrients out of the capillaries and carbon dioxide and wastes in), blood returns to the right atrium and the cycle is repeated.*

The blood exiting the systemic capillaries is lower in oxygen concentration than when it entered. The capillaries will ultimately unite to form venules, joining to form ever-larger veins, eventually flowing into the two major systemic veins, the **superior vena cava** and the **inferior vena cava**, which return blood to the right atrium. The blood in the superior and inferior venae cavae flows into the right atrium, which pumps blood into the right ventricle. This process of blood circulation continues as long as the individual remains alive. Understanding the flow of blood through the pulmonary and systemic circuits is critical to all health professions (Figure 8).

Membranes, Surface Features, and Layers: Our exploration of more in-depth heart structures begins by examining the membrane that surrounds the heart, the prominent surface features of the heart, and the layers that form the wall of the heart. Each of these components plays its own unique role in terms of function.

Membranes: The membrane that directly surrounds the heart and defines the pericardial cavity is called the pericardium or pericardial sac. It also surrounds the "roots" of the major vessels, or the areas of closest proximity to the heart. The **pericardium**, which literally translates as "around the heart," consists of two distinct sublayers. The sturdy outer layer is the **fibrous pericardium**, made of tough, dense connective tissue that protects the heart and maintains its position in the thorax. The inner **serous pericardium** consists of two layers:

the outer **parietal pericardium**, which is fused to the fibrous pericardium, and an inner **visceral pericardium**, or **epicardium**, which is fused to the heart and is part of the heart wall. The pericardial cavity, filled with lubricating serous fluid, lies between the epicardium and the pericardium.

In most organs within the body, visceral serous membranes such as the epicardium are microscopic. However, in the case of the heart, it is not a microscopic layer but rather a macroscopic layer, consisting of a simple squamous epithelium called a **mesothelium**, reinforced with loose, irregular, or areolar connective tissue that attaches to the pericardium (Figure 9). This mesothelium secretes the lubricating serous fluid that fills the pericardial cavity and reduces friction as the heart contracts.

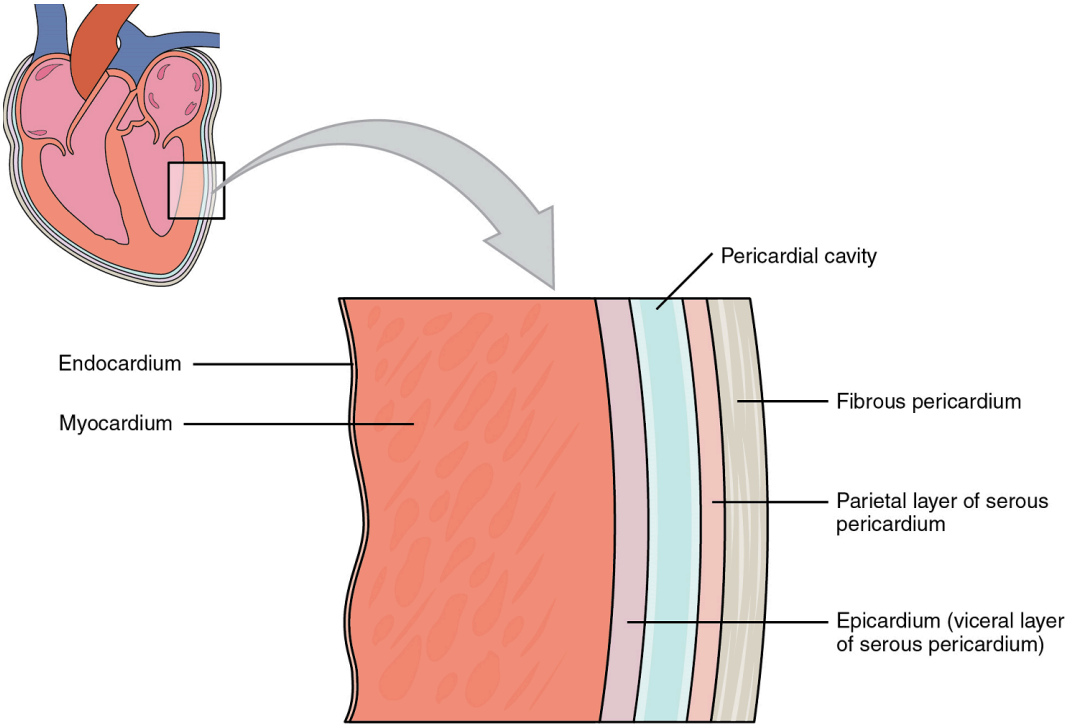


Figure 9. Pericardial Membranes and Layers of the Heart Wall. *The pericardial membrane that surrounds the heart consists of three layers and the pericardial cavity. The heart wall also consists of three layers. The pericardial membrane and the heart wall share the epicardium.*

Surface Features of the Heart: Inside the pericardium, the surface features of the heart are visible, including the four chambers (Figure 10). There is a superficial leaf-like extension of the atria near the superior surface of the heart, one on each side, called an **auricle**—a name that means "ear like"—because its shape resembles the external ear of a human. Auricles are relatively thin-walled structures that can fill with blood and empty into the atria or upper chambers of the heart. You may also hear them referred to as atrial appendages.

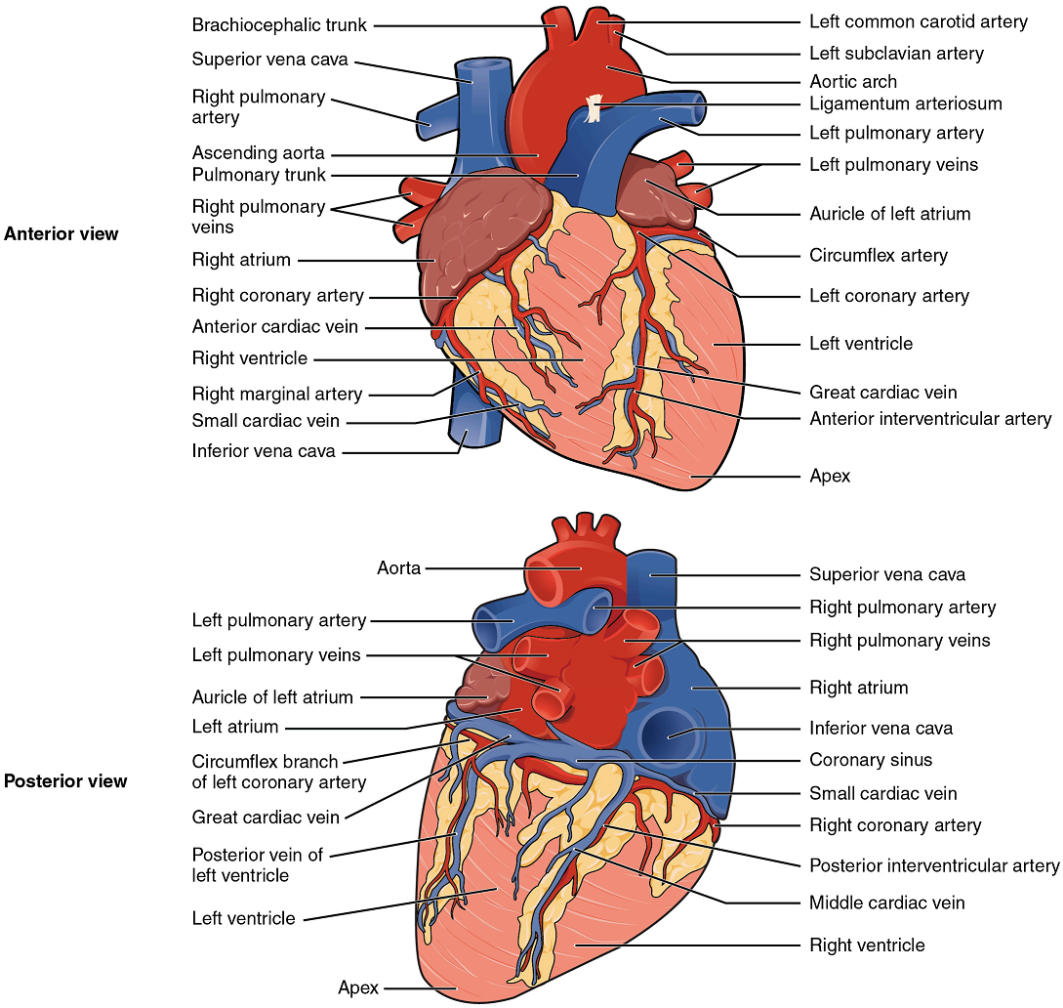


Figure 10. External Anatomy of the Heart. *Inside the pericardium, the surface features of the heart are visible.*

Layers: The wall of the heart is composed of three layers of unequal thickness. From superficial to deep, these are the epicardium, the myocardium, and the endocardium (Figure 9). The outermost layer of the wall of the heart is also the innermost layer of the pericardium, the epicardium, or the visceral pericardium discussed earlier.

The middle and thickest layer is the **myocardium**, made largely of cardiac muscle cells. It is built upon a framework of collagenous fibers, plus the blood vessels that supply the myocardium and the nerve fibers that help regulate the heart. It is the contraction of the myocardium that pumps blood through the heart and into the major arteries. The muscle pattern is elegant and complex, as the muscle cells swirl and spiral around the chambers of the heart (Figure 11). They form a figure 8 pattern around the atria and around the bases of the great vessels. Deeper ventricular muscles also form a figure 8 around the two ventricles and proceed toward the apex. More superficial layers of ventricular muscle wrap around both ventricles. This complex swirling pattern allows the heart to pump blood more effectively than a simple linear pattern would.

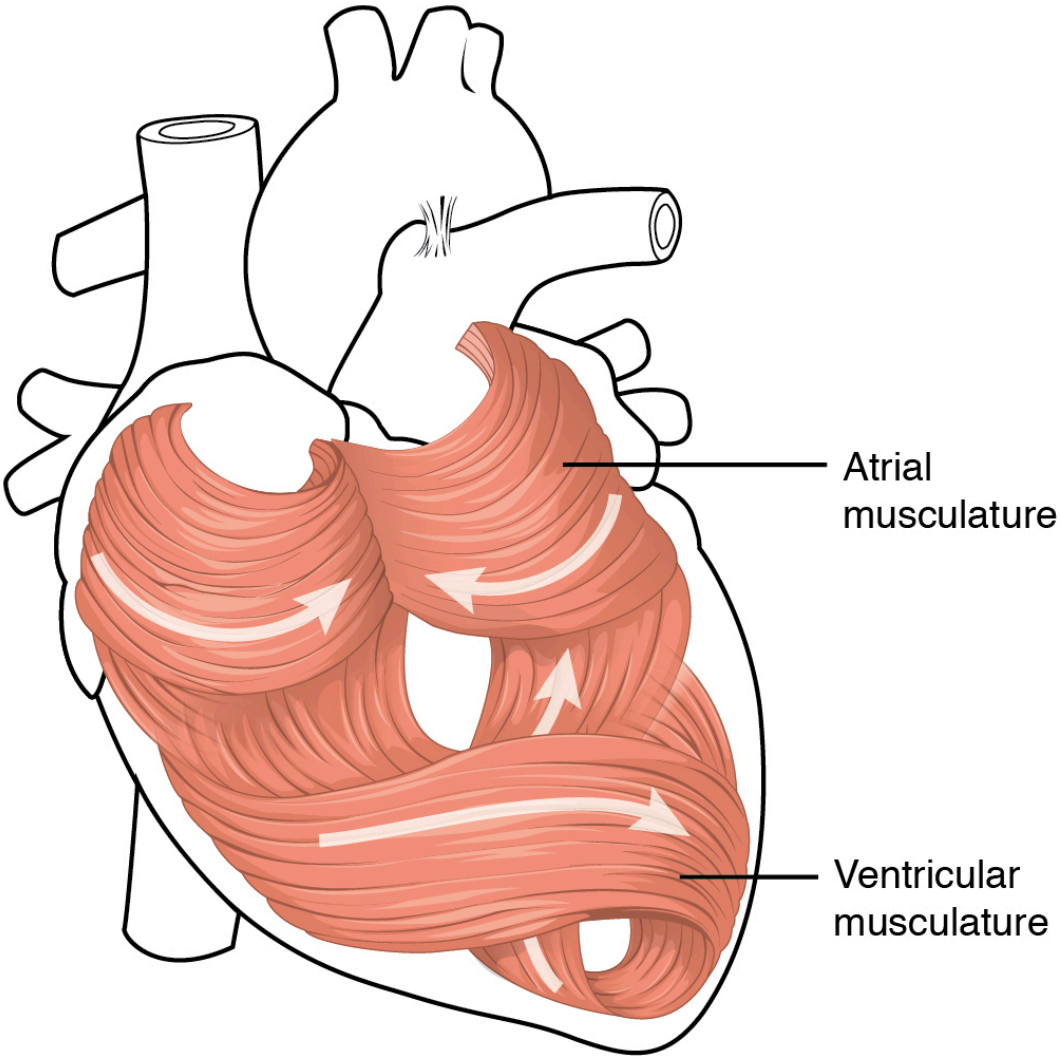


Figure 11. Heart Musculature. *The swirling pattern of cardiac muscle tissue contributes significantly to the heart's ability to pump blood effectively.*

Although the ventricles on the right and left sides pump the same amount of blood per contraction, the muscle of the left ventricle is much thicker and better developed than that of the right ventricle (Figure 12). In order to overcome the high resistance required to pump blood into the long systemic circuit, the left ventricle must generate a great amount of pressure. The right ventricle does not need to generate as much pressure, since the pulmonary circuit is shorter and provides less resistance.

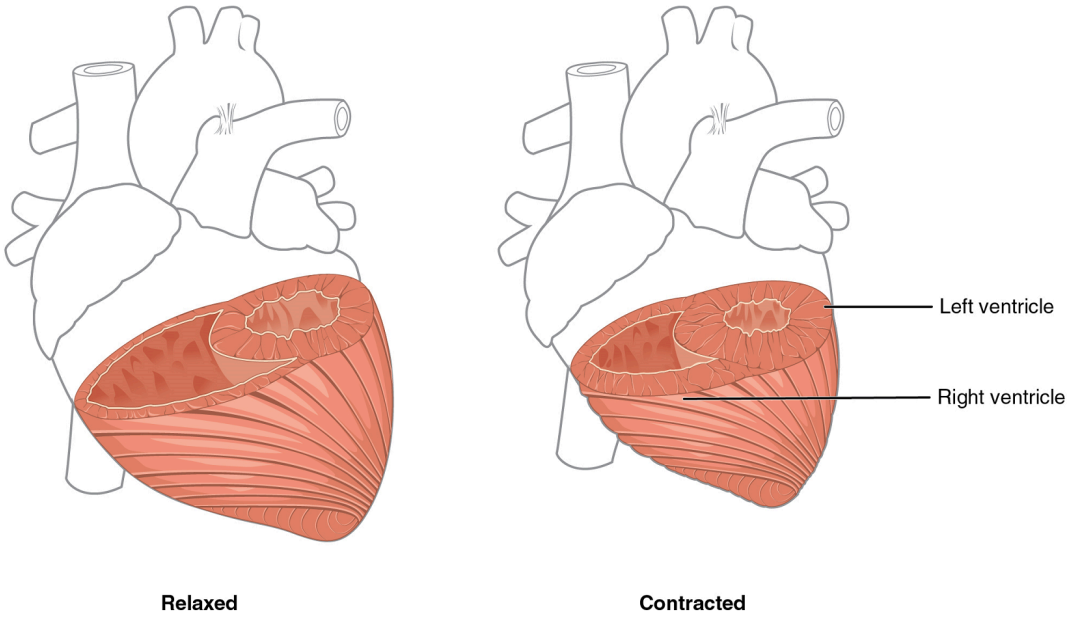


Figure 12. Differences in Ventricular Muscle Thickness. *The myocardium in the left ventricle is significantly thicker than that of the right ventricle. Both ventricles pump the same amount of blood, but the left ventricle must generate a much greater pressure to overcome greater resistance in the systemic circuit. The ventricles are shown in both relaxed and contracting states. Note the differences in the relative size of the lumens, the region inside each ventricle where the blood is contained.*

The innermost layer of the heart wall, the **endocardium**, is joined to the myocardium with a thin layer of connective tissue. The endocardium lines the chambers where the blood circulates and covers the heart valves. It is made of simple squamous epithelium called **endothelium**, which is continuous with the endothelial lining of the blood vessels (Figure 9).

Once regarded as a simple lining layer, recent evidence indicates that the endothelium of the endocardium and the coronary capillaries may play active roles in regulating the contraction of the muscle within the myocardium.

Internal Structure of the Heart: Recall that the heart's contraction cycle follows a dual pattern of circulation—the pulmonary and systemic circuits—because of the pairs of chambers that pump blood into the circulation. In order to develop a more precise understanding of cardiac function, it is first necessary to explore the internal anatomical structures in more detail.

Septa of the Heart: The word septum is derived from the Latin for "something that encloses;" in this case, a **septum** (plural = septa) refers to a wall or partition that divides the heart into chambers. The septa are physical extensions of the myocardium lined with endocardium. Located between the two atria is the **interatrial septum**. Normally in an adult heart, the interatrial septum bears an oval-shaped depression known as the **fossa ovalis**, a remnant of an opening in the fetal heart known as the **foramen ovale**. The foramen ovale allowed blood in the fetal heart to pass directly from the right atrium to the left atrium, allowing some blood to bypass the pulmonary circuit. Within seconds after birth, a flap of tissue known as the septum **primum** that previously acted as a valve closes the foramen ovale and establishes the typical cardiac circulation pattern.

Between the two ventricles is a second septum known as the **interventricular septum** (Figure 13). Unlike the interatrial septum, the interventricular septum is normally intact after its formation during fetal development. It is substantially thicker than the interatrial septum, since the ventricles generate far greater pressure when they contract.

The septum between the atria and ventricles is known as the **atrioventricular septum**. It is marked by the presence of four openings that allow blood to move from the atria into the ventricles and from the ventricles into the pulmonary trunk and aorta. Located in each of these openings between the atria and ventricles is a **valve**, a specialized structure that ensures one-way flow of blood. The valves between the atria and ventricles

are known generically as **atrioventricular valves**. The valves at the openings that lead to the pulmonary trunk and aorta are known generically as **semilunar valves**. Since these openings and valves structurally weaken the atrioventricular septum, the remaining tissue is heavily reinforced with dense connective tissue called the **cardiac skeleton**, or skeleton of the heart. It includes four rings that surround the openings between the atria and ventricles, and the openings to the pulmonary trunk and aorta, and serve as the point of attachment for the valves of the heart. The cardiac skeleton also provides an important boundary in the heart electrical conduction system.

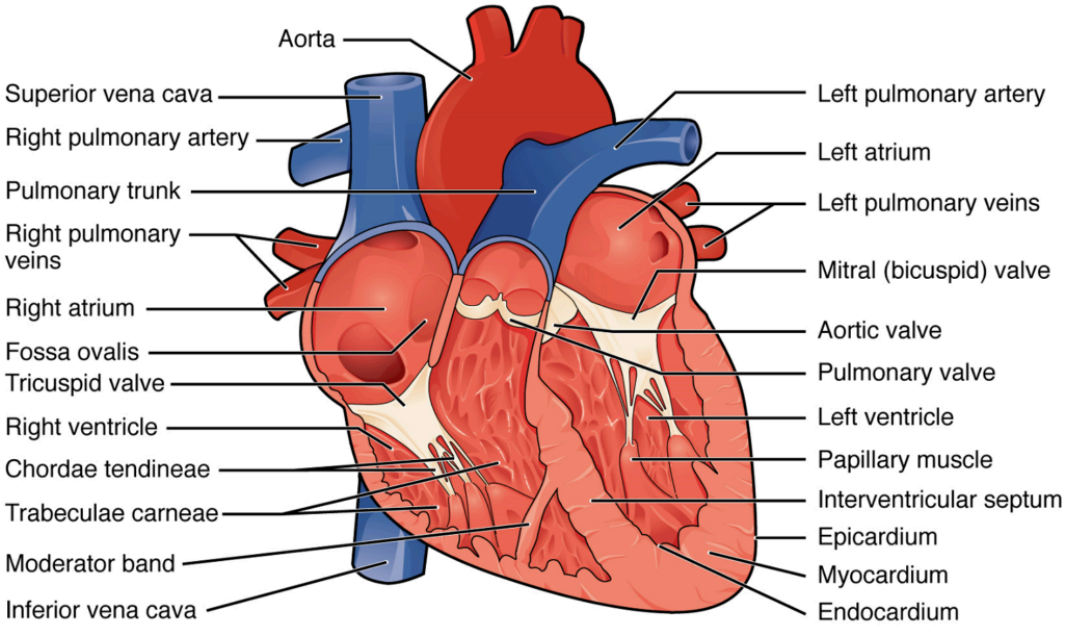


Figure 13. Internal Structures of the Heart. *This frontal section of the heart (anterior view) shows the four chambers, the major vessels and their early branches, as well as the valves. The presence of the pulmonary trunk and aorta covers the interatrial septum, and the atrioventricular septum is cut away to show the atrioventricular valves.*

Right Atrium: The right atrium serves as the receiving chamber for blood returning to the heart from the systemic circulation (Figure 13). The two major systemic veins, the superior and inferior venae cavae, and the large coronary vein called the **coronary sinus** that drains the heart myocardium empty into the right atrium (Figure 18). The superior vena cava drains blood from regions superior to the diaphragm: the head, neck, upper limbs, and the thoracic region. It empties into the superior and posterior portions of the right atrium. The inferior vena cava drains blood from areas inferior to the diaphragm: the lower limbs and abdominopelvic region of the body. It, too, empties into the posterior portion of the atria, but inferior to the opening of the superior vena cava. Immediately superior and slightly medial to the opening of the inferior vena cava on the posterior surface of the atrium is the opening of the coronary sinus. This thin-walled vessel drains most of the coronary veins that return systemic blood from the heart.

The atria receive venous blood on a nearly continuous basis, preventing venous flow from stopping while the ventricles are contracting. While most ventricular filling occurs while the atria are relaxed, they do demonstrate a contractile phase and actively pump blood into the ventricles just prior to ventricular contraction. The opening between the atrium and ventricle is guarded by the tricuspid valve.

Right Ventricle: The right ventricle receives blood from the right atrium through the tricuspid valve (Figure 13). Each flap of the valve is attached to strong strands of connective tissue, the **chordae tendineae**, literally "tendinous cords," or sometimes more poetically referred to as "heart strings" (Figure 14). There are several chordae tendineae associated with each of the flaps. They are composed of approximately 80% collagenous fibers with the remainder consisting of elastic fibers and endothelium. They connect each of the flaps to a **papillary muscle** that extends from the inferior ventricular surface (Figure 14). There are three papillary muscles

in the right ventricle, called the anterior, posterior, and septal muscles, which correspond to the three sections of the valves.

When the myocardium of the ventricle contracts, pressure within the ventricular chamber rises. Blood, like any fluid, flows from higher pressure to lower pressure areas, in this case, toward the pulmonary trunk and the atrium. To prevent any potential backflow, the papillary muscles also contract, generating tension on the chordae tendineae. This prevents the flaps of the valves from being forced into the atria and regurgitation of the blood back into the atria during ventricular contraction.

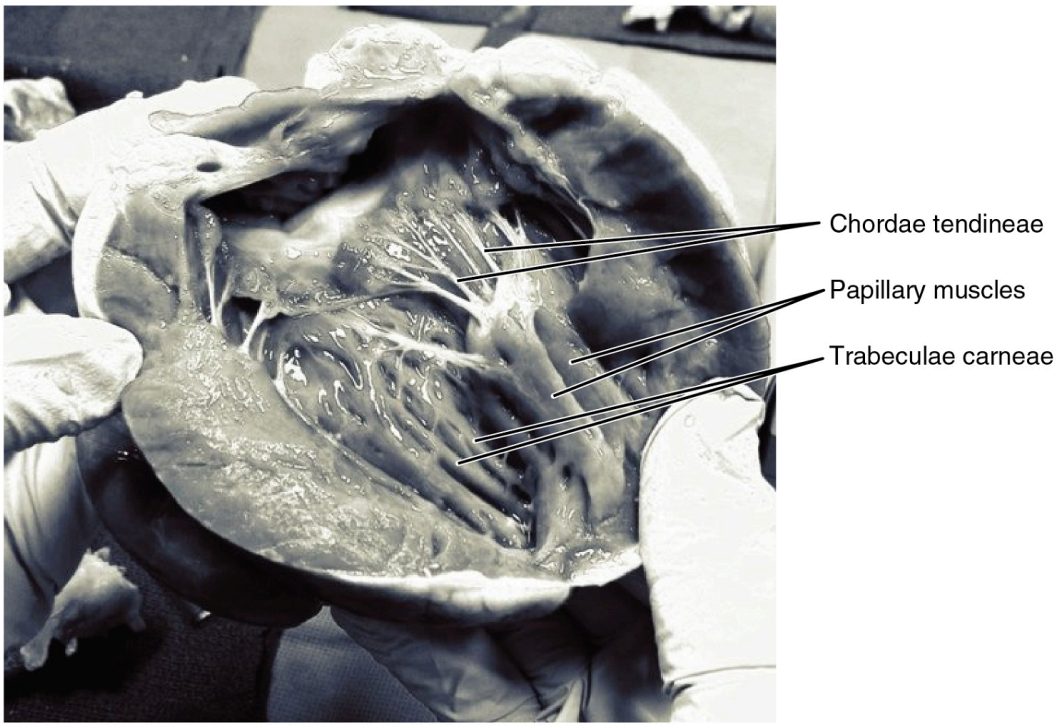


Figure 14. Chordae Tendineae and Papillary Muscles. *In this frontal section, you can see papillary muscles attached to the tricuspid valve on the right as well as the mitral valve on the left via chordae tendineae. (credit: modification of work by "PV KS"/flickr.com)*

Left Atrium: After exchange of gases in the pulmonary capillaries, blood returns to the left atrium high in oxygen via one of the four pulmonary veins (Figure 13). Blood flows nearly continuously from the pulmonary veins back into the atrium, which acts as the receiving chamber, and from here through an opening into the left ventricle. Most blood flows passively into the heart while both the atria and ventricles are relaxed, but toward the end of the ventricular relaxation period, the left atrium will contract, pumping blood into the ventricle. This atrial contraction accounts for approximately 20% of ventricular filling. The opening between the left atrium and ventricle is guarded by the mitral valve.

Left Ventricle: Recall that, although both sides of the heart will pump the same amount of blood, the muscular layer is much thicker in the left ventricle compared to the right (Figure 13). Like the right ventricle, the left also has trabeculae carneae. The mitral valve is connected to papillary muscles via chordae tendineae. There are two papillary muscles on the left – the anterior and posterior – as opposed to three on the right.

The left ventricle is the major pumping chamber for the systemic circuit; it ejects blood into the aorta through the aortic semilunar valve.

Heart Valve Structure and Function: A transverse section through the heart slightly above the level of the atrioventricular septum reveals all four heart valves along the same plane (Figure 15). The valves ensure unidirectional blood flow through the heart. Between the right atrium and the right ventricle is the **right atrioventricular valve**, or **tricuspid valve**. It typically consists of three flaps, or leaflets, made of endocardium

reinforced with additional connective tissue. The flaps are connected by chordae tendineae to the papillary muscles, which control the opening and closing of the valves.

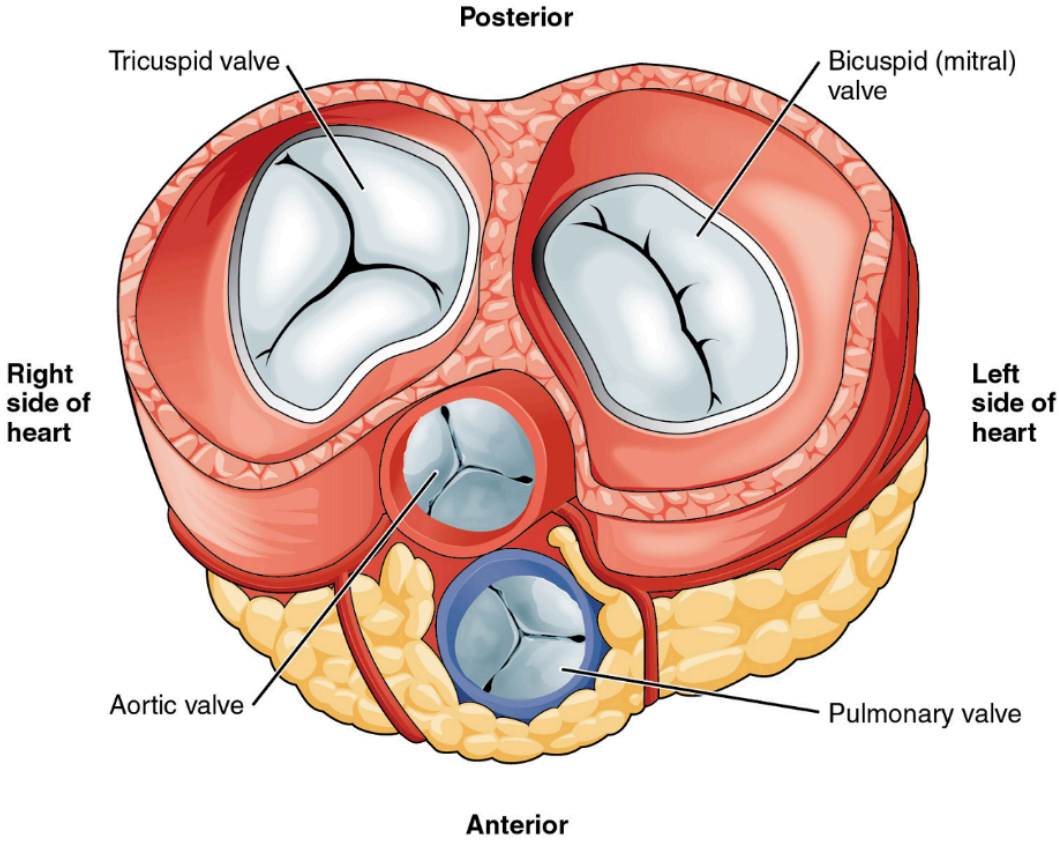


Figure 15. Heart Valves. *With the atria and major vessels removed, all four valves are clearly visible, although it is difficult to distinguish the three separate cusps of the tricuspid valve.*

Emerging from the right ventricle at the base of the pulmonary trunk is the pulmonary semilunar valve, or the **pulmonary valve**; it is also known as the pulmonic valve or the right semilunar valve. The pulmonary valve is comprised of three small flaps of endothelium reinforced with connective tissue. When the ventricle relaxes, the pressure differential causes blood to flow back into the ventricle from the pulmonary trunk. This flow of blood fills the pocket-like flaps of the pulmonary valve, causing the valve to close and producing an audible sound. Unlike the atrioventricular valves, there are no papillary muscles or chordae tendineae associated with the pulmonary valve.

Located at the opening between the left atrium and left ventricle is the **mitral valve**, also called the **bicuspid valve** or the left **atrioventricular valve**. Structurally, this valve consists of two cusps, known as the anterior medial cusp and the posterior medial cusp, compared to the three cusps of the tricuspid valve. In a clinical setting, the valve is referred to as the mitral valve, rather than the bicuspid valve. The two cusps of the mitral valve are attached by chordae tendineae to two papillary muscles that project from the wall of the ventricle.

At the base of the aorta is the aortic semilunar valve, or the **aortic valve**, which prevents backflow from the aorta. It normally is composed of three flaps. When the ventricle relaxes and blood attempts to flow back into the ventricle from the aorta, blood will fill the cusps of the valve, causing it to close and producing an audible sound.

When both atria and ventricles are relaxed, and when the atria contract to pump blood into the ventricles, the atrioventricular valves are open and the semilunar valves are closed (Figure 16).

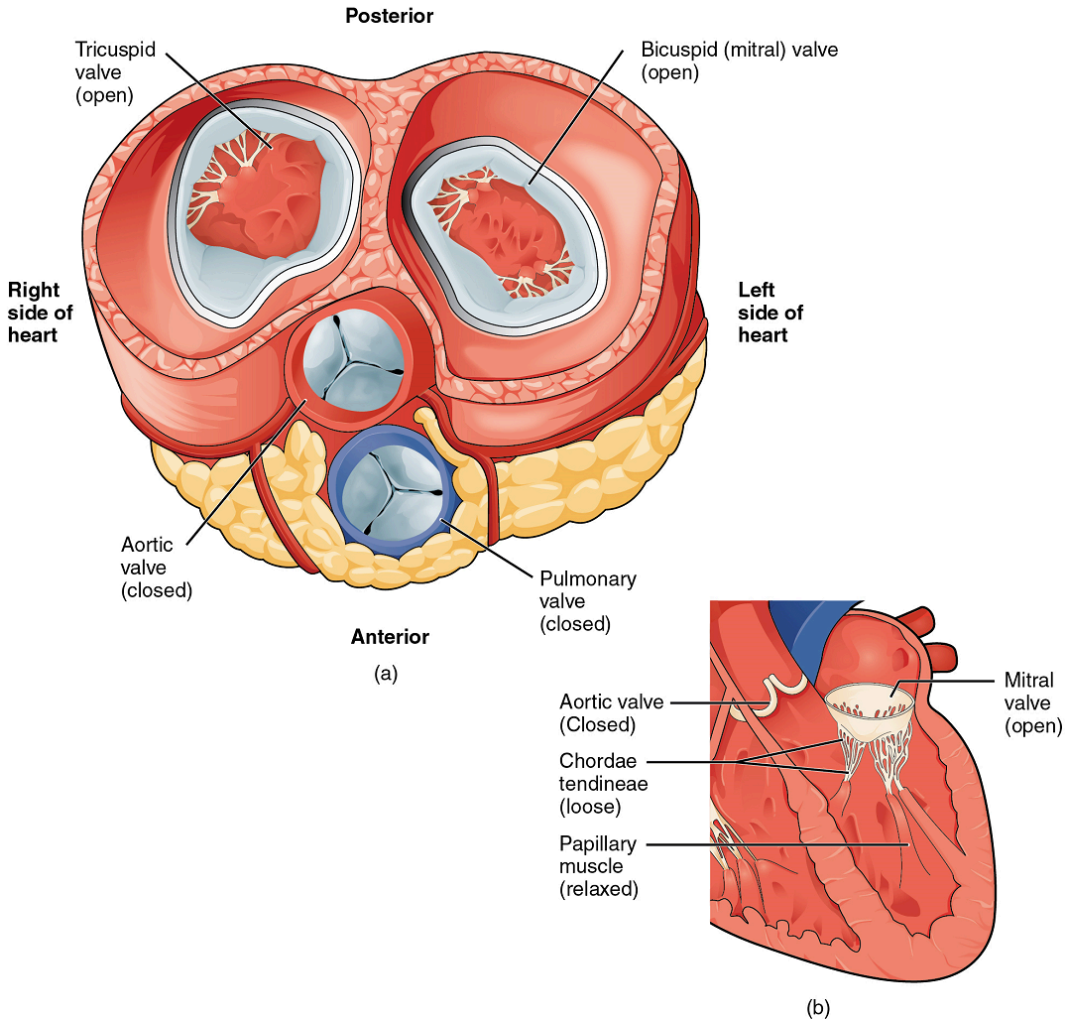


Figure 16. Blood Flow from the Left Atrium to the Left Ventricle. *(a) A transverse section through the heart illustrates the four heart valves. The two atrioventricular valves are open; the two semilunar valves are closed. The atria and vessels have been removed. (b) A frontal section through the heart illustrates blood flow through the mitral valve. When the mitral valve is open, it allows blood to move from the left atrium to the left ventricle. The aortic semilunar valve is closed to prevent backflow of blood from the aorta to the left ventricle.*

When the ventricles contract to eject blood into the pulmonary trunk and aorta, the atrioventricular valves close and the two semilunar valves open (Figure 17). Closure of the two atrioventricular valves prevents blood from being forced back into the atria.

When the ventricles begin to contract, pressure within the ventricles rises and blood flows toward the area of lowest pressure, which is initially in the atria. This backflow causes the cusps of the tricuspid and mitral (bicuspid) valves to close. These valves are tied down to the papillary muscles by chordae tendineae. During the relaxation phase of the cardiac cycle, the papillary muscles are also relaxed and the tension on the chordae tendineae is slight (Figure 16b). However, as the myocardium of the ventricle contracts, so do the papillary muscles. This creates tension on the chordae tendineae (Figure 17b), helping to hold the cusps of the atrioventricular valves in place and preventing them from being blown back into the atria.

The aortic and pulmonary semilunar valves lack the chordae tendineae and papillary muscles associated with the atrioventricular valves. Instead, they consist of pocket-like folds of endocardium reinforced with additional connective tissue. When the ventricles relax and the change in pressure forces the blood toward the ventricles, the blood presses against these cusps and seals the openings.

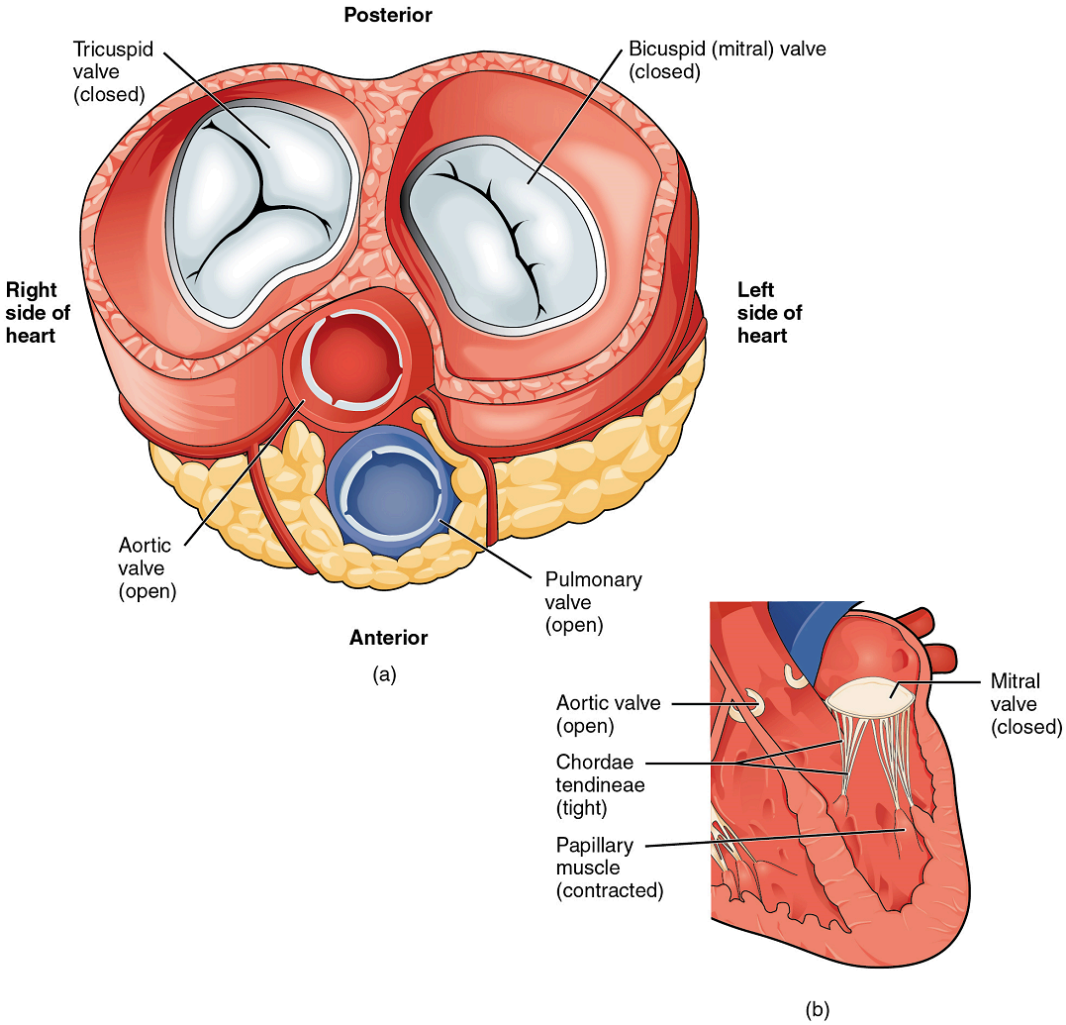


Figure 17. Blood Flow from the Left Ventricle into the Great Vessels. *(a) A transverse section through the heart illustrates the four heart valves during ventricular contraction. The two atrioventricular valves are closed, but the two semilunar valves are open. The atria and vessels have been removed. (b) A frontal view shows the closed mitral (bicuspid) valve that prevents backflow of blood into the left atrium. The aortic semilunar valve is open to allow blood to be ejected into the aorta.*

Coronary Circulation: You will recall that the heart is a remarkable pump composed largely of cardiac muscle cells that are incredibly active throughout life. Like all other cells, a **cardiomyocyte** requires a reliable supply of oxygen and nutrients, and a way to remove wastes, so it needs a dedicated, complex, and extensive coronary circulation. And because of the critical and nearly ceaseless activity of the heart throughout life, this need for a blood supply is even greater than for a typical cell. However, coronary circulation is not continuous; rather, it cycles, reaching a peak when the heart muscle is relaxed and nearly ceasing while it is contracting.

Coronary Arteries: Coronary arteries supply blood to the myocardium and other components of the heart. The first portion of the aorta after it arises from the left ventricle gives rise to the coronary arteries (Figure 18).

The left coronary artery distributes blood to the left side of the heart, the left atrium and ventricle, and the interventricular septum. The **circumflex artery** arises from the left coronary artery and follows the coronary sulcus to the **left anterior descending artery (LAD).** A coronary artery blockage often results in death of the cells (myocardial infarction) supplied by the particular vessel.

The right coronary artery proceeds along the coronary sulcus and distributes blood to the right atrium, portions of both ventricles, and the heart conduction system (Figure 19).

Coronary Veins: Coronary veins drain the heart and generally parallel the large surface arteries (Figure 18). Most drain into the coronary sinus. The coronary sinus is a large, thin-walled vein on the posterior surface of the heart lying within the coronary sulcus and emptying directly into the right atrium.

Watch this CrashCourse video for an overview of the heart! Direct link: https://youtu.be/X9ZZ6tcxArI

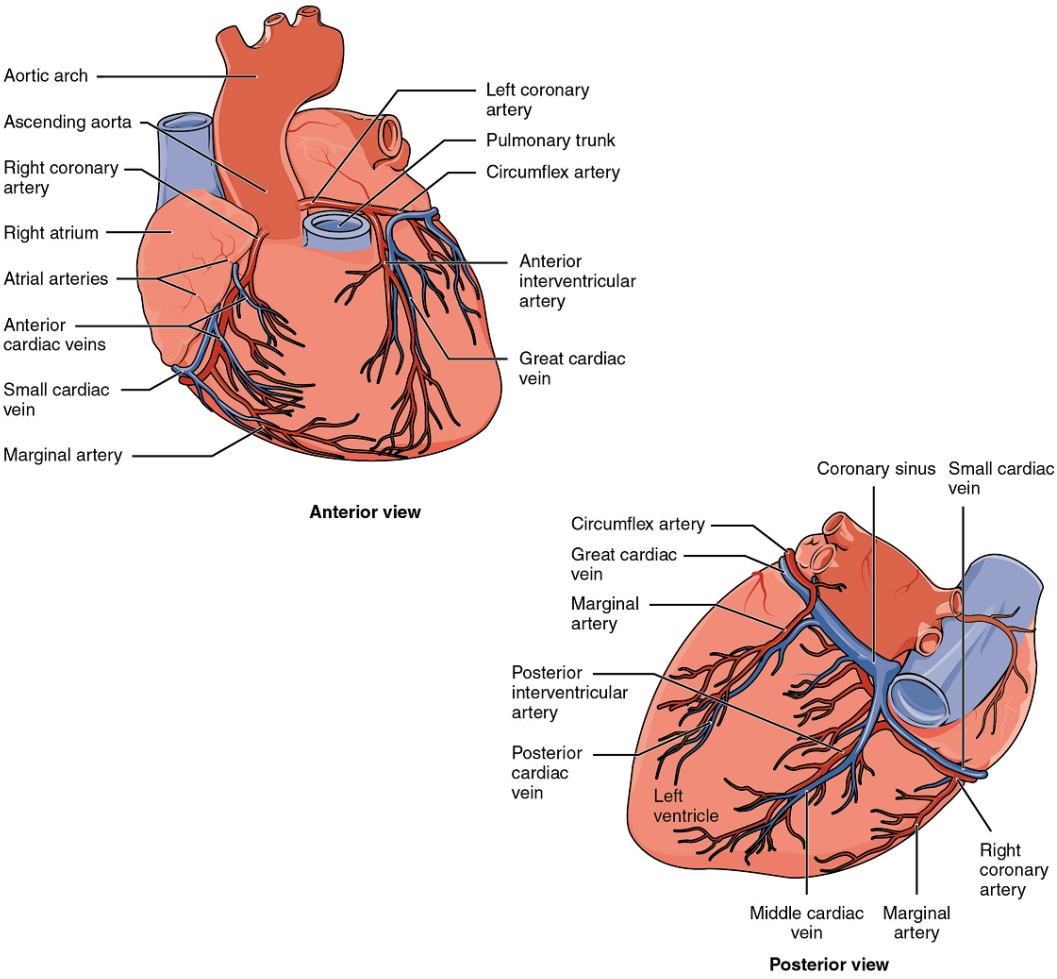


Figure 18. Coronary Circulation. *The anterior view of the heart shows the prominent coronary surface vessels. The posterior view of the heart shows the prominent coronary surface vessels.*

Part 2: Cardiac Muscle and Electrical Activity

Recall that cardiac muscle shares a few characteristics with both skeletal muscle and smooth muscle, but it has some unique properties of its own. Not the least of these exceptional properties is its ability to initiate an electrical potential at a fixed rate that spreads rapidly from cell to cell to trigger the contractile mechanism. This property is known as **autorhythmicity**. Neither smooth nor skeletal muscle can do this. Even though cardiac muscle has autorhythmicity, heart rate is modulated by the endocrine and nervous systems.

There are two major types of cardiac muscle cells: myocardial contractile cells and myocardial conducting cells. The **myocardial contractile cells** constitute the bulk (99%) of the cells in the atria and ventricles.

Contractile cells conduct impulses and are responsible for contractions that pump blood through the body. The **myocardial conducting cells** (1% of the cells) form the conduction system of the heart. Except for Purkinje fibers, they are generally much smaller than the contractile cells and have few of the myofibrils or filaments needed for contraction. Their function is similar in many respects to neurons, although they are specialized muscle cells. Myocardial conduction cells initiate and propagate the action potential (the electrical impulse) that travels throughout the heart and triggers the contractions that propel the blood.

Conduction System of the Heart: If embryonic heart cells are separated into a Petri dish and kept alive, each is capable of generating its own electrical impulse followed by contraction. When two independently beating embryonic cardiac muscle cells are placed together, the cell with the higher inherent rate sets the pace, and the impulse spreads from the faster to the slower cell to trigger a contraction. As more cells are joined together, the fastest cell continues to assume control of the rate. A fully developed adult heart maintains the capability of generating its own electrical impulse, triggered by the fastest cells, as part of the cardiac conduction system. The components of the cardiac conduction system include the **sinoatrial node (SA node)**, the **atrioventricular node (AV node)**, the **atrioventricular bundle (bundle of His)**, the **atrioventricular bundle branches**, and the **Purkinje fibers** (Figure 19).

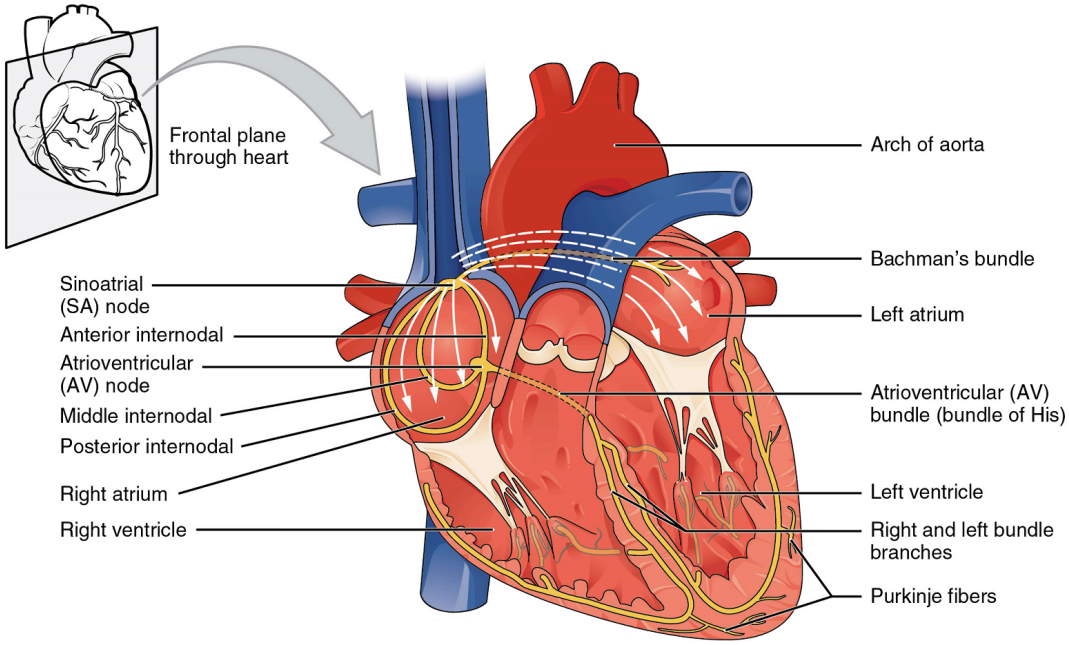


Figure 19. Conduction System of the Heart. *Specialized conducting components of the heart include the sinoatrial node, the internodal pathways, the atrioventricular node, the atrioventricular bundle, the right and left bundle branches, and the Purkinje fibers.*

Sinoatrial (SA) Node: Normal cardiac rhythm is established by the sinoatrial (SA) node, a specialized clump of myocardial conducting cells located in the superior and posterior walls of the right atrium in close proximity to the orifice of the superior vena cava. The sinoatrial node has the highest inherent rate of depolarization and is known as the **pacemaker** of the heart. It initiates the **sinus rhythm**, or normal electrical pattern followed by contraction of the heart.

This impulse spreads from its initiation in the sinoatrial node throughout the atria to the atrial myocardial contractile cells and the atrioventricular node (Figure 19). The impulse takes approximately 50 ms (milliseconds) to travel between these two nodes. When the impulse reaches the atrioventricular septum, the connective tissue of the cardiac skeleton prevents the impulse from spreading into the myocardial cells in the ventricles except at the atrioventricular node.

The electrical event, the wave of depolarization, is the trigger for muscular contraction. The wave of

depolarization begins in the right atrium, and the impulse spreads across the superior portions of both atria and then down through the contractile cells. The contractile cells then begin contraction from the superior to the inferior portions of the atria, efficiently pumping blood into the ventricles.

Atrioventricular (AV) Node: The **atrioventricular (AV) node** is a second clump of specialized myocardial conductive cells, located in the inferior portion of the right atrium within the atrioventricular septum. The septum prevents the impulse from spreading directly to the ventricles without passing through the atrioventricular node. There is a critical pause before the atrioventricular node depolarizes and transmits the impulse to the atrioventricular bundle (Figure 20).

This delay in transmission is partially attributable to the small diameter of the cells of the node, which slow the impulse. Also, conduction between nodal cells is less efficient than between cardiomyocytes to complete their contraction that pumps blood into the ventricles before the impulse is transmitted to the cells of the ventricle itself. With extreme stimulation by the sinoatrial node, the atrioventricular node can transmit impulses maximally at 220 per minute. This establishes the typical maximum heart rate in a healthy young individual. Damaged hearts or those stimulated by drugs can contract at higher rates, but at these rates, the heart can no longer effectively pump blood. It takes the impulse approximately 100 ms to pass through the atrioventricular node. This pause is critical to heart function, as it allows the atria to empty their blood into the ventricles.

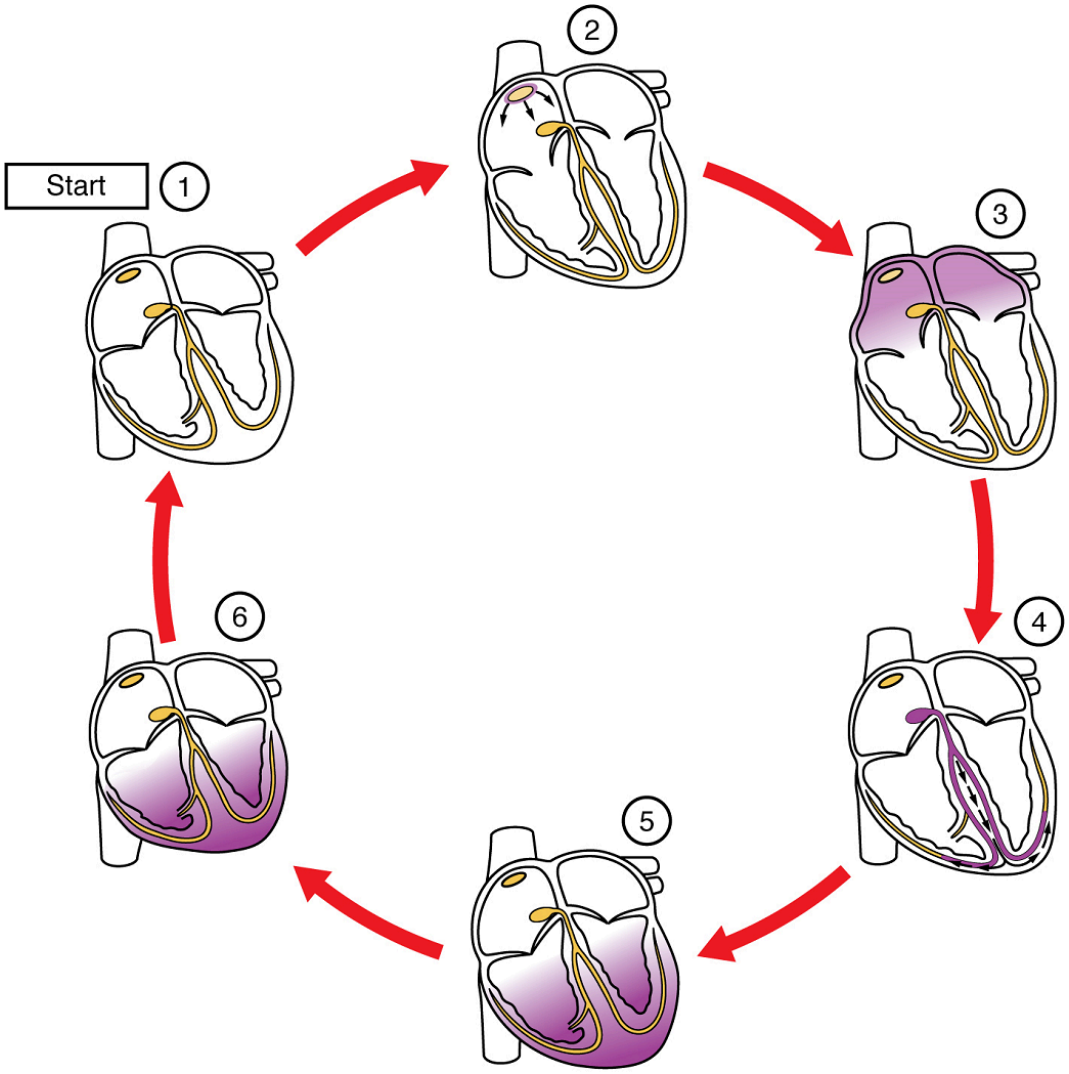


Figure 20. Cardiac Conduction. *(1) The sinoatrial (SA) node and the remainder of the conduction system are at rest. (2) The sinoatrial node initiates the action potential, which sweeps across the atria. (3) After reaching the atrioventricular node, there is a delay of approximately 100 ms that allows the atria to complete pumping blood before the impulse is transmitted to the atrioventricular bundle. (4) Following the delay, the impulse travels through the atrioventricular bundle and bundle branches to the Purkinje fibers, and also reaches the right papillary muscle via the moderator band. (5) The impulse spreads to the contractile fibers of the ventricle. (6) Ventricular contraction begins.*

Atrioventricular Bundle (Bundle of His), Bundle Branches, and Purkinje Fibers: Arising from the atrioventricular node, the **atrioventricular bundle**, or **bundle of His**, proceeds through the interventricular septum before dividing into two **atrioventricular bundle branches**, commonly called the left and right bundle branches. The left bundle branch supplies the left ventricle, and the right bundle branch the right ventricle. Both bundle branches descend and reach the apex of the heart where they connect with the Purkinje fibers (Figure 20). This passage takes approximately 25 ms.

The **Purkinje fibers** are additional myocardial conductive fibers that spread the impulse to the myocardial contractile cells in the ventricles. They extend throughout the myocardium from the apex of the heart toward the atrioventricular septum and the base of the heart. The Purkinje fibers have a fast inherent conduction rate, and the electrical impulse reaches all of the ventricular muscle cells in about 75 ms (Figure 20). Since the electrical stimulus begins at the apex, the contraction also begins at the apex and travels toward the base of the heart, similar to squeezing a tube of toothpaste from the bottom. However, the contraction of the ventricles is asynchronous with the right ventricle contracting slightly ahead of the left ventricle at the apex. This causes a twisting of the ventricles pushing blood towards major vessels leaving the heart. This allows the blood to be pumped out of the ventricles and into the aorta and pulmonary trunk in a more efficient manner. The total time elapsed from the initiation of the impulse in the sinoatrial node until depolarization of the ventricles is approximately 225 ms.

Electrocardiogram

By careful placement of surface electrodes on the body, it is possible to record the complex, compound electrical signal of the heart. This tracing of the electrical signal is the **electrocardiogram (ECG)**, also commonly abbreviated EKG (for "elektrokardiogramm", the German term for the test). Careful analysis of the ECG reveals a detailed picture of both normal and abnormal heart function, and is an indispensable clinical diagnostic tool. The standard electrocardiograph (the instrument that generates an ECG) uses 3, 5, or 12 leads. The greater the number of leads an electrocardiograph uses, the more information the ECG provides. The term "lead" may be used to refer to the cable from the electrode to the electrical recorder, but it typically describes the voltage difference between two of the electrodes. The 12-lead electrocardiograph uses 10 electrodes placed in standard locations on the patient's skin (Figure 21).

There are five prominent points on the ECG: the P wave, the QRS complex, and the T wave (Figure 22). The small **P wave** represents the depolarization of the atria. The atria begin contracting approximately 25 ms after the start of the P wave. The large **QRS complex** represents the depolarization of the ventricles, which requires a much stronger electrical signal because of the larger size of the ventricular cardiac muscle. The ventricles begin to contract as the QRS reaches the peak of the R wave. Lastly, the **T wave** represents the repolarization of the ventricles. The repolarization of the atria occurs during the QRS complex, which masks it on an ECG.

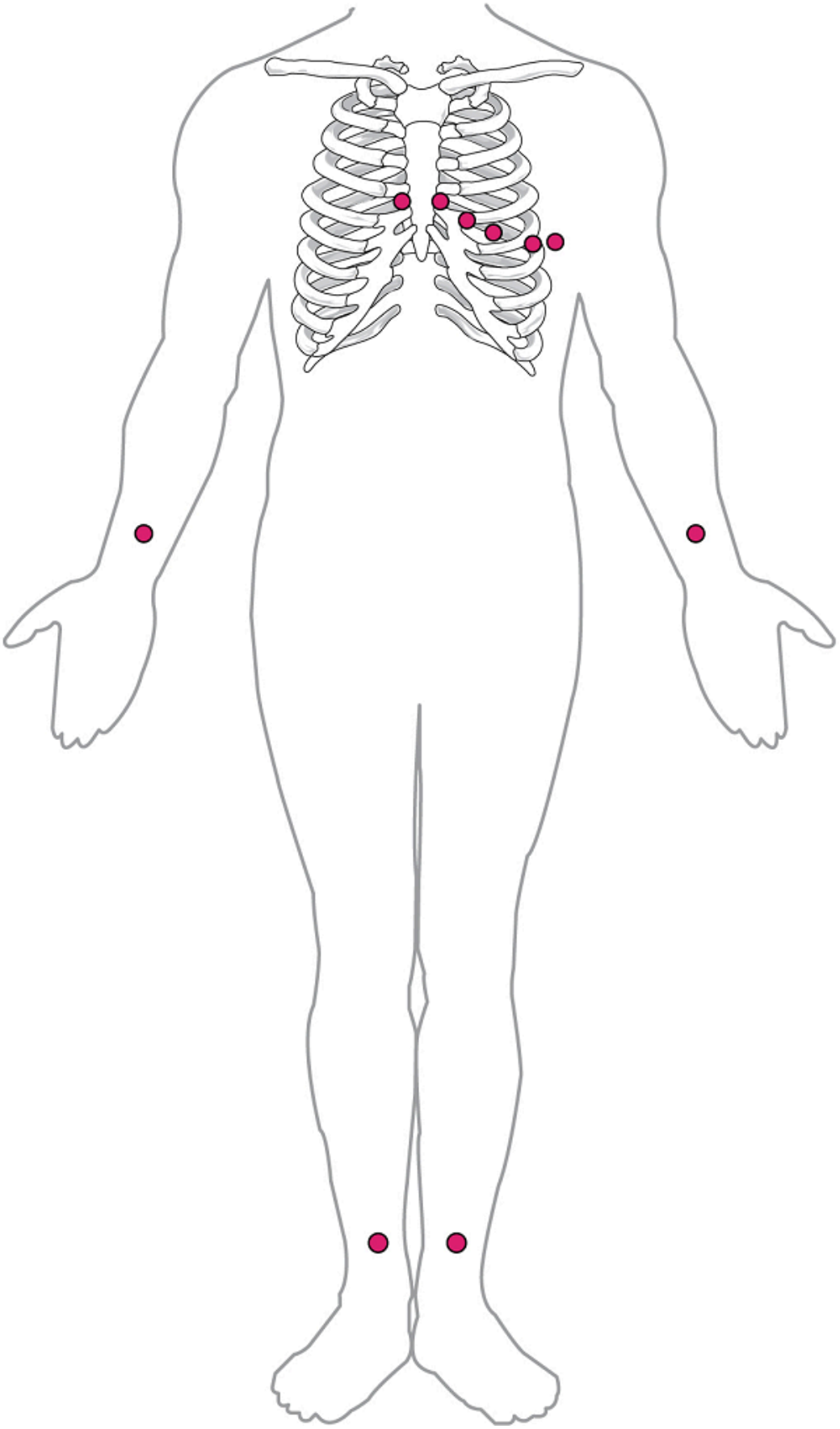


Figure 21. Standard Placement of ECG Leads. *In a 12-lead ECG, six electrodes are placed on the chest, and four electrodes are placed on the limbs.*

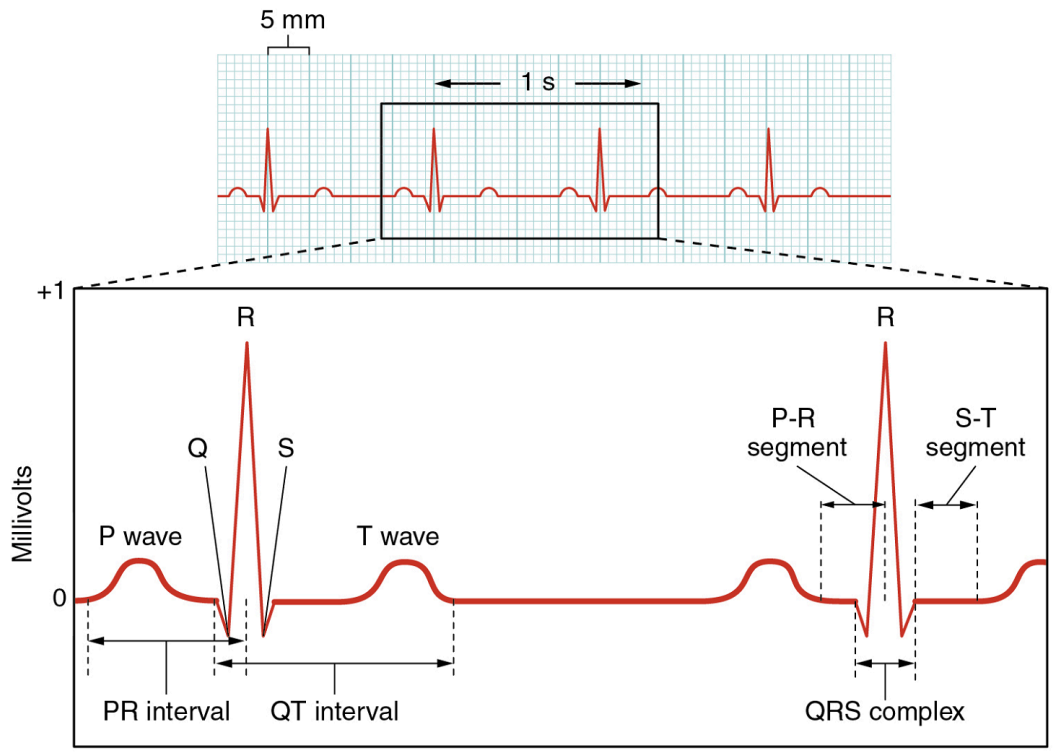


Figure 22. Electrocardiogram. *A normal tracing shows the P wave, QRS complex, and T wave. Also indicated are the PR and QT intervals, plus the P-R and S-T segments, which are medically useful pieces of information but the details of which are beyond the scope of this text.*

The depolarization events that appear on an ECG tracing should result in contraction of the corresponding chambers (Figure 23). Repolarization events measured then correspond with relaxation of the corresponding chambers.

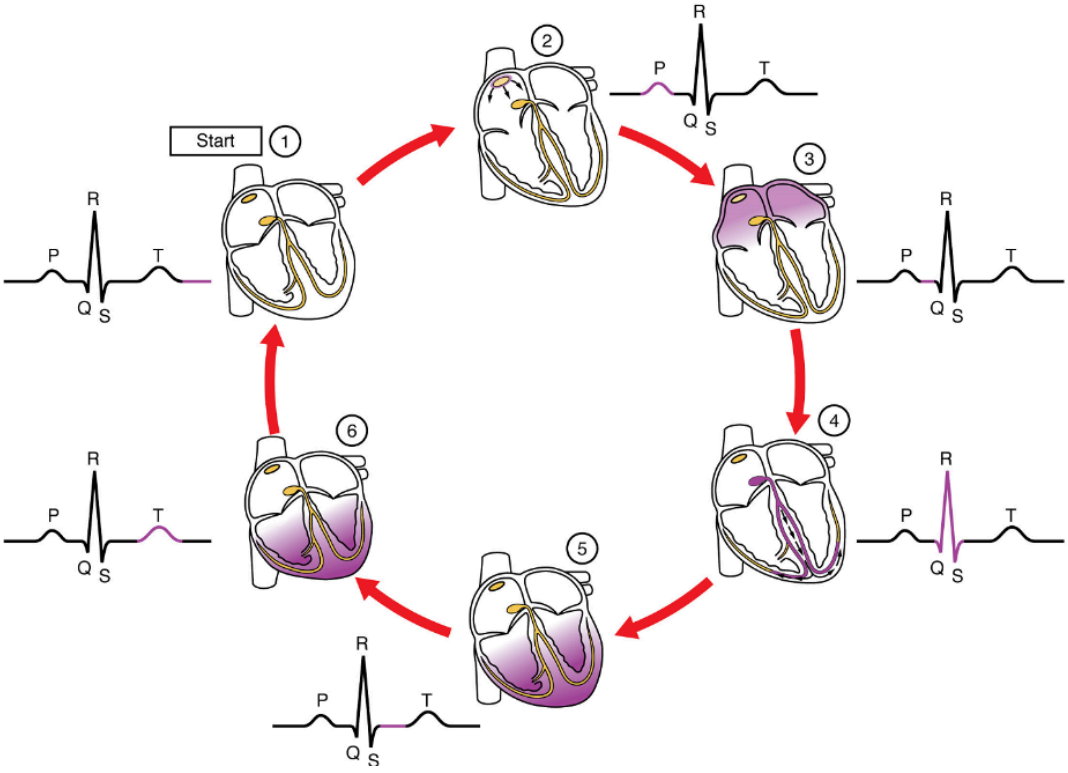


Figure 23. ECG Tracing Correlated to the Cardiac Cycle. *This diagram correlates an ECG tracing with the electrical and mechanical events of a heart contraction. Each segment of an ECG tracing corresponds to one event in the cardiac cycle.*

Part 3: Cardiac Cycle

The period of time that begins with contraction of the atria and ends with ventricular relaxation is known as the **cardiac cycle** (Figure 24). The period of contraction that the heart undergoes while it pumps blood into circulation is called **systole**. The period of relaxation that occurs as the chambers fill with blood is called **diastole**. Both the atria and ventricles undergo systole and diastole, and it is essential that these components be carefully regulated and coordinated to ensure blood is pumped efficiently to the body.

Pressures and Flow: Fluids, whether gases or liquids, are materials that flow according to pressure gradients—that is, they move from regions that are higher in pressure to regions that are lower in pressure. Accordingly, when the heart chambers are relaxed (diastole), blood will flow into the atria from the veins, which are higher in pressure. As blood flows into the atria, the pressure will rise, so the blood will initially move passively from the atria into the ventricles. When the action potential triggers the muscles in the atria to contract (atrial systole), the pressure within the atria rises further, pumping blood into the ventricles. During ventricular systole, pressure rises in the ventricles, pumping blood into the pulmonary trunk from the right ventricle and into the aorta from the left ventricle. Again, as you consider this flow and relate it to the conduction pathway, the elegance of the system should become apparent.

Phases of the Cardiac Cycle: At the beginning of the cardiac cycle, both the atria and ventricles are relaxed (diastole). Blood is flowing into the right atrium from the superior and inferior venae cavae and the coronary sinus. Blood flows into the left atrium from the four pulmonary veins. The two atrioventricular valves, the tricuspid and mitral valves, are both open, so blood flows unimpeded from the atria and into the ventricles. Approximately 70–80% of ventricular filling occurs by this method. The two semilunar valves, the pulmonary and aortic valves, are closed, preventing backflow of blood into the right and left ventricles from the pulmonary trunk on the right and the aorta on the left.

Atrial Systole and Diastole: Contraction of the atria follows depolarization, represented by the P wave of the ECG (Figure 25). As the atrial muscles contract from the superior portion of the atria toward the atrioventricular septum, pressure rises within the atria and blood is pumped into the ventricles through the open atrioventricular (tricuspid, and mitral or bicuspid) valves. At the start of atrial systole, the ventricles are normally filled with approximately 70–80% of their capacity due to inflow during diastole. Atrial contraction, also referred to as the "atrial kick," contributes the remaining 20–30% of filling (Figure 24). Atrial systole lasts approximately 100 ms and ends prior to ventricular systole, as the atrial muscle returns to diastole.

Ventricular Systole: Ventricular systole (Figure 24) follows the depolarization of the ventricles and is represented by the QRS complex in the ECG (Figure 25).

Initially, as the muscles in the ventricle contract, the pressure of the blood within the chamber rises, but it is not yet high enough to open the semilunar (pulmonary and aortic) valves and be ejected from the heart. However, blood pressure quickly rises above that of the atria that are now relaxed and in diastole. This increase in pressure causes blood to flow back toward the atria, closing the tricuspid and mitral valves.

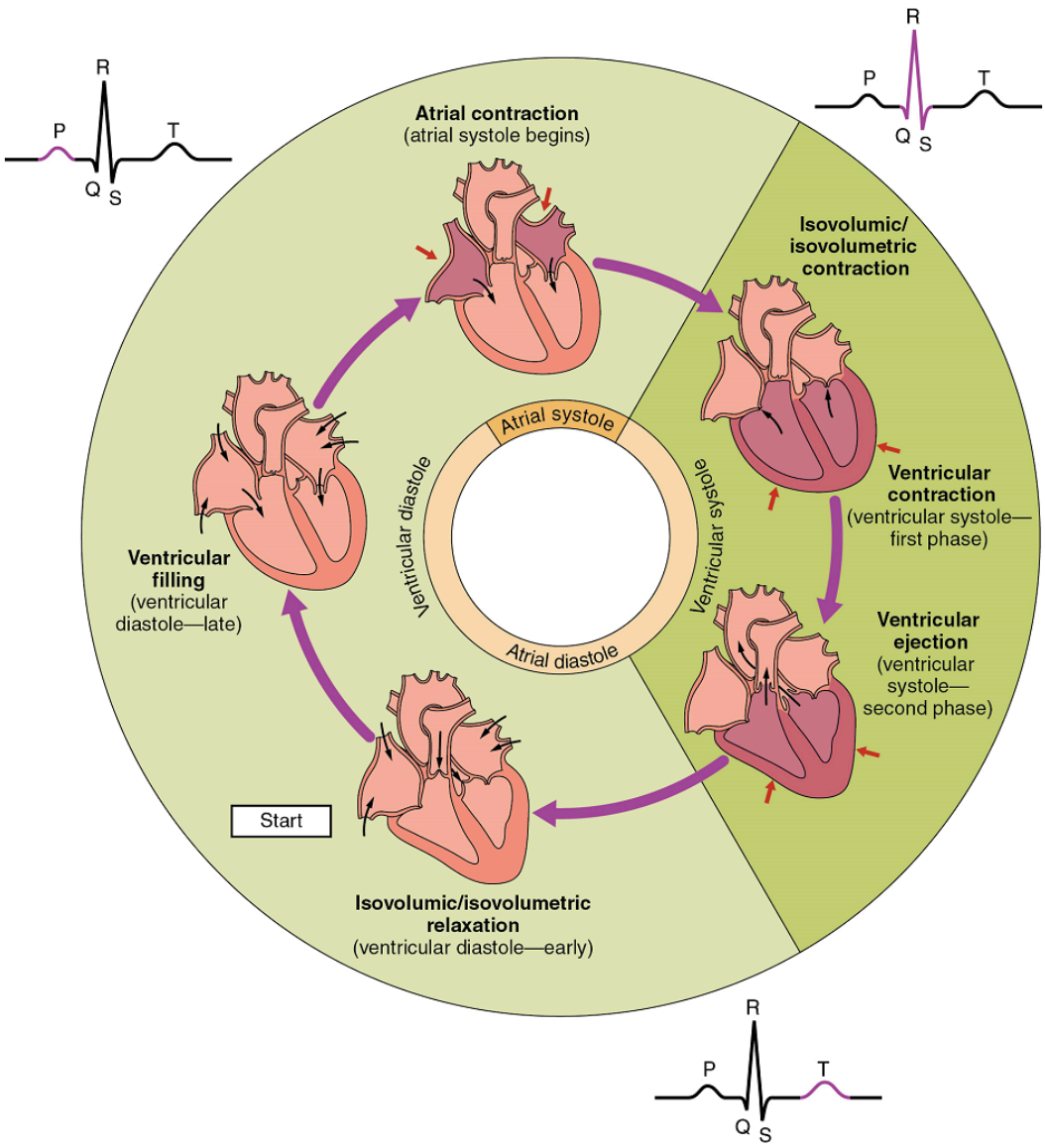


Figure 24. Overview of the Cardiac Cycle. *The cardiac cycle begins with atrial systole and progresses to ventricular systole, atrial diastole, and ventricular diastole, when the cycle begins again. Correlations to the ECG are highlighted.*

Eventually, the contraction of the ventricular muscle has raised the pressure within the ventricle to the point that it is greater than the pressures in the pulmonary trunk and the aorta. Blood is pumped from the heart, pushing open the pulmonary and aortic semilunar valves. Pressure generated by the left ventricle will be appreciably greater than the pressure generated by the right ventricle, since the existing pressure in the aorta will be so much higher. Nevertheless, both ventricles pump the same amount of blood. This quantity is referred to as stroke volume.

Ventricular Diastole: Ventricular relaxation, or diastole, follows repolarization of the ventricles and is represented by the T wave of the ECG (Figure 25). As the ventricular muscle relaxes, pressure on the remaining blood within the ventricle begins to fall. When pressure within the ventricles drops below pressure in both the pulmonary trunk and aorta, the semilunar valves close to prevent backflow into the heart (Figure 24).

As the ventricular muscle relaxes further, pressure on the blood within the ventricles drops even further. Eventually, it drops below the pressure in the atria. When this occurs, blood flows from the atria into the ventricles, pushing open the tricuspid and mitral valves. As pressure drops within the ventricles, blood flows from the major veins into the relaxed atria and from there into the ventricles. Both chambers are in diastole,

the atrioventricular valves are open, and the semilunar valves remain closed (Figure 24). The cardiac cycle is complete.

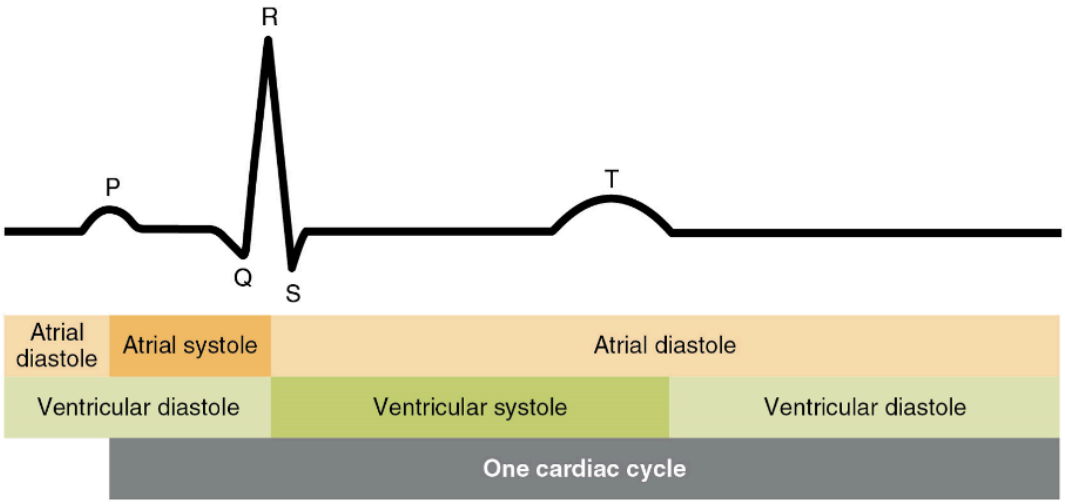


Figure 25. Relationship between the Cardiac Cycle and ECG. *Initially, both the atria and ventricles are relaxed (diastole). The P wave represents depolarization of the atria and is followed by atrial contraction (systole). Atrial systole extends until the QRS complex, at which point, the atria relax. The QRS complex represents depolarization of the ventricles and is followed by ventricular contraction. The T wave represents the repolarization of the ventricles and marks the beginning of ventricular relaxation.*

Heart Sounds: One of the simplest, yet effective, diagnostic techniques applied to assess the state of a patient's heart is auscultation using a stethoscope.

In a normal, healthy heart, there are only two audible **heart sounds**: S_1 and S_2. S_1 is the sound created by the closing of the atrioventricular valves during ventricular contraction and is normally described as a "lub," or first heart sound. The second heart sound, S_2, is the sound of the closing of the semilunar valves during ventricular diastole and is described as a "dub" (Figure 26). In both cases, as the valves close, the openings within the atrioventricular septum guarded by the valves will become reduced, and blood flow through the opening will become more turbulent until the valves are fully closed.

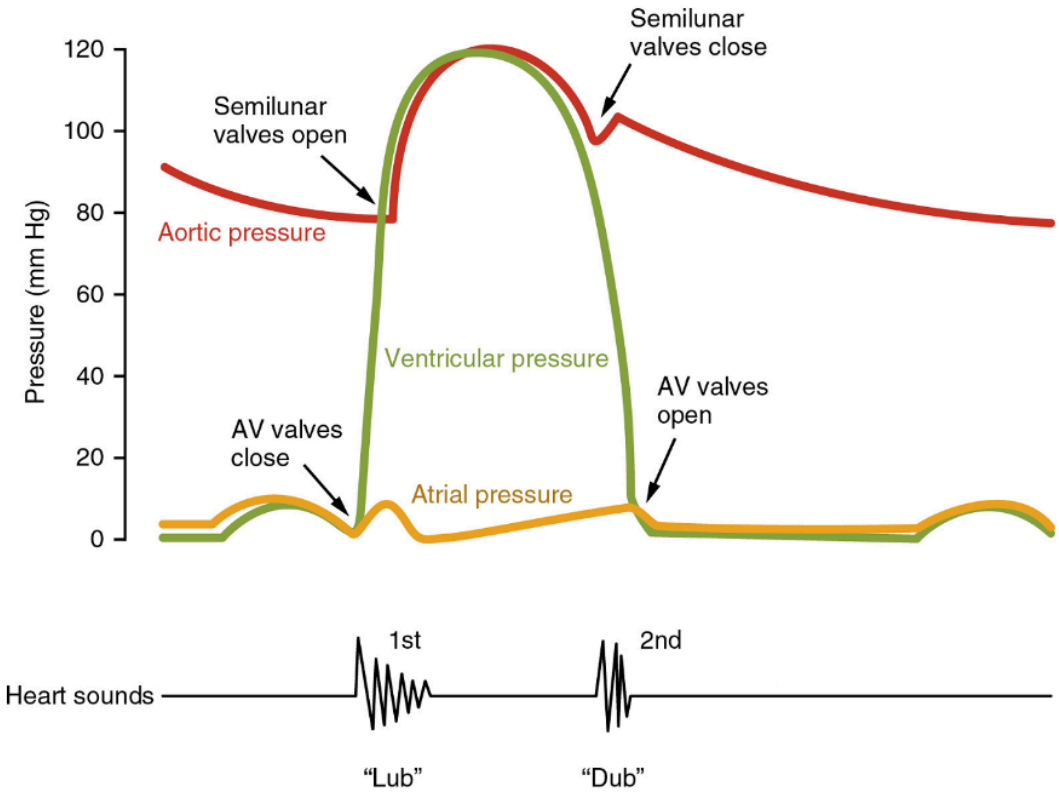


Figure 26. Heart Sounds and the Cardiac Cycle. *In this illustration, the x-axis reflects time with a recording of the heart sounds. The y-axis represents pressure.*

Cardiac output (CO) is a measurement of the amount of blood pumped by each ventricle in one minute. To calculate this value, multiply **stroke volume (SV)**, the amount of blood pumped by each ventricle, by **heart rate (HR)**, in contractions per minute (or beats per minute, bpm). It can be represented mathematically by the following equation:

$$CO = HR \times SV$$

Stroke volume (SV) is normally measured using an echocardiogram to record end diastolic volume (EDV) and end systolic volume (ESV), and calculating the difference: SV = EDV – ESV. Stroke volume can also be measured using a specialized catheter, but this is an invasive procedure and far more dangerous to the patient. A mean stroke volume for a resting 70 kg (150 lb) individual would be approximately 70 mL. There are several important variables, including size of the heart, physical and mental condition of the individual, sex, contractility, duration of contraction, preload or EDV, and afterload or resistance. Normal range for stroke volume would be 55–100 mL. An average resting heart rate is approximately 75 bpm but can range from 60–100 in some individuals.

Part 4: Cardiac Physiology

The autorhythmicity inherent in cardiac cells keeps the heart beating at a regular pace; however, the heart is regulated by and responds to outside influences as well (Figure 27). Neural and endocrine controls are vital to the regulation of cardiac function. In addition, the heart is sensitive to several environmental factors, including electrolytes.

Heart Rates: Heart rates vary considerably, not only with exercise and fitness levels, but also with age. Newborn resting heart rates may be 120 bpm. Heart rate gradually decreases until young adulthood and then gradually increases again with age.

Maximum heart rates are normally in the range of 200–220 bpm, although there are some extreme cases in which they may reach higher levels. As one ages, the ability to generate maximum rates decreases. This may be estimated by taking the maximal value of 220 bpm and subtracting the individual's age. So a 40-year-old

individual would be expected to hit a maximum rate of approximately 180, and a 60-year-old person would achieve a heart rate of 160.

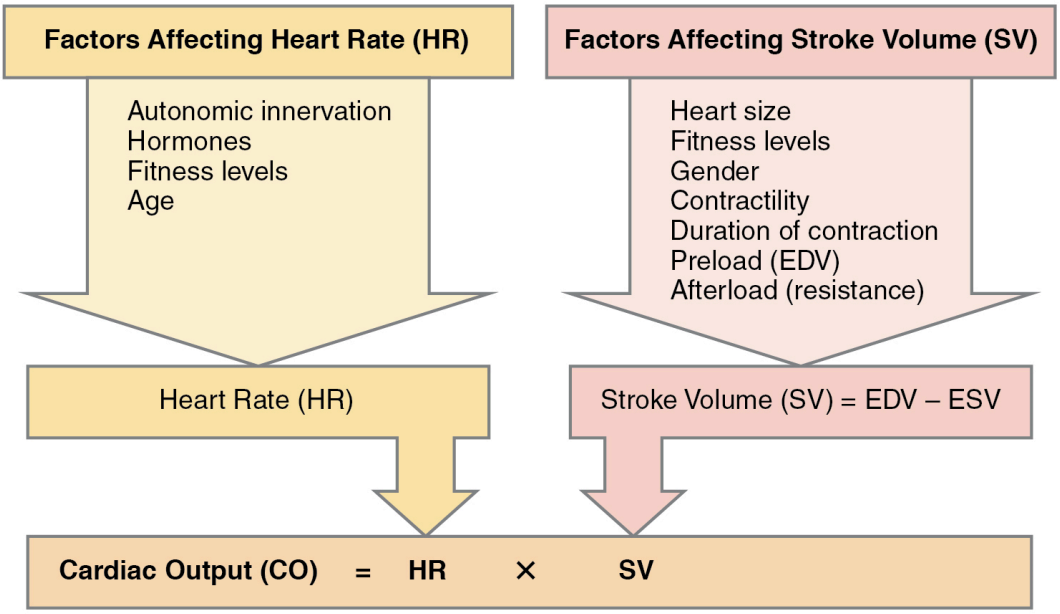


Figure 27. Major Factors Influencing Cardiac Output. *Cardiac output is influenced by heart rate and stroke volume, both of which are also variable.*

Bradycardia (resting heart rate below 60 bpm) may be caused by either inherent factors or causes external to the heart. While the condition may be inherited, typically it is acquired in older individuals. Inherent causes include abnormalities in either the sinoatrial or atrioventricular node. If the condition is serious, a pacemaker may be required. Other causes include ischemia to the heart muscle or diseases of the heart vessels or valves. External causes include metabolic disorders, pathologies of the endocrine system often involving the thyroid, electrolyte imbalances, neurological disorders including inappropriate autonomic responses, autoimmune pathologies, over-prescription of beta blocker drugs that reduce heart rate, recreational drug use, or even prolonged bed rest. Treatment relies upon establishing the underlying cause of the disorder and may necessitate supplemental oxygen.

Tachycardia (resting heart rate above 100 bpm) is not normal in a resting patient but may be detected in pregnant women or individuals experiencing extreme stress. In the latter case, it would likely be triggered by stimulation from the limbic system or disorders of the autonomic nervous system. In some cases, tachycardia may involve only the atria. Some individuals may remain asymptomatic, but when present, symptoms may include dizziness, shortness of breath, lightheadedness, rapid pulse, heart palpations, chest pain, or fainting (syncope). While tachycardia is defined as a heart rate above 100 bpm, there is considerable variation among people. Further, the normal resting heart rates of children are often above 100 bpm, but this is not considered to be tachycardia Many causes of tachycardia may be benign, but the condition may also be correlated with fever, anaemia, hypoxia, hyperthyroidism, hypersecretion of catecholamines, some cardiomyopathies, some disorders of the valves, and acute exposure to radiation. Elevated rates in an exercising or resting patient are normal and expected. Resting rate should always be taken after recovery from exercise. Treatment depends upon the underlying cause but may include medications, implantable cardioverter defibrillators, ablation, or surgery.

Correlation Between Heart Rates and Cardiac Output: Initially, physiological conditions that cause heart rate to increase also trigger an increase in stroke volume. During exercise, the rate of blood returning to the heart increases. However as the heart rate rises, there is less time spent in diastole and consequently less time for the ventricles to fill with blood. Even though there is less filling time, stroke volume will initially remain high. However, as heart rate continues to increase, stroke volume gradually decreases due to decreased filling

time. Cardiac output will initially stabilize as the increasing heart rate compensates for the decreasing stroke volume , but at very high rates, cardiac output will eventually decrease as increasing rates are no longer able to compensate for the decreasing stroke volume. Consider this phenomenon in a healthy young individual. Initially, as heart rate increases from resting to approximately 120 bpm, cardiac output will rise. As heart rate increases from 120 to 160 bpm, cardiac output remains stable, since the increase in rate is offset by decreasing ventricular filling time and, consequently, stroke volume. As heart rate continues to rise above 160 bpm, cardiac output actually decreases as stroke volume falls faster than heart rate increases. So although aerobic exercises are critical to maintain the health of the heart, individuals are cautioned to monitor their heart rate to ensure they stay within the **target heart rate range** of between 120 and 160 bpm, so cardiac output is maintained. The target heart rate is loosely defined as the range in which both the heart and lungs receive the maximum benefit from the aerobic workout and is dependent upon age.

Cardiovascular Centres: Nervous control over heart rate is centralized within the two paired cardiovascular centres of the medulla oblongata (Figure 28). The cardioacceleratory centre stimulates activity via sympathetic stimulation of the cardioacceleratory nerves, and the cardioinhibitory centre inhibits heart rate via parasympathetic stimulation as one component of the vagus nerve, cranial nerve X.

During rest, both centres provide slight stimulation to the heart, contributing to **autonomic tone**. This is a similar concept to tone in skeletal muscles. Normally, vagal stimulation predominates; left unregulated, the sinoatrial node would initiate a sinus rhythm of approximately 100 bpm.

The ventricles are more richly innervated by sympathetic fibers than parasympathetic fibers. Sympathetic stimulation causes the release of the neurotransmitter norepinephrine (NE) at the neuromuscular junction of the cardiac nerves. Norepinephrine shortens the repolarization period, thus speeding the rate of depolarization and force of contraction, which results in an increase in heart rate.

Norepinephrine binds mainly to the beta-1 receptors in the heart but there are also beta-2 receptors and norepinephrine's effect is similar on these. Some cardiac medications (for example, beta blockers) work by blocking these receptors, thereby slowing heart rate and are one possible treatment for hypertension. Overuse of these drugs may lead to bradycardia and even stoppage of the heart with chronic use.

Parasympathetic stimulation originates from the cardioinhibitory region with impulses traveling via the vagus nerve (cranial nerve X). The vagus nerve sends branches to both the sinoatrial and atrioventricular nodes, and to portions of both the atria and ventricles. Parasympathetic stimulation releases the neurotransmitter acetylcholine (ACh) at the neuromuscular junction, and ACh slows heart rate. Without any nervous stimulation, the sinoatrial node would establish a sinus rhythm of approximately 100 bpm. Since resting rates are considerably less than this, it becomes evident that parasympathetic stimulation normally slows heart rate.

Input to the Cardiovascular Centre: The cardiovascular centre receives input from a series of visceral receptors with impulses traveling through visceral sensory fibers within the vagus and sympathetic nerves via the cardiac plexus. Among these receptors are various proprioreceptors, baroreceptors, and chemoreceptors, plus stimuli from the limbic system. Collectively, these inputs normally enable the cardiovascular centres to regulate heart function precisely, a process known as **cardiac reflexes**. Increased physical activity results in increased rates of firing by various proprioreceptors located in muscles, joint capsules, and tendons. Any such increase in physical activity would logically warrant increased blood flow. The cardiac centres monitor these increased rates of firing, and suppress parasympathetic stimulation and increase sympathetic stimulation as needed in order to increase blood flow.

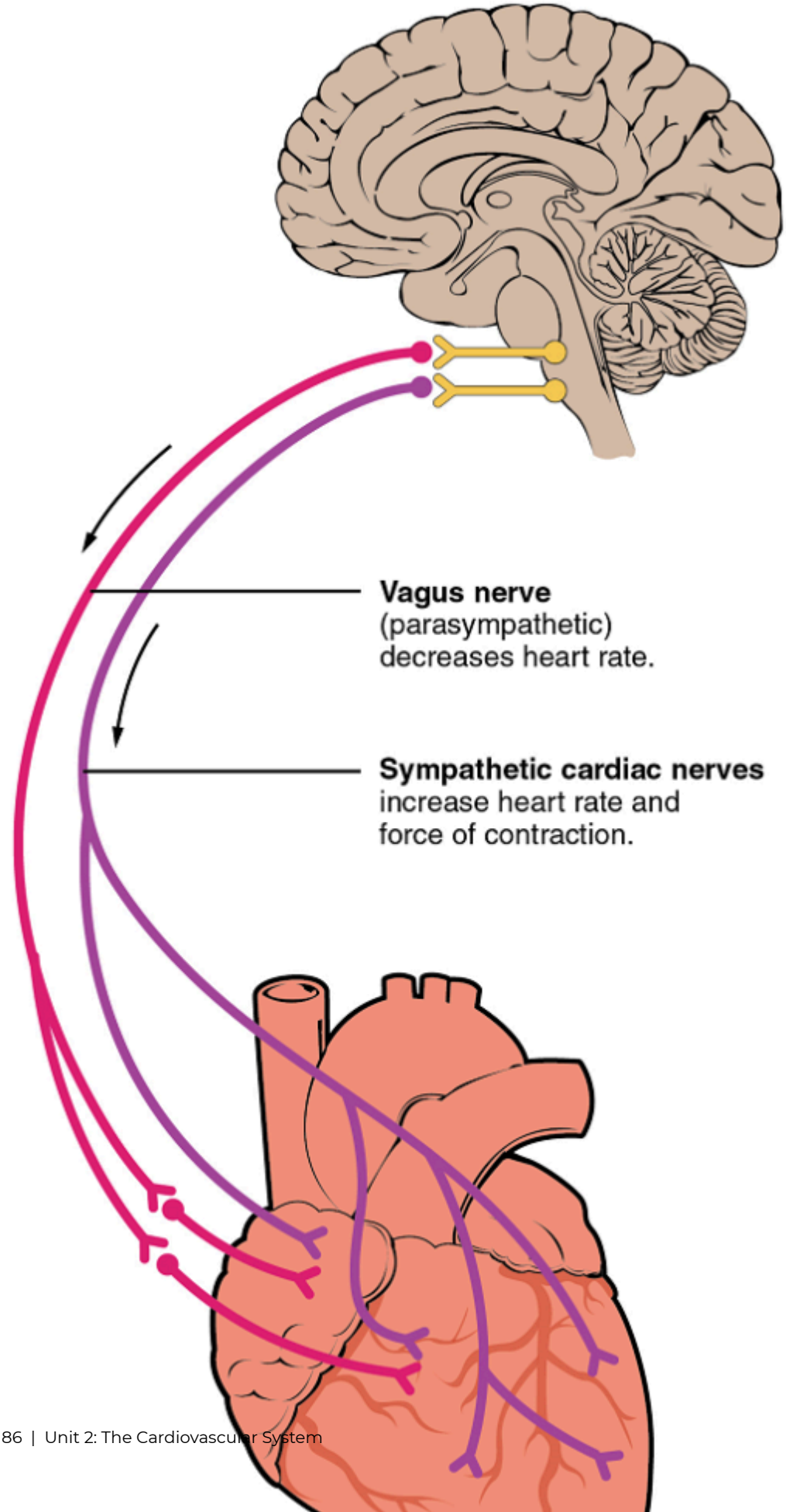


Figure 28. Autonomic Innervation of the Heart.
Cardioacceleratory and cardioinhibitory centres of the brain innervate the heart via sympathetic cardiac nerves that increase cardiac activity and part of the vagus (parasympathetic) nerve that slows cardiac activity.

Similarly, baroreceptors are stretch receptors located in the aortic sinus, carotid bodies, the venae cavae, and other locations, including pulmonary vessels and the right side of the heart itself. Rates of firing from the baroreceptors represent blood pressure, level of physical activity, and the relative distribution of blood. The cardiac centres monitor baroreceptor firing to maintain cardiac homeostasis, a mechanism called the baroreceptor reflex. With increased pressure and stretch, the rate of baroreceptor firing increases, and the cardiac centres decrease sympathetic stimulation and increase parasympathetic stimulation. As pressure and stretch decrease, the rate of baroreceptor firing decreases, and the cardiac centres increase sympathetic stimulation and decrease parasympathetic stimulation (Table 4 and Table 5).

Table 4: Cardiac response to decreasing blood flow and pressure due to decreasing cardiac output

	Baroreceptors (aorta, carotid arteries, venae cavae, and atria)	**Chemoreceptors (both central nervous system and in proximity to baroreceptors)**
Sensitive to	Decreasing stretch	Decreasing O_2 and increasing CO_2, H^+, and lactic acid
Target	Parasympathetic stimulation suppressed	Sympathetic stimulation increased
Response of heart	Increasing heart rate and increasing stroke volume	Increasing heart rate and increasing stroke volume
Overall effect	Increasing blood flow and pressure due to increasing cardiac output	Increasing blood flow and pressure due to increasing cardiac output

Table 5: Cardiac response to increasing blood flow and pressure due to increasing cardiac output

	Baroreceptors (aorta, carotid arteries, venae cavae, and atria)	**Chemoreceptors (both central nervous system and in proximity to baroreceptors)**
Sensitive to	Increasing stretch	Increasing O_2 and decreasing CO_2, H^+, and lactic acid
Target	Parasympathetic stimulation increased	Sympathetic stimulation suppressed
Response of heart	Decreasing heart rate and decreasing stroke volume	Decreasing heart rate and decreasing stroke volume
Overall effect	Decreasing blood flow and pressure due to decreasing cardiac output	Decreasing blood flow and pressure due to decreasing cardiac output

Increased metabolic byproducts associated with increased activity, such as carbon dioxide, hydrogen ions, and lactic acid, plus falling oxygen levels, are detected by a suite of chemoreceptors innervated by the glossopharyngeal and vagus nerves. These chemoreceptors provide feedback to the cardiac centres about the need for increased or decreased blood flow, based on the relative levels of these substances (Table 4 and Table 5).

Other Factors Influencing Heart Rate

Using a combination of autorhythmicity and innervation, the cardiac centres are able to provide relatively precise control over heart rate. However, there are a number of other factors that have an impact on heart rate as well, including epinephrine, norepinephrine, and thyroid hormones; levels of various ions including calcium, potassium, and sodium; body temperature; hypoxia; and pH (Table 6).

Epinephrine and Norepinephrine: The catecholamines (epinephrine and norepinephrine) secreted by the

adrenal medulla form one component of the extended fight-or-flight mechanism. The other component is sympathetic stimulation. Epinephrine and norepinephrine have similar effects. There is no parasympathetic stimulation to the adrenal medulla.

Thyroid Hormones: In general, increased levels of thyroid hormone, or thyroxin, increase both the heart rate and the force of contraction (contractility). The impact of thyroid hormone is typically of a much longer duration than that of the catecholamines.

Calcium: Calcium ion levels have great impacts upon both heart rate and contractility; as the levels of calcium ions increase, so do heart rate and contractility. Extremely high levels of calcium may induce cardiac arrest.

Table 6: Major factors affecting heart rate and force of contraction

Factor	Effect
Increased sympathetic nervous system activity	Increased heart rate and contractility; norepinephrine release
Increased parasympathetic nervous system activity	Decreased heart rate
Epinephrine/adrenaline and norepinephrine/ noradrenaline	Increased heart rate and contractility
Thyroxine	Increased heart rate and contractility
Blood Ca^{2+} concentration	High levels increase heart rate and contractility; low levels decrease heart rate and contractility
Blood K^{+} concentration	Low levels increase heart rate and decrease contractility; high levels decrease heart rate and contractility
High body temperature	Increased heart rate
Low body temperature	Decreased heart rate and contractility

Blood Vessels and Circulation

In this section, you will learn about the vascular part of the cardiovascular system, that is, the vessels that transport blood throughout the body and provide the physical site where gases, nutrients, and other substances are exchanged with body cells. When vessel functioning is reduced, blood-borne substances do not circulate effectively throughout the body. As a result, tissue injury occurs, metabolism is impaired, and the functions of every bodily system are threatened.

Part 1: Structure and Function of Blood Vessels

Blood is carried through the body via blood vessels. An artery is a blood vessel that carries blood away from the heart, where it branches into ever-smaller vessels. Eventually, the smallest arteries, vessels called arterioles, further branch into tiny capillaries, where nutrients and wastes are exchanged, and then combine with other vessels that exit capillaries to form venules, small blood vessels that carry blood to a vein, a larger blood vessel that returns blood to the heart.

Arteries and veins transport blood in two distinct circuits: the systemic circuit and the pulmonary circuit (Figure 29). Systemic arteries provide blood rich in oxygen to the body's tissues. The blood returned to the heart through systemic veins has less oxygen, since much of the oxygen carried by the arteries has been delivered to the cells. In contrast, in the pulmonary circuit, arteries carry blood low in oxygen exclusively to the lungs for gas exchange. Pulmonary veins then return freshly oxygenated blood from the lungs to the heart to be pumped back out into systemic circulation. Although arteries and veins differ structurally and functionally, they share certain features.

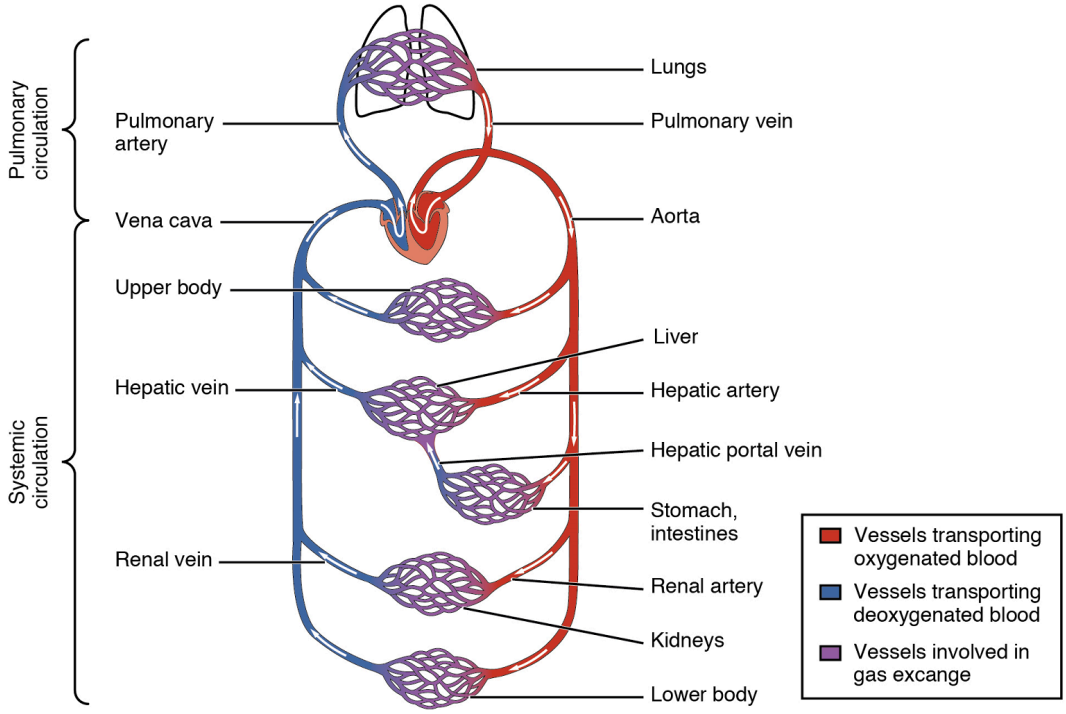


Figure 29. Cardiovascular Circulation. *The pulmonary circuit moves blood from the right side of the heart to the lungs and back to the heart. The systemic circuit moves blood from the left side of the heart to the head and body and returns it to the right side of the heart to repeat the cycle. The arrows indicate the direction of blood flow, and the colours show the relative levels of oxygen concentration.*

Shared Structures: Different types of blood vessels vary slightly in their structures, but they share the same general features. Arteries and arterioles have thicker walls than veins and venules because they are closer to the heart and receive blood that is surging at a far greater pressure (Figure 30). Each type of vessel has a **lumen**—a hollow passageway through which blood flows. Arteries have smaller lumens than veins, a characteristic that helps to maintain the pressure of blood moving through the system. Together, their thicker walls and smaller diameters give arterial lumens a more rounded appearance in cross section than the lumens of veins.

By the time blood has passed through capillaries and entered venules, the pressure initially exerted upon it by heart contractions has diminished. In other words, in comparison to arteries, venules and veins withstand a much lower pressure from the blood that flows through them. Their walls are considerably thinner and their lumens are correspondingly larger in diameter, allowing more blood to flow with less vessel resistance. In addition, many veins of the body, particularly those of the limbs, contain valves that assist the unidirectional flow of blood toward the heart. This is critical because blood flow becomes sluggish in the extremities, as a result of the lower pressure and the effects of gravity.

Both arteries and veins have the same three distinct tissue layers, called tunics (from the Latin term tunica), for the garments first worn by ancient Romans; the term tunic is also used for some modern garments. From the most interior layer to the outer, these tunics are the tunica intima, the tunica media, and the tunica externa (Figure 30 and Table 7).

Tunica Intima: The tunica intima (also called the tunica interna) is composed of epithelial and connective tissue layers. Lining the tunica intima is the specialized simple squamous epithelium called the endothelium, which is continuous throughout the entire vascular system, including the lining of the chambers of the heart. Damage to this endothelial lining and exposure of blood to the collagenous fibers beneath is one of the primary causes of clot formation. Until recently, the endothelium was viewed simply as the boundary between the blood in the lumen and the walls of the vessels. Recent studies, however, have shown that it is physiologically critical to such activities as helping to regulate capillary exchange and altering blood flow. The endothelium releases local chemicals called endothelins that can constrict the smooth muscle within the walls of the vessel to increase

blood pressure. Uncompensated overproduction of endothelins may contribute to hypertension (high blood pressure) and cardiovascular disease.

Next to the endothelium is the basement membrane, or basal lamina, that effectively binds the endothelium to the connective tissue. The basement membrane provides strength while maintaining flexibility, and it is permeable, allowing materials to pass through it. The thin outer layer of the tunica intima contains a small amount of areolar connective tissue that consists primarily of elastic fibers to provide the vessel with additional flexibility; it also contains some collagenous fibers to provide additional strength.

In larger arteries, there is also a thick, distinct layer of elastic fibers known as **the internal elastic membrane** (also called the internal elastic lamina) at the boundary with the tunica media. Like the other components of the tunica intima, the internal elastic membrane provides structure while allowing the vessel to stretch. It is permeated with small openings that allow exchange of materials between the tunics. The internal elastic membrane is not apparent in veins. In addition, many veins, particularly in the lower limbs, contain valves formed by sections of thickened endothelium that are reinforced with connective tissue, extending into the lumen.

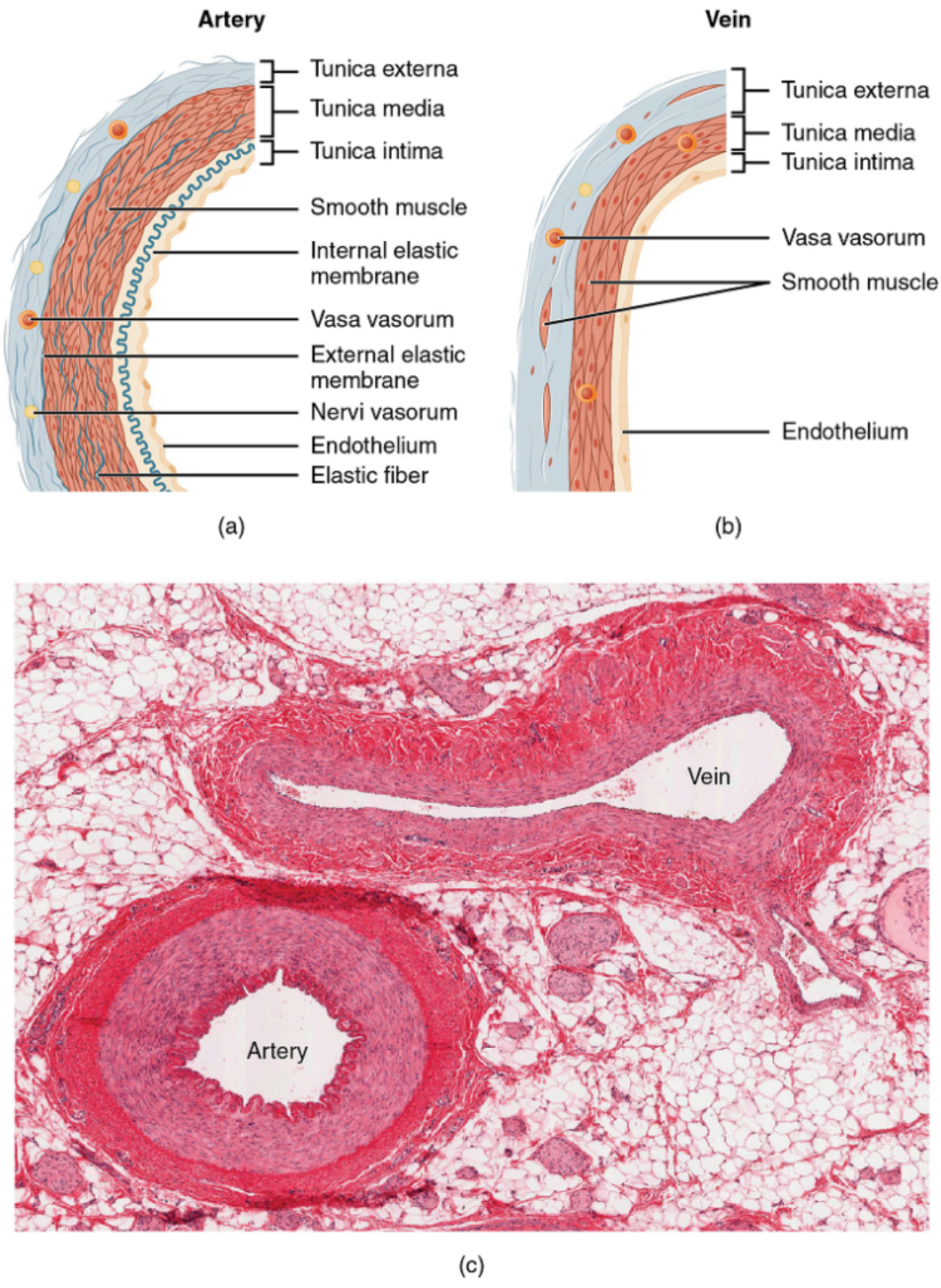


Figure 30. Structure of Blood Vessels. *(a) Arteries and (b) veins share the same general features, but the walls of arteries are much thicker because of the higher pressure of the blood that flows through them. (c) A micrograph shows the relative differences in thickness. LM × 160. (Micrograph provided by the Regents of the University of Michigan Medical School © 2012)*

Tunica Media: The **tunica media** is the substantial middle layer of the vessel wall (Figure 30). It is generally the thickest layer in arteries, and it is much thicker in arteries than it is in veins. The tunica media consists of layers of smooth muscle supported by connective tissue that is primarily made up of elastic fibers, most of which are arranged in circular sheets. Toward the outer portion of the tunic, there are also layers of longitudinal muscle. Contraction and relaxation of the circular muscles decrease and increase the diameter of the vessel lumen, respectively. Specifically, in arteries, **vasoconstriction** decreases blood flow as the smooth muscle in the walls of

the tunica media contracts, making the lumen narrower and increasing blood pressure. Similarly, **vasodilation** increases blood flow as the smooth muscle relaxes, allowing the lumen to widen and blood pressure to drop.

The smooth muscle layers of the tunica media are supported by a framework of collagenous fibers that also binds the tunica media to the inner and outer tunics. Along with the collagenous fibers are large numbers of elastic fibers that appear as wavy lines in prepared slides.

Table 7: Comparison of wall layers in arteries, veins, and capillaries

	Arteries	Veins	Capillaries
General appearance	Thick walls with small lumens	Thin walls with large lumens	Very (microscopically) thin walls and very small lumens
	Generally appear rounded	Generally appear flattened	Generally round
Tunica intima	Endothelium usually appears wavy due to constriction of smooth muscle	Endothelium appears smooth	Endothelium appears smooth
	Internal elastic membrane present in larger vessels	Internal elastic membrane absent	Internal elastic membrane absent
Tunica media	Normally the thickest layer in arteries	Normally thinner than the tunica externa	Tunica media absent
	Smooth muscle cells and elastic fibers predominate (exact proportions vary with distance from the heart)	Smooth muscle cells and collagenous fibers predominate	
	External elastic membrane present in larger vessels	External elastic membrane absent	
		Nervi vasorum and vasa vasorum present	
Tunica externa	Normally thinner than tunica media in all but the largest arteries	Normally the thickest layer in veins	Tunica externa absent
	Collagenous and elastic fibers	Collagenous and smooth fibers predominate	
	Nervi vasorum and vasa vasorum present	Nervi vasorum and vasa vasorum present	

Tunica Externa: The outer tunic, the **tunica externa** (also called the tunica adventitia), is a substantial sheath of connective tissue composed primarily of collagenous fibers. Some bands of elastic fibers are found here as well. The tunica externa in veins also contains groups of smooth muscle fibers. This is normally the thickest tunic in veins and may be thicker than the tunica media in some larger arteries.

Arteries: An **artery** is a blood vessel that conducts blood away from the heart. All arteries have relatively thick walls that can withstand the high pressure of blood ejected from the heart.

Arterioles: An **arteriole** is a very small artery that leads to a capillary. Arterioles have the same three tunics as the larger vessels, but the thickness of each is greatly diminished. The critical endothelial lining of the tunica intima is intact. The tunica media is restricted to one or two smooth muscle cell layers in thickness. The tunica externa remains but is very thin (Figure 39).

The importance of the arterioles is that they will be the primary site of both resistance and regulation of blood pressure. The precise diameter of the lumen of an arteriole at any given moment is determined by neural and chemical controls, and vasoconstriction and vasodilation in the arterioles are the primary mechanisms for distribution of blood flow.

Capillaries: A capillary is a microscopic channel that supplies blood to the tissues themselves, a process called **perfusion**. Exchange of gases and other substances occurs in the capillaries between the blood and the surrounding cells and their tissue fluid (interstitial fluid). The diameter of a capillary lumen is from 5-10 μm; the

smallest are just barely wide enough for an erythrocyte to squeeze through. Flow through capillaries is often described as microcirculation.

Unlike the walls of veins and arteries, the wall of a capillary consists of an endothelial layer surrounded by a basement membrane with occasional smooth muscle fibers. There is some variation in wall structure: in a large capillary, several endothelial cells bordering each other may line the lumen; in a small capillary, there may be only a single cell layer that wraps around to contact itself.

Venules: A venule is an extremely small vein, generally 8–100 μm in diameter. Postcapillary venules join multiple capillaries exiting from a capillary bed. Multiple venules join to form veins. The walls of venules consist of endothelium, a thin middle layer with a few muscle cells and elastic fibers, plus an outer layer of connective tissue fibers that constitute a very thin tunica externa. Venules as well as capillaries are the primary sites of emigration or diapedesis, in which the leukocytes adhere to the endothelial lining of the vessels and then squeeze through adjacent cells to enter the tissue fluid.

Veins: A vein is a blood vessel that conducts blood toward the heart. Compared to arteries, veins are thin-walled vessels with large and irregular lumens (Figure 42). Because they are low-pressure vessels, larger veins are commonly equipped with valves that promote the unidirectional flow of blood toward the heart and prevent backflow toward the capillaries caused by the inherent low blood pressure in veins as well as the pull of gravity. Table 8 compares the features of arteries and veins.

Table 8: Comparison of arteries and veins

	Arteries	Veins
Direction of blood flow	Conducts blood away from the heart	Conducts blood toward the heart
General appearance	Rounded	Irregular, often collapsed
Pressure	High	Low
Wall thickness	Thick	Thin
Relative oxygen concentration	Higher in systemic arteries; lower in pulmonary arteries	Lower in systemic veins; higher in pulmonary venis
Valves	Not present	Present most commonly in limbs and in veins inferior to the heart

Part 2: Blood Flow, Blood Pressure, and Resistance

Blood flow refers to the movement of blood through a vessel, tissue, or organ, and is usually expressed in terms of volume of blood per unit of time. It is initiated by the contraction of the ventricles of the heart. Ventricular contraction ejects blood into the major arteries, resulting in flow from regions of higher pressure to regions of lower pressure, as blood encounters smaller arteries and arterioles, then capillaries, then the venules and veins of the venous system. This section discusses a number of critical variables that contribute to blood flow throughout the body. It also discusses the factors that impede or slow blood flow, a phenomenon known as **resistance**.

As noted earlier, hydrostatic pressure is the force exerted by a fluid due to gravitational pull, usually against the wall of the container in which it is located. One form of hydrostatic pressure is **blood pressure**, the force exerted by blood upon the walls of the blood vessels or the chambers of the heart. Blood pressure may be measured in capillaries and veins, as well as the vessels of the pulmonary circulation; however, the term blood pressure without any specific descriptors typically refers to systemic arterial blood pressure—that is, the pressure of blood flowing in the arteries of the systemic circulation. In clinical practice, this pressure is measured in mm Hg and is usually obtained using the brachial artery of the arm.

Components of Arterial Blood Pressure: Arterial blood pressure in the larger vessels consists of several distinct components (Figure 43): systolic and diastolic pressures, pulse pressure, and mean arterial pressure.

Systolic and Diastolic Pressures: When systemic arterial blood pressure is measured, it is recorded as a ratio of two numbers (e.g., 120/80 is a normal adult blood pressure), expressed as systolic pressure over diastolic

pressure. The **systolic pressure** is the higher value (typically around 120 mm Hg) and reflects the arterial pressure resulting from the ejection of blood during ventricular contraction, or systole. The **diastolic pressure** is the lower value (usually about 80 mm Hg) and represents the arterial pressure of blood during ventricular relaxation, or diastole.

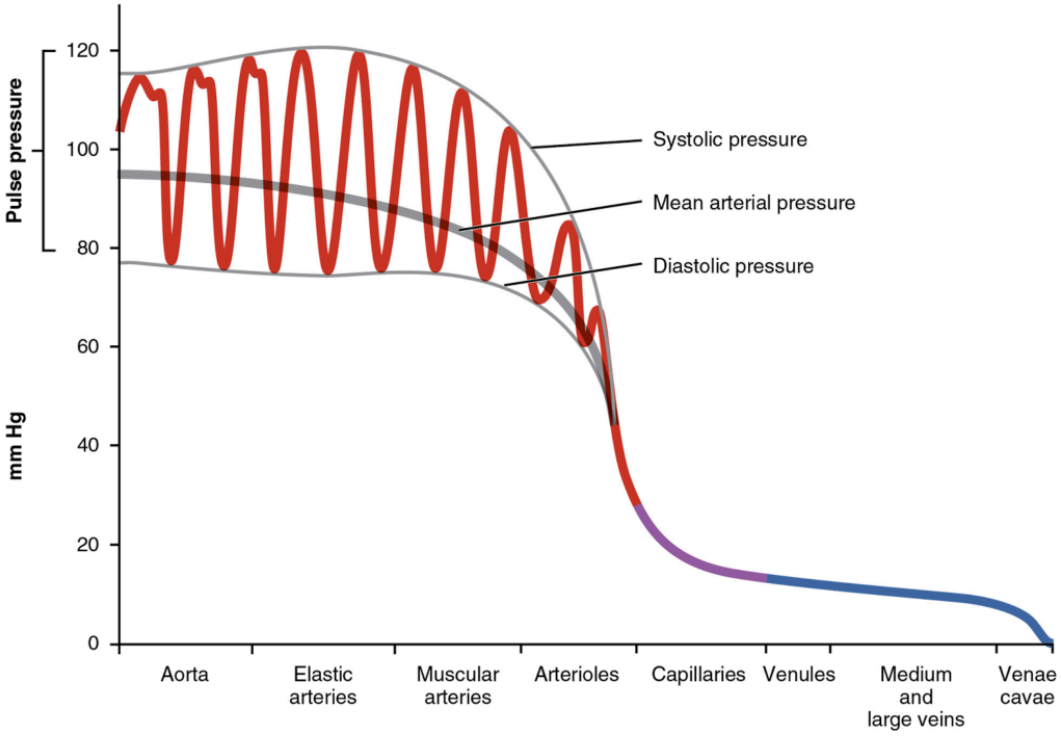


Figure 31. Systemic Blood Pressure. *The graph shows the components of blood pressure throughout the blood vessels, including systolic, diastolic, mean arterial, and pulse pressures.*

Mean Arterial Pressure: Mean arterial pressure (MAP) represents the "average" pressure of blood in the arteries, that is, the average force driving blood into vessels that serve the tissues. Mean is a statistical concept and is calculated by taking the sum of the values divided by the number of values. Although complicated to measure directly and complicated to calculate, MAP can be approximated by adding the diastolic pressure to one-third of the pulse pressure or systolic pressure minus the diastolic pressure:

MAP = diastolic BP + ((systolic-diastolic BP) / 3)

Normally, the MAP falls within the range of 70–110 mm Hg. If the value falls below 60 mm Hg for an extended time, blood pressure will not be high enough to ensure circulation to and through the tissues, which results in **ischemia**, or insufficient blood flow. A condition called hypoxia, inadequate oxygenation of tissues, commonly accompanies ischemia. The term hypoxemia refers to low levels of oxygen in systemic arterial blood.

Measurement of Blood Pressure: Blood pressure is one of the critical parameters measured on virtually every patient in every healthcare setting. The technique used today was developed more than 100 years ago by a pioneering Russian physician, Dr. Nikolai Korotkoff. Turbulent blood flow through the vessels can be heard as a soft ticking while measuring blood pressure; these sounds are known as **Korotkoff sounds**. The technique of measuring blood pressure requires the use of a **sphygmomanometer** (a blood pressure cuff attached to a measuring device) and a stethoscope. The technique is as follows:

- The clinician wraps an inflatable cuff tightly around the patient's arm at about the level of the heart.
- The clinician squeezes a rubber pump to inject air into the cuff, raising pressure around the artery and temporarily cutting off blood flow into the patient's arm.
- The clinician places the stethoscope on the patient's antecubital region and, while gradually allowing air

within the cuff to escape, listens for the Korotkoff sounds.

The first sound heard through the stethoscope—the first Korotkoff sound—indicates systolic pressure. As more air is released from the cuff, blood is able to flow freely through the brachial artery and all sounds disappear. The point at which the last sound is heard is recorded as the patient's diastolic pressure.

Pulse: After blood is ejected from the heart, elastic fibers in the arteries help maintain a high-pressure gradient as they expand to accommodate the blood, then recoil. This expansion and recoiling effect, known as the **pulse**, can be palpated manually or measured electronically. Although the effect diminishes over distance from the heart, elements of the systolic and diastolic components of the pulse are still evident down to the level of the arterioles.

Because pulse indicates heart rate, it is measured clinically to provide clues to a patient's state of health. It is recorded as beats per minute. Both the rate and the strength of the pulse are important clinically. A high or irregular pulse rate can be caused by physical activity or other temporary factors, but it may also indicate a heart condition. The pulse strength indicates the strength of ventricular contraction and cardiac output. If the pulse is strong, then systolic pressure is high. If it is weak, systolic pressure has fallen, and medical intervention may be warranted.

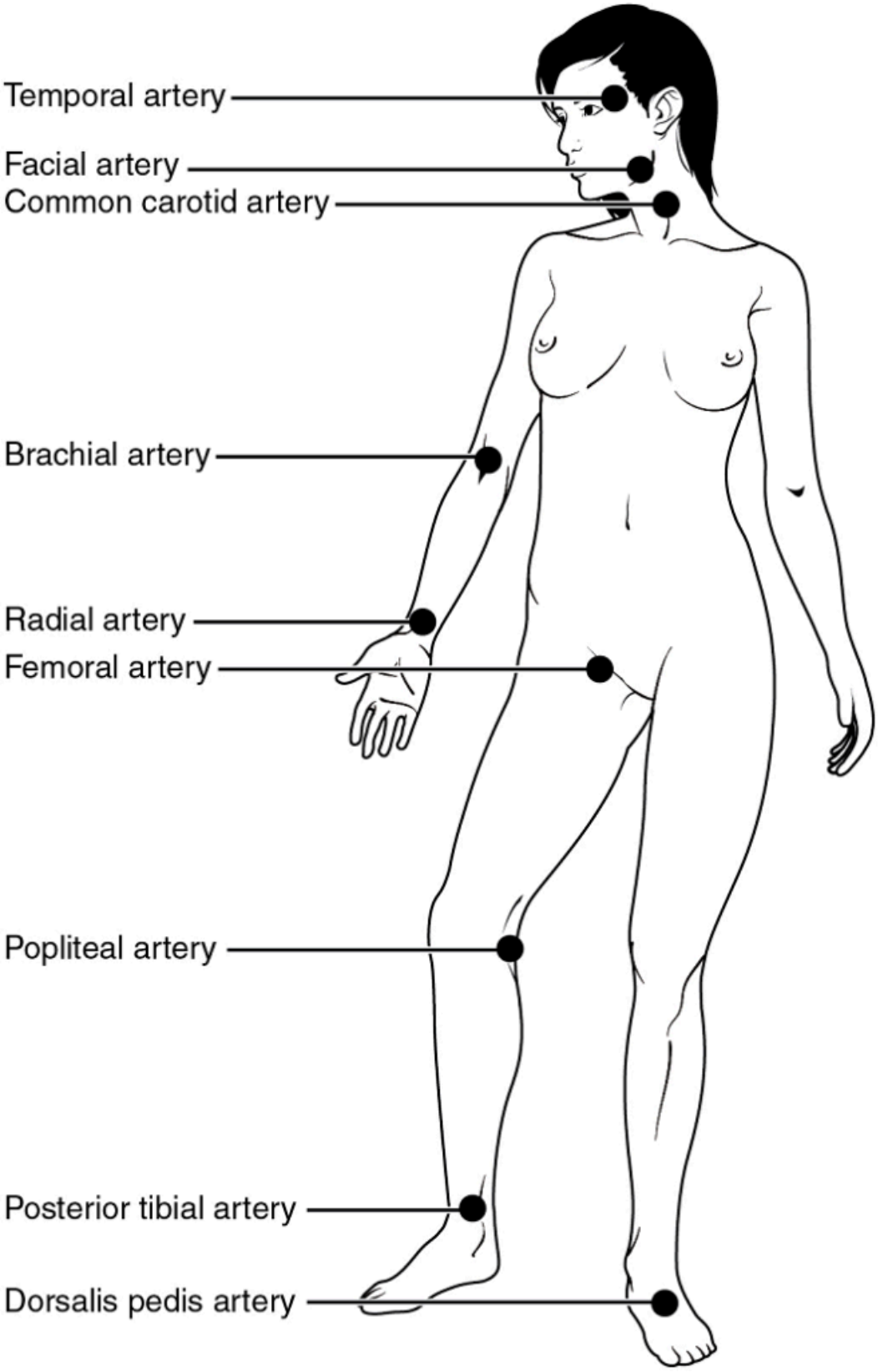


Figure 32. Pulse Sites. *The pulse is most readily measured at the radial artery, but can be measured at any of the pulse points shown.*

Pulse can be palpated manually by placing the tips of the fingers across an artery that runs close to the body

surface and pressing lightly. While this procedure is normally performed using the radial artery in the wrist or the common carotid artery in the neck, any superficial artery that can be palpated may be used (Figure 32). Common sites to find a pulse include temporal and facial arteries in the head, brachial arteries in the upper arm, femoral arteries in the thigh, popliteal arteries behind the knees, posterior tibial arteries near the medial tarsal regions, and dorsalis pedis arteries in the feet. A variety of commercial electronic devices are also available to measure pulse.

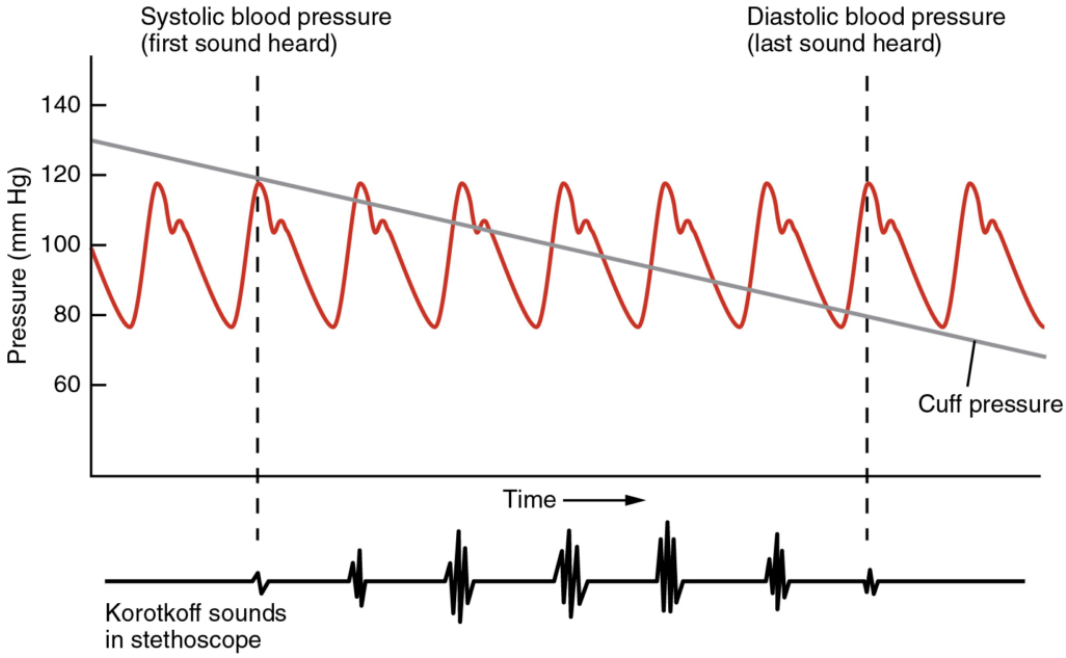


Figure 33. Blood Pressure Measurement. *When pressure in a sphygmomanometer cuff is released, a clinician can hear the Korotkoff sounds. In this graph, a blood pressure tracing is aligned to a measurement of systolic and diastolic pressures.*

Variables Affecting Blood Flow and Blood Pressure: Five variables influence blood flow and blood pressure:

- Cardiac output
- Compliance
- Volume of the blood
- Viscosity of the blood
- Blood vessel length and diameter

Recall that blood moves from higher pressure to lower pressure. It is pumped from the heart into the arteries at high pressure. If you increase pressure in the arteries (afterload), and cardiac function does not compensate, blood flow will actually decrease. In the venous system, the opposite relationship is true. Increased pressure in the veins does not decrease flow as it does in arteries, but actually increases flow. Since pressure in the veins is normally relatively low, for blood to flow back into the heart, the pressure in the atria during atrial diastole must be even lower. It normally approaches zero, except when the atria contract (Figure 33).

Cardiac Output: Cardiac output is the measurement of blood flow from the heart through the ventricles, and is usually measured in liters per minute. Any factor that causes cardiac output to increase, by elevating heart rate or stroke volume or both, will elevate blood pressure and promote blood flow. These factors include sympathetic stimulation, the catecholamines epinephrine and norepinephrine, thyroid hormones, and increased calcium ion levels. Conversely, any factor that decreases cardiac output, by decreasing heart rate or stroke volume or both, will decrease arterial pressure and blood flow. These factors include parasympathetic stimulation, elevated or decreased potassium ion levels, decreased calcium levels, anoxia, and acidosis.

Compliance: Compliance is the ability of any compartment to expand to accommodate increased content. A metal pipe, for example, is not compliant, whereas a balloon is. The greater the compliance of an artery, the more effectively it is able to expand to accommodate surges in blood flow without increased resistance or blood pressure. Veins are more compliant than arteries and can expand to hold more blood. When vascular disease causes stiffening of arteries, compliance is reduced and resistance to blood flow is increased. The result is more turbulence, higher pressure within the vessel, and reduced blood flow. This increases the work of the heart.

Blood Volume: The relationship between blood volume, blood pressure, and blood flow is intuitively obvious. Water may merely trickle along a creek bed in a dry season, but rush quickly and under great pressure after a heavy rain. Similarly, as blood volume decreases, pressure and flow decrease. As blood volume increases, pressure and flow increase.

Blood Viscosity: Viscosity is the thickness of fluids that affects their ability to flow. Clean water, for example, is less viscous than mud. The viscosity of blood is directly proportional to resistance and inversely proportional to flow; therefore, any condition that causes viscosity to increase will also increase resistance (and therefore blood pressure) and decrease flow. For example, imagine sipping milk, then a milkshake, through the same size straw. You experience more resistance and therefore less flow from the milkshake. Conversely, any condition that causes viscosity to decrease (such as when the milkshake melts) will decrease resistance and increase flow.

Normally the viscosity of blood does not change over short periods of time. The two primary determinants of blood viscosity are the formed elements and plasma proteins. Since the vast majority of formed elements are erythrocytes, any condition affecting erythropoiesis, such as polycythemia or anaemia, can alter viscosity. Viscosity generally increases with increasing numbers of formed elements relative to the amount of plasma. If the concentration of proteins in the plasma is increased, this would also increase viscosity. Since most plasma proteins are produced by the liver, any condition affecting liver function can also change the viscosity and therefore affect blood flow. Liver abnormalities include hepatitis, cirrhosis, alcohol damage, and drug toxicities. While leukocytes and platelets are normally a small component of the formed elements, there are some rare conditions in which there is such a great overproduction of these that viscosity increases.

Vessel Length and Diameter: The length of a vessel is directly proportional to its resistance: the longer the vessel, the greater the resistance and the lower the flow. As with blood volume, this makes intuitive sense, since the increased surface area of the vessel will impede the flow of blood. Likewise, if the vessel is shortened, the resistance will decrease and flow will increase.

In contrast to length, the diameter of blood vessels changes throughout the body, according to the type of vessel, as we discussed earlier. The diameter of any given vessel may also change frequently throughout the day in response to neural and chemical signals that trigger vasodilation and vasoconstriction. The **vascular tone** of the vessel is the contractile state of the smooth muscle and the primary determinant of diameter, and thus of resistance and flow. The effect of vessel diameter on resistance is inverse: Given the same volume of blood, an increased diameter means there is less blood contacting the vessel wall, thus lower friction and lower resistance, subsequently increasing flow. A decreased diameter means more of the blood contacts the vessel wall, and resistance increases, subsequently decreasing flow.

Vasodilation and vasoconstriction of arterioles play more significant roles in regulating blood pressure than do the vasodilation and vasoconstriction of other vessels.

Venous System: The pumping action of the heart propels the blood into the arteries, from an area of higher pressure toward an area of lower pressure. If blood is to flow from the veins back into the heart, the pressure in the veins must be greater than the pressure in the atria of the heart. Two factors help maintain this pressure gradient between the veins and the heart. First, the pressure in the atria during diastole is very low, often approaching zero when the atria are relaxed (atrial diastole). Second, two physiologic “pumps” increase pressure in the venous system. The use of the term “pump” implies a physical device that speeds flow. These physiological pumps are less obvious.

Skeletal Muscle Pump: In many body regions, the pressure within the veins can be increased by the contraction of the surrounding skeletal muscle. This mechanism, known as the **skeletal muscle pump** (Figure

34), helps the lower-pressure veins counteract the force of gravity, increasing pressure to move blood back to the heart. As leg muscles contract, for example during walking or running, they exert pressure on nearby veins with their numerous one-way valves. This increased pressure causes blood to flow upward, opening valves superior to the contracting muscles so blood flows through. Simultaneously, valves inferior to the contracting muscles close; thus, blood should not seep back downward toward the feet. Military recruits are trained to flex their legs slightly while standing at attention for prolonged periods. Failure to do so may allow blood to pool in the lower limbs rather than returning to the heart. Consequently, the brain will not receive enough oxygenated blood, and the individual may lose consciousness.

Respiratory Pump: The respiratory pump aids blood flow through the veins of the thorax and abdomen. During inhalation, the volume of the thorax increases, largely through the contraction of the diaphragm, which moves downward and compresses the abdominal cavity. The elevation of the chest caused by the contraction of the external intercostal muscles also contributes to the increased volume of the thorax. The volume increase causes air pressure within the thorax to decrease, allowing us to inhale. Additionally, as air pressure within the thorax drops, blood pressure in the thoracic veins also decreases, falling below the pressure in the abdominal veins. This causes blood to flow along its pressure gradient from veins outside the thorax, where pressure is higher, into the thoracic region, where pressure is now lower. This in turn promotes the return of blood from the thoracic veins to the atria. During exhalation, when air pressure increases within the thoracic cavity, pressure in the thoracic veins increases, speeding blood flow into the heart while valves in the veins prevent blood from flowing backward from the thoracic and abdominal veins. Also notice that, as blood moves from venules to veins, the average blood pressure drops.

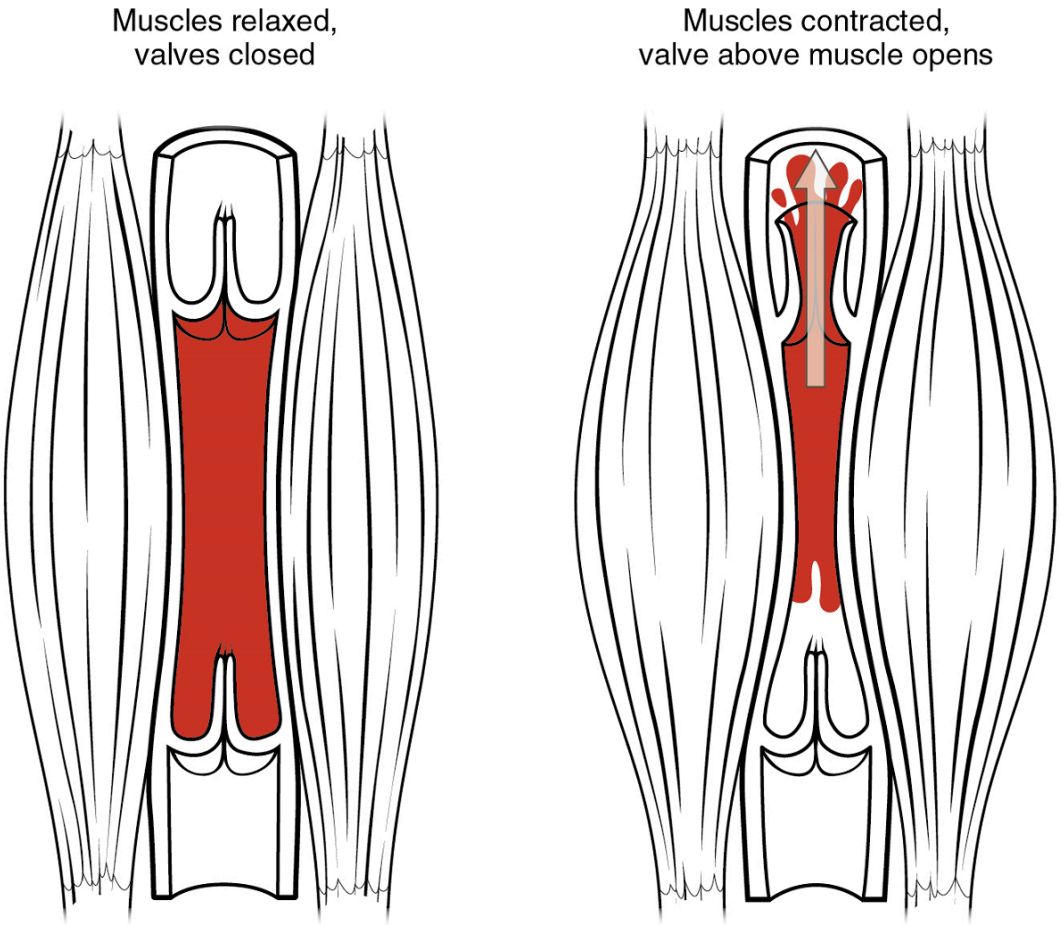


Figure 34. Skeletal Muscle Pump. *The contraction of skeletal muscles surrounding a vein compresses the blood and increases the pressure in that area. This action forces blood closer to the heart where venous pressure is lower. Note the importance of the one-way valves to assure that blood flows only in the proper direction.*

Part 3: Capillary Exchange

The primary purpose of the cardiovascular system is to circulate gases, nutrients, wastes, and other substances to and from the cells of the body. Small molecules, such as gases, lipids, and lipid-soluble molecules, can diffuse directly through the membranes of the endothelial cells of the capillary wall. Glucose, amino acids, and ions—including sodium, potassium, calcium, and chloride—use transporters to move through specific channels in the membrane by facilitated diffusion. Glucose, ions, and larger molecules may also leave the blood through intercellular clefts. Larger molecules can pass through the pores of fenestrated capillaries, and even large plasma proteins can pass through the great gaps in the sinusoids. Some large proteins in blood plasma can move into and out of the endothelial cells packaged within vesicles by endocytosis and exocytosis. Water moves by osmosis.

Part 4: Homeostatic Regulation of the Vascular System

To maintain homeostasis in the cardiovascular system and provide adequate blood to the tissues, blood flow must be redirected continually to the tissues as they become more active. In a very real sense, the cardiovascular system engages in resource allocation, because there is not enough blood flow to distribute blood equally to all tissues simultaneously. For example, when an individual is exercising, more blood will be directed to skeletal muscles, the heart, and the lungs. Following a meal, more blood is directed to the digestive system. Only the brain receives a more or less constant supply of blood whether you are active, resting, thinking, or engaged in any other activity.

Table 9 provides the distribution of systemic blood at rest and during exercise. Although most of the data appears logical, the values for the distribution of blood to the integument may seem surprising. During exercise, the body distributes more blood to the body surface where it can dissipate the excess heat generated by increased activity into the environment. Three homeostatic mechanisms ensure adequate blood flow, blood pressure, distribution, and ultimately perfusion: neural, endocrine, and autoregulatory mechanisms (Figure 35).

Table 9: Systemic blood flow during rest, mild exercise, and maximal exercise in a healthy young individual

Organ	Resting (mL/min)	Mild exercise (mL/min)	Maximal exercise (mL/min)
Skeletal muscle	1200	4500	12,500
Heart	250	350	750
Brain	750	750	750
Integument	500	1500	1900
Kidney	1100	900	600
Gastrointestinal	1400	1100	600
Others (e.g., liver, spleen)	600	400	400
Total	5800	9500	17,500

Neural Regulation: The nervous system plays a critical role in the regulation of vascular homeostasis. The primary regulatory sites include the cardiovascular centres in the brain that control both cardiac and vascular functions. In addition, more generalized neural responses from the limbic system and the autonomic nervous system are factors.

The Cardiovascular Centres in the Brain: Neurological regulation of blood pressure and flow depends on the cardiovascular centres located in the medulla oblongata. This cluster of neurons responds to changes in blood pressure as well as blood concentrations of oxygen, carbon dioxide, and hydrogen ions. The cardiovascular centre contains three distinct components:

- The cardioacceleratory centre stimulates cardiac function by regulating heart rate and stroke volume via sympathetic stimulation from the cardiac accelerator nerve.
- The cardioinhibitory centre slows cardiac function by decreasing heart rate via parasympathetic stimulation

from the vagus nerve.

- The vasomotor centre controls vessel tone or contraction of the smooth muscle in the tunica media. Changes in diameter affect peripheral resistance, pressure, and flow, which affect cardiac output. The majority of these neurons act via the release of the neurotransmitter norepinephrine from sympathetic neurons.

Although each centre functions independently, they are not anatomically distinct.

There is also a small population of neurons that control vasodilation in the vessels of the brain and skeletal muscles by relaxing the smooth muscle fibers in the vessel tunics. Many of these are cholinergic neurons, that is, they release acetylcholine, which in turn stimulates the vessels' endothelial cells to release nitric oxide (NO), which causes vasodilation. Others release norepinephrine that binds to β_2 receptors. A few neurons release NO directly as a neurotransmitter.

Baroreceptor Reflexes: Baroreceptors are specialized stretch receptors located within thin areas of blood vessels and heart chambers that respond to the degree of stretch caused by the presence of blood. They send impulses to the cardiovascular centres to regulate blood pressure. Vascular baroreceptors are found primarily in sinuses (small cavities) within the aorta and carotid arteries: The **aortic sinuses** are found in the walls of the ascending aorta just superior to the aortic valve, whereas the **carotid sinuses** are in the base of the internal carotid arteries. There are also low-pressure baroreceptors located in the walls of the venae cavae and right atrium.

When blood pressure increases, the baroreceptors are stretched more tightly and initiate action potentials at a higher rate. At lower blood pressures, the degree of stretch is lower and the rate of firing is slower. When the cardiovascular centres in the medulla oblongata receives this input, they triggers a reflex that maintains homeostasis (Figure 36):

- When blood pressure rises too high, the baroreceptors fire at a higher rate and trigger parasympathetic stimulation of the heart. As a result, cardiac output falls. Sympathetic stimulation of the peripheral arterioles will also decrease, resulting in vasodilation. Combined, these activities cause blood pressure to fall.
- When blood pressure drops too low, the rate of baroreceptor firing decreases. This will trigger an increase in sympathetic stimulation of the heart, causing cardiac output to increase. It will also trigger sympathetic stimulation of the peripheral vessels, resulting in vasoconstriction. Combined, these activities cause blood pressure to rise.

The baroreceptors in the venae cavae and right atrium monitor blood pressure as the blood returns to the heart from the systemic circulation. Normally, blood flow into the aorta is the same as blood flow back into the right atrium. If blood is returning to the right atrium more rapidly than it is being ejected from the left ventricle, the atrial receptors will stimulate the cardiovascular centres to increase sympathetic firing and increase cardiac output until homeostasis is achieved. The opposite is also true. This mechanism is referred to as the atrial reflex.

Chemoreceptor Reflexes: In addition to the baroreceptors are chemoreceptors that monitor levels of oxygen, carbon dioxide, and hydrogen ions (pH), and thereby contribute to vascular homeostasis. Chemoreceptors monitoring the blood are located in close proximity to the baroreceptors in the aortic and carotid sinuses. They signal the cardiovascular centres as well as the respiratory centres in the medulla oblongata.

Since tissues consume oxygen and produce carbon dioxide and acids as waste products, when the body is more active, oxygen levels fall and carbon dioxide levels rise as cells undergo cellular respiration to meet the energy needs of activities. This causes more hydrogen ions to be produced, causing the blood pH to drop. When the body is resting, oxygen levels are higher, carbon dioxide levels are lower, more hydrogen is bound, and pH rises.

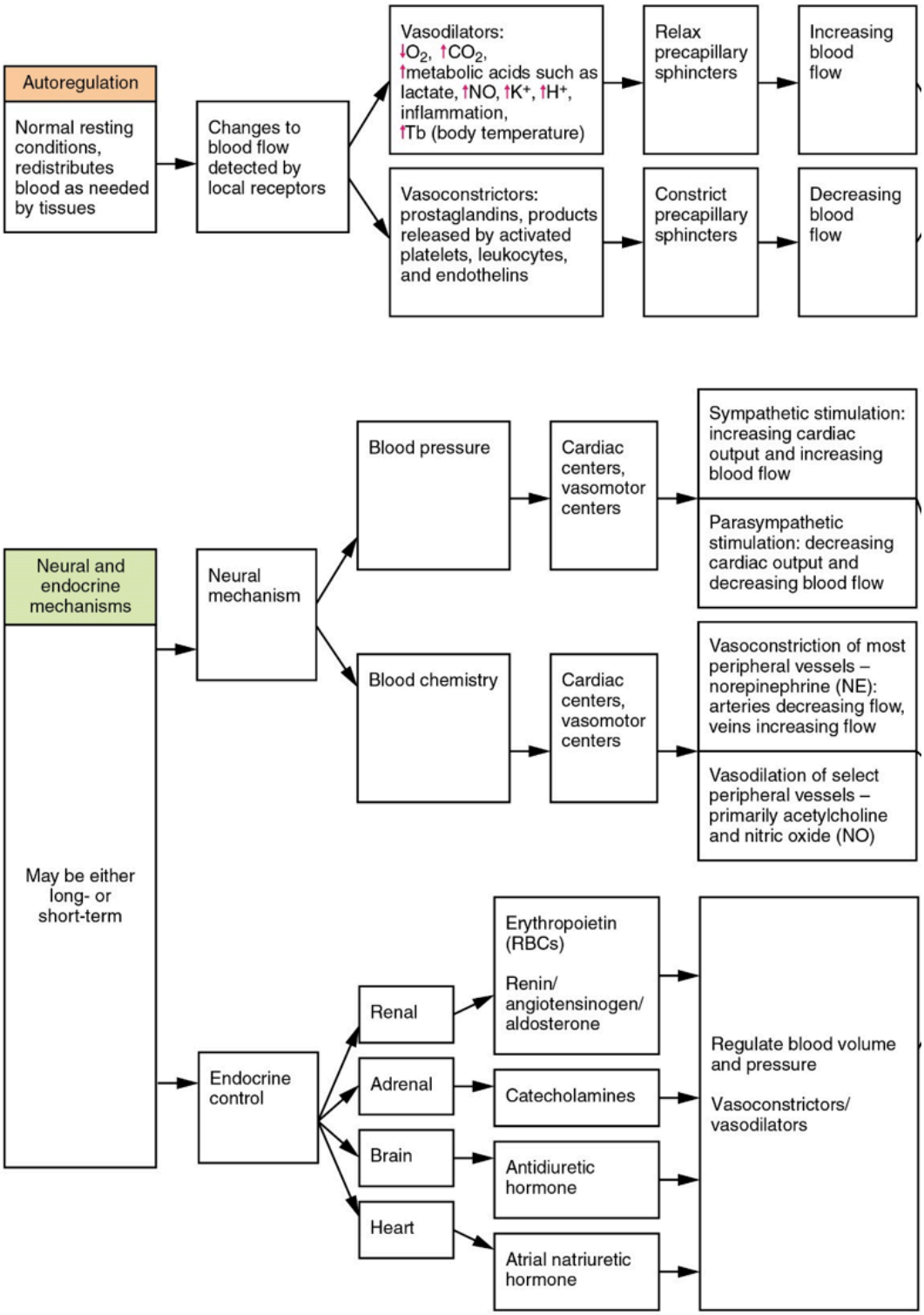


Figure 35. Summary of Factors Maintaining Vascular Homeostasis. *Adequate blood flow, blood pressure, distribution, and perfusion involve autoregulatory, neural, and endocrine mechanisms.*

The chemoreceptors respond to increasing carbon dioxide and hydrogen ion levels (falling pH) by stimulating the cardioacceleratory and vasomotor centres, increasing cardiac output and constricting peripheral vessels. The cardioinhibitory centre is suppressed. With falling carbon dioxide and hydrogen ion levels (increasing pH), the cardioinhibitory centre is stimulated, and the cardioacceleratory and vasomotor centres are suppressed,

decreasing cardiac output and causing peripheral vasodilation. In order to maintain adequate supplies of oxygen to the cells and remove waste products such as carbon dioxide, it is essential that the respiratory system respond to changing metabolic demands. In turn, the cardiovascular system will transport these gases to the lungs for exchange, again in accordance with metabolic demands. This interrelationship of cardiovascular and respiratory control cannot be overemphasized.

Other neural mechanisms can also have affect cardiovascular function. These include the limbic system that links physiological responses to psychological stimuli, as well as generalized sympathetic and parasympathetic stimulation.

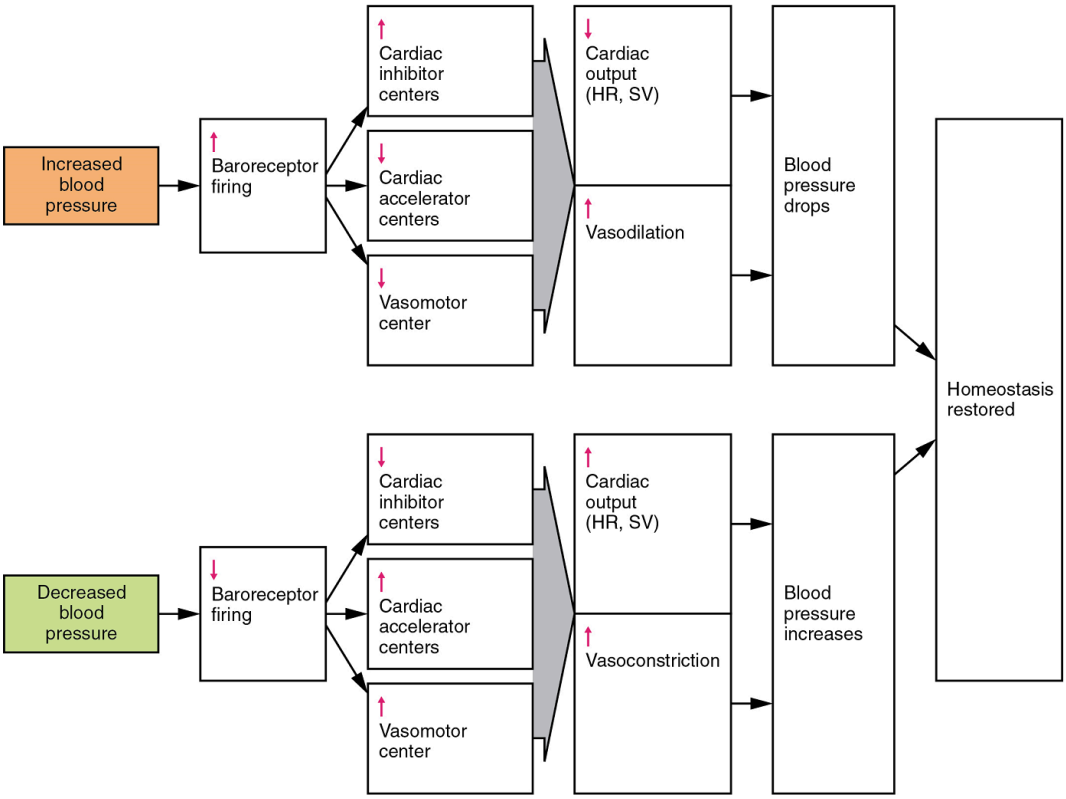


Figure 36. Baroreceptor Reflexes for Maintaining Vascular Homeostasis. *Increased blood pressure results in increased rates of baroreceptor firing, whereas decreased blood pressure results in slower rates of fire, both initiating the homeostatic mechanism to restore blood pressure.*

Endocrine Regulation: Endocrine control over the cardiovascular system involves the catecholamines, epinephrine and norepinephrine, as well as several hormones that interact with the kidneys in the regulation of blood volume.

Epinephrine and Norepinephrine: The catecholamines epinephrine and norepinephrine are released by the adrenal medulla, and enhance and extend the body's sympathetic or "fight-or-flight" response (Figure 37). They increase heart rate and force of contraction, while temporarily constricting blood vessels to organs not essential for flight-or-fight responses and redirecting blood flow to the liver, muscles, and heart.

Antidiuretic Hormone: Antidiuretic hormone (ADH), also known as vasopressin, is secreted by the cells in the hypothalamus and transported via the hypothalamic-hypophyseal tracts to the posterior pituitary where it is stored until released upon nervous stimulation. The primary trigger prompting the hypothalamus to release antidiuretic hormone is increasing osmolarity of tissue fluid, usually in response to significant loss of blood volume (Figure 38). ADH signals its target cells in the kidneys to reabsorb more water, thus preventing the loss of additional fluid in the urine. This will increase overall fluid levels and help restore blood volume and pressure. In addition, antidiuretic hormone constricts peripheral vessels.

Renin-Angiotensin-Aldosterone Mechanism: The renin-angiotensin-aldosterone mechanism has a major effect upon the cardiovascular system (Figure 37). Renin is an enzyme, although because of its importance

in the renin-angiotensin-aldosterone pathway, some sources identify it as a hormone. Specialized cells in the kidneys found in the juxtaglomerular apparatus respond to decreased blood flow by secreting renin into the blood. Renin converts the plasma protein angiotensinogen, which is produced by the liver, into its active form—angiotensin I. Angiotensin I circulates in the blood and is then converted into angiotensin II in the lungs. This reaction is catalyzed by the enzyme angiotensin-converting enzyme (ACE).

Angiotensin II is a powerful vasoconstrictor, greatly increasing blood pressure. It also stimulates the release of antiduiretic hormone and aldosterone, a hormone produced by the adrenal cortex. Aldosterone increases the reabsorption of sodium into the blood by the kidneys. Since water follows sodium, this increases the reabsorption of water. This in turn increases blood volume, raising blood pressure. Angiotensin II also stimulates the thirst centre in the hypothalamus, so an individual will likely consume more fluids, again increasing blood volume and pressure.

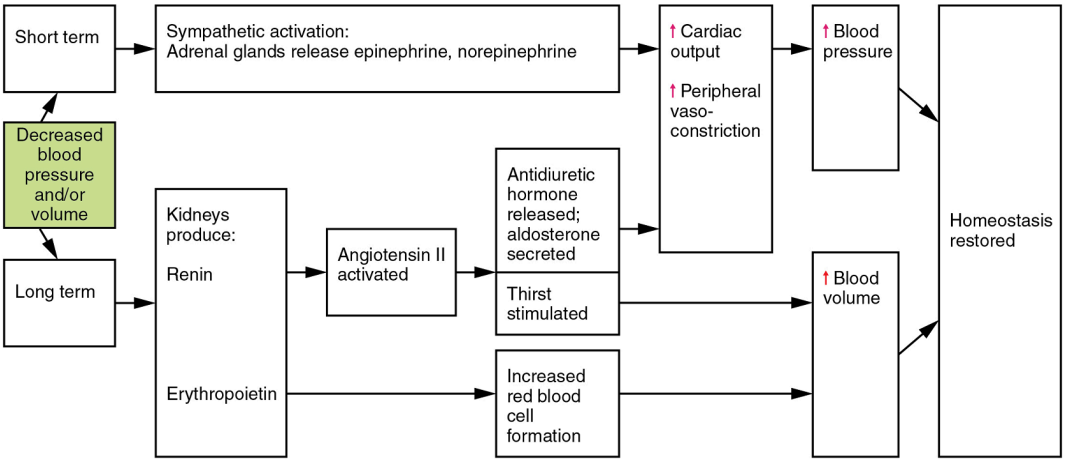


Figure 37. Hormones Involved in Renal Control of Blood Pressure. *In the renin-angiotensin-aldosterone mechanism, increasing angiotensin II will stimulate the production of antidiuretic hormone and aldosterone. In addition to renin, the kidneys produce erythropoietin, which stimulates the production of red blood cells, further increasing blood volume.*

Erythropoietin: Erythropoietin (EPO) is released by the kidneys when blood flow and/or oxygen levels decrease. Erythropoietin stimulates the production of erythrocytes within the bone marrow. Erythrocytes are the major formed element of the blood and may contribute 40% or more to blood volume, a significant factor of viscosity, resistance, pressure, and flow. In addition, erythropoietin is a vasoconstrictor. Overproduction of erythropoietin or excessive intake of synthetic erythropoietin, often to enhance athletic performance, will increase viscosity, resistance, and pressure, and decrease flow in addition to its contribution as a vasoconstrictor.

Autoregulation of Perfusion: Autoregulation mechanisms require neither specialized nervous stimulation nor endocrine control. Rather, these are local, self-regulatory mechanisms that allow each region of tissue to adjust its blood flow, and thus its perfusion. These local mechanisms include chemical signals and myogenic controls.

Chemical Signals Involved in Autoregulation: Chemical signals work at the level of the precapillary sphincters to trigger either constriction or relaxation. Opening a precapillary sphincter allows blood to flow into that particular capillary, whereas constricting a precapillary sphincter temporarily shuts off blood flow to that region. The factors involved in regulating the precapillary sphincters include the following:

- Opening of the sphincter is triggered in response to decreased oxygen concentrations; increased carbon dioxide concentrations; increasing levels of lactic acid or other byproducts of cellular metabolism; increasing concentrations of potassium ions or hydrogen ions (falling pH); inflammatory chemicals such as histamines; and increased body temperature. These conditions in turn stimulate the release of NO, a powerful vasodilator, from endothelial cells.

- Contraction of the precapillary sphincter is triggered by the opposite levels of the regulators, which prompt the release of endothelins, powerful vasoconstricting peptides secreted by endothelial cells. Platelet secretions and certain prostaglandins may also trigger constriction.

Again, these factors alter tissue perfusion via their effects on the precapillary sphincter mechanism, which regulates blood flow to capillaries. Since the amount of blood is limited, not all capillaries can fill at once, so blood flow is allocated based upon the needs and metabolic state of the tissues as reflected in these parameters. Bear in mind, however, that dilation and constriction of the arterioles feeding the capillary beds is the primary control mechanism.

The Myogenic Response: The myogenic response is a reaction to the stretching of the smooth muscle in the walls of arterioles as changes in blood flow occur through the vessel. This may be viewed as a largely protective function against dramatic fluctuations in blood pressure and blood flow to maintain homeostasis. If perfusion of an organ is too low (ischemia), the tissue will experience low levels of oxygen (hypoxia). In contrast, excessive perfusion could damage the organ's smaller and more fragile vessels. The myogenic response is a localized process that serves to stabilize blood flow in the capillary network that follows that arteriole. When blood flow is low, the vessel's smooth muscle will be only minimally stretched. In response, it relaxes, allowing the vessel to dilate and thereby increase the movement of blood into the tissue. When blood flow is too high, the smooth muscle will contract in response to the increased stretch, prompting vasoconstriction that reduces blood flow.

Part 5: Circulatory Pathways

Virtually every cell, tissue, organ, and system in the body is impacted by the circulatory system. This includes the generalized and more specialized functions of transport of materials, capillary exchange, maintaining health by transporting leukocytes and various immunoglobulins (antibodies), hemostasis, regulation of body temperature, and helping to maintain acid-base balance. In addition to these shared functions, many systems enjoy a unique relationship with the circulatory system (Figure 39).

As you learn about the vessels of the systemic and pulmonary circuits, notice that many arteries and veins share the same names, parallel one another throughout the body, and are very similar on the right and left sides of the body. For example, you will find a pair of femoral arteries and a pair of femoral veins, with one vessel on each side of the body. In contrast, some vessels closer to the midline of the body, such as the aorta, are unique. Another phenomenon that can make the study of vessels challenging is that names of vessels can change with location. Like a street that changes name as it passes through an intersection, an artery or vein can change names as it passes an anatomical landmark. For example, the left subclavian artery becomes the axillary artery as it passes through the body wall and into the axillary region, and then becomes the brachial artery as it flows from the axillary region into the upper arm (or brachium).

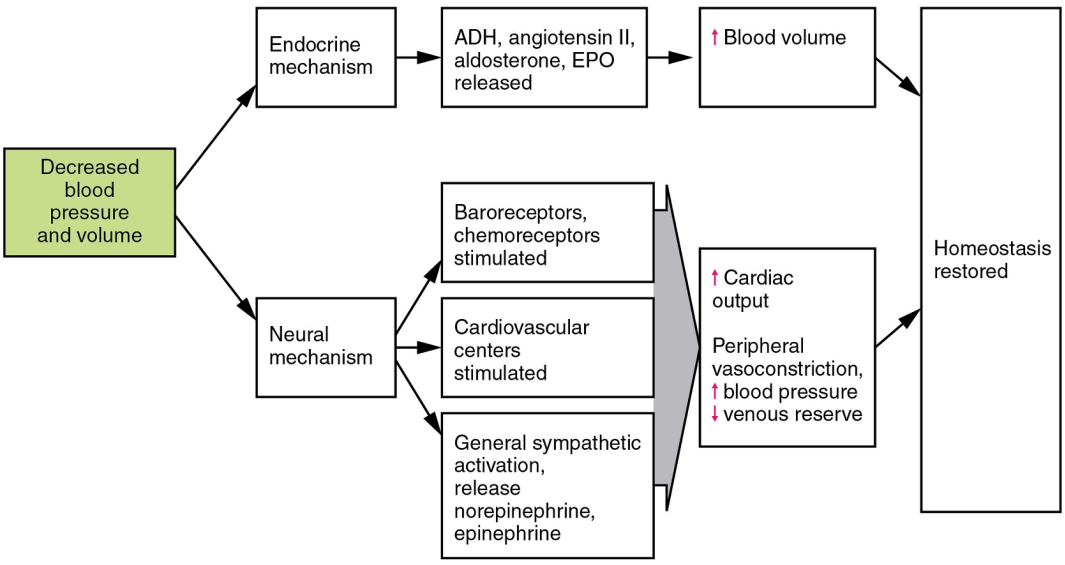


Figure 38. Homeostatic Responses to Loss of Blood Volume

Pulmonary Circulation: Recall that blood returning from the systemic circuit enters the right atrium (Figure 40) via the **superior and inferior venae cavae** and the **coronary sinus**, which drains the blood supply of the heart muscle. These vessels will be described more fully later in this section. This blood is relatively low in oxygen and relatively high in carbon dioxide, since much of the oxygen has been extracted for use by the tissues and the waste gas carbon dioxide was picked up to be transported to the lungs for elimination. From the right atrium, blood moves into the right ventricle, which pumps it to the lungs for gas exchange. This system of vessels is referred to as the **pulmonary circuit**.

The single vessel exiting the right ventricle is the **pulmonary trunk**. At the base of the pulmonary trunk is the pulmonary semilunar valve, which prevents backflow of blood into the right ventricle during ventricular diastole. As the pulmonary trunk reaches the superior surface of the heart, it curves posteriorly and rapidly bifurcates (divides) into two branches, a left and a right **pulmonary artery**. To prevent confusion between these vessels, it is important to refer to the vessel exiting the heart as the pulmonary trunk, rather than also calling it a pulmonary artery.

The pulmonary arteries in turn branch many times within the lung, forming a series of smaller arteries and arterioles that eventually lead to the pulmonary capillaries. The pulmonary capillaries surround lung structures known as alveoli that are the sites of oxygen and carbon dioxide exchange.

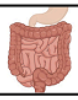
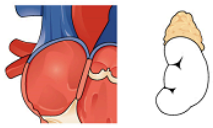
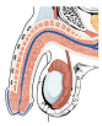
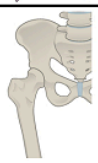
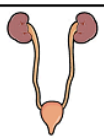
System	Role of Circulatory System
Digestive 	Absorbs nutrients and water; delivers nutrients (except most lipids) to liver for processing by hepatic portal vein; provides nutrients essential for hematopoiesis and building hemoglobin
Endocrine 	Delivers hormones: atrial natriuretic hormone (peptide) secreted by the heart atrial cells to help regulate blood volumes and pressures; epinephrine, ANH, angiotensin II, ADH, and thyroxine to help regulate blood pressure; estrogen to promote vascular health in women and men
Integumentary 	Carries clotting factors, platelets, and white blood cells for hemostasis, fighting infection, and repairing damage; regulates temperature by controlling blood flow to the surface, where heat can be dissipated; provides some coloration of integument; acts as a blood reservoir
Lymphatic 	Transports various white blood cells, including those produced by lymphatic tissue, and immunoglobulins (antibodies) throughout the body to maintain health; carries excess tissue fluid not able to be reabsorbed by the vascular capillaries back to the lymphatic system for processing
Muscular 	Provides nutrients and oxygen for contraction; removes lactic acid and distributes heat generated by contraction; muscular pumps aid in venous return; exercise contributes to cardiovascular health and helps to prevent atherosclerosis
Nervous 	Produces cerebrospinal fluid (CSF) within choroid plexuses; contributes to blood–brain barrier; cardiac and vasomotor centers regulate cardiac output and blood flow through vessels via autonomic system
Reproductive 	Aids in erection of genitalia in both sexes during sexual arousal; transports gonadotropic hormones that regulate reproductive functions
Respiratory 	Provides blood for critical exchange of gases to carry oxygen needed for metabolic reactions and carbon dioxide generated as byproducts of these processes
Skeletal 	Provides calcium, phosphate, and other minerals critical for bone matrix; transports hormones regulating buildup and absorption of matrix including growth hormone (somatotropin), thyroid hormone, calcitonins, and parathyroid hormone; erythropoietin stimulates myeloid cell hematopoiesis; some level of protection for select vessels by bony structures
Urinary 	Delivers 20% of resting circulation to kidneys for filtering, reabsorption of useful products, and secretion of excesses; regulates blood volume and pressure by regulating fluid loss in the form of urine and by releasing the enzyme renin that is essential in the renin-angiotensin-aldosterone mechanism

Figure 39. Interaction of the Circulatory System with Other Body Systems

Once gas exchange is completed, oxygenated blood flows from the pulmonary capillaries into a series of pulmonary venules that eventually lead to a series of larger **pulmonary veins**. Four pulmonary veins, two on the left and two on the right, return blood to the left atrium. At this point, the pulmonary circuit is complete. Table 10 defines the major arteries and veins of the pulmonary circuit discussed in the text.

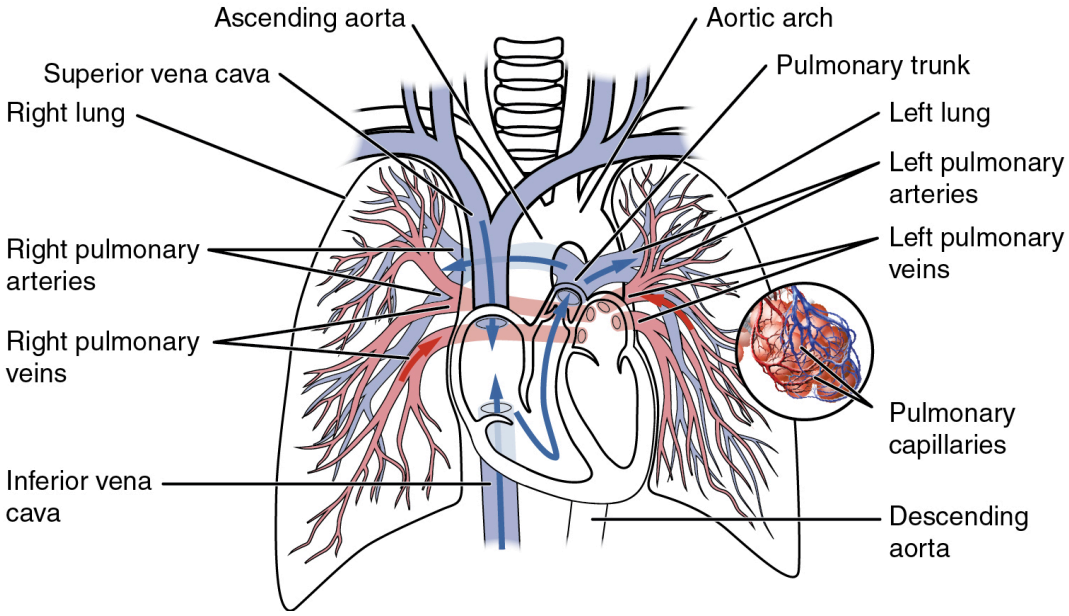


Figure 40. Pulmonary Circuit. *Blood exiting from the right ventricle flows into the pulmonary trunk, which bifurcates into the two pulmonary arteries. These vessels branch to supply blood to the pulmonary capillaries, where gas exchange occurs within the lung alveoli. Blood returns via the pulmonary veins to the left atrium.*

Watch this CrashCourse video to learn more about the blood vessels! Direct link: https://youtu.be/v43ej5lCeBo

Overview of Systemic Arteries: Blood relatively high in oxygen concentration is returned from the pulmonary circuit to the left atrium via the four pulmonary veins. From the left atrium, blood moves into the left ventricle, which pumps blood into the aorta. The aorta and its branches—the systemic arteries—send blood to virtually every organ of the body (Figure 41).

Table 10: Pulmonary arteries and veins

Vessel	Description
Pulmonary trunk	Single large vessel exiting the right ventricle (divides to form the right and left pulmonary arteries)
Pulmonary arteries (left pulmonary artery, right pulmonary artery)	Two vessels that form from the pulmonary trunk and lead to smaller arterioles and eventually to the pulmonary capillaries
Pulmonary veins (left superior pulmonary vein, left inferior pulmonary vein, right superior pulmonary vein, right inferior pulmonary vein)	Two sets of paired vessels (one pair from each side) that are formed from venules, leading blood away from the pulmonary capillaries to flow into the left atrium

The Aorta: The **aorta** is the largest artery in the body (Figure 42). It arises from the left ventricle and eventually descends to the abdominal region, where it bifurcates at the level of the fourth lumbar vertebra into the two common iliac arteries. The aorta consists of the ascending aorta, the aortic arch, and the descending aorta

(Table 11) which passes through the diaphragm, a landmark that divides into the superior thoracic and inferior abdominal components. Arteries originating from the aorta ultimately distribute blood to virtually all tissues of the body. At the base of the aorta is the aortic semilunar valve that prevents backflow of blood into the left ventricle while the heart is relaxing.

After exiting the heart, the **ascending aorta** moves in a superior direction for approximately 5 cm and ends at the sternal angle. Following this ascent, it reverses direction, forming a graceful arc to the left, called the **aortic arch**. The aortic arch descends toward the inferior portions of the body and ends at the level of the intervertebral disk between the fourth and fifth thoracic vertebrae. Beyond this point, the **descending aorta** continues close to the bodies of the vertebrae and passes through an opening in the diaphragm. Superior to the diaphragm, the aorta is called the **thoracic aorta**, and inferior to the diaphragm, it is called the **abdominal aorta**. The abdominal aorta terminates when it bifurcates into the two common iliac arteries at the level of the fourth lumbar vertebra. See Figure 55 for an illustration of the ascending aorta, the aortic arch, and the initial segment of the descending aorta plus major branches.

Coronary Circulation: The first vessels that branch from the ascending aorta are the paired coronary arteries (see Figure 42), which arise from two of the three sinuses in the ascending aorta just superior to the aortic semilunar valve. These sinuses contain the aortic baroreceptors and chemoreceptors critical to maintain cardiac function. The left coronary artery arises from the left posterior aortic sinus. The right coronary artery arises from the anterior aortic sinus. Normally, the right posterior aortic sinus does not give rise to a vessel.

The coronary arteries encircle the heart, forming a ring-like structure that divides into the next level of branches that supplies blood to the heart tissues.

Aortic Arch Branches: There are three major branches of the aortic arch: the **brachiocephalic** artery, the **left common carotid artery**, and the **left subclavian** (literally "under the clavicle") **artery**. As you would expect based upon proximity to the heart, each of these vessels is classified as an elastic artery.

The brachiocephalic artery is located only on the right side of the body; there is no corresponding artery on the left. The brachiocephalic artery branches into the **right subclavian artery** and the **right common carotid artery**. The left subclavian and left common carotid arteries arise independently from the aortic arch but otherwise follow a similar pattern and distribution to the corresponding arteries on the right side (see Figure 42).

Each **subclavian artery** supplies blood to the arms, chest, shoulders, back, and central nervous system.

The **common carotid** artery divides into internal and external carotid arteries. The right common carotid artery arises from the brachiocephalic artery and the left common carotid artery arises directly from the aortic arch. The **branches of the carotid arteries** supply blood to numerous structures within the head and neck. Each internal carotid artery initially forms an expansion known as the carotid sinus, containing the carotid baroreceptors and chemoreceptors. Like their counterparts in the aortic sinuses, the information provided by these receptors is critical to maintaining cardiovascular homeostasis (see Figure 41).

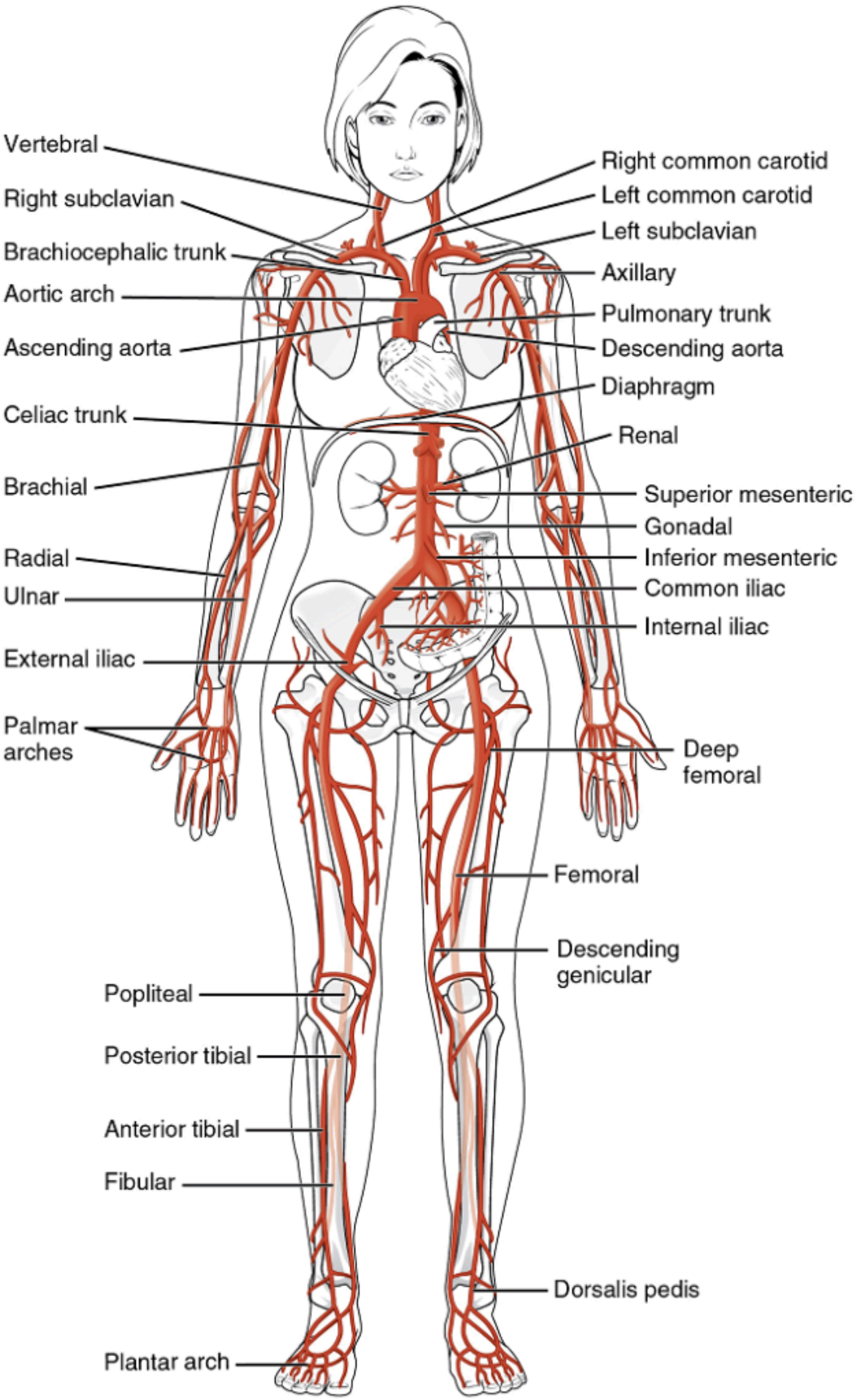


Figure 41. Systemic Arteries. *The major systemic arteries shown here deliver oxygenated blood throughout the body.*

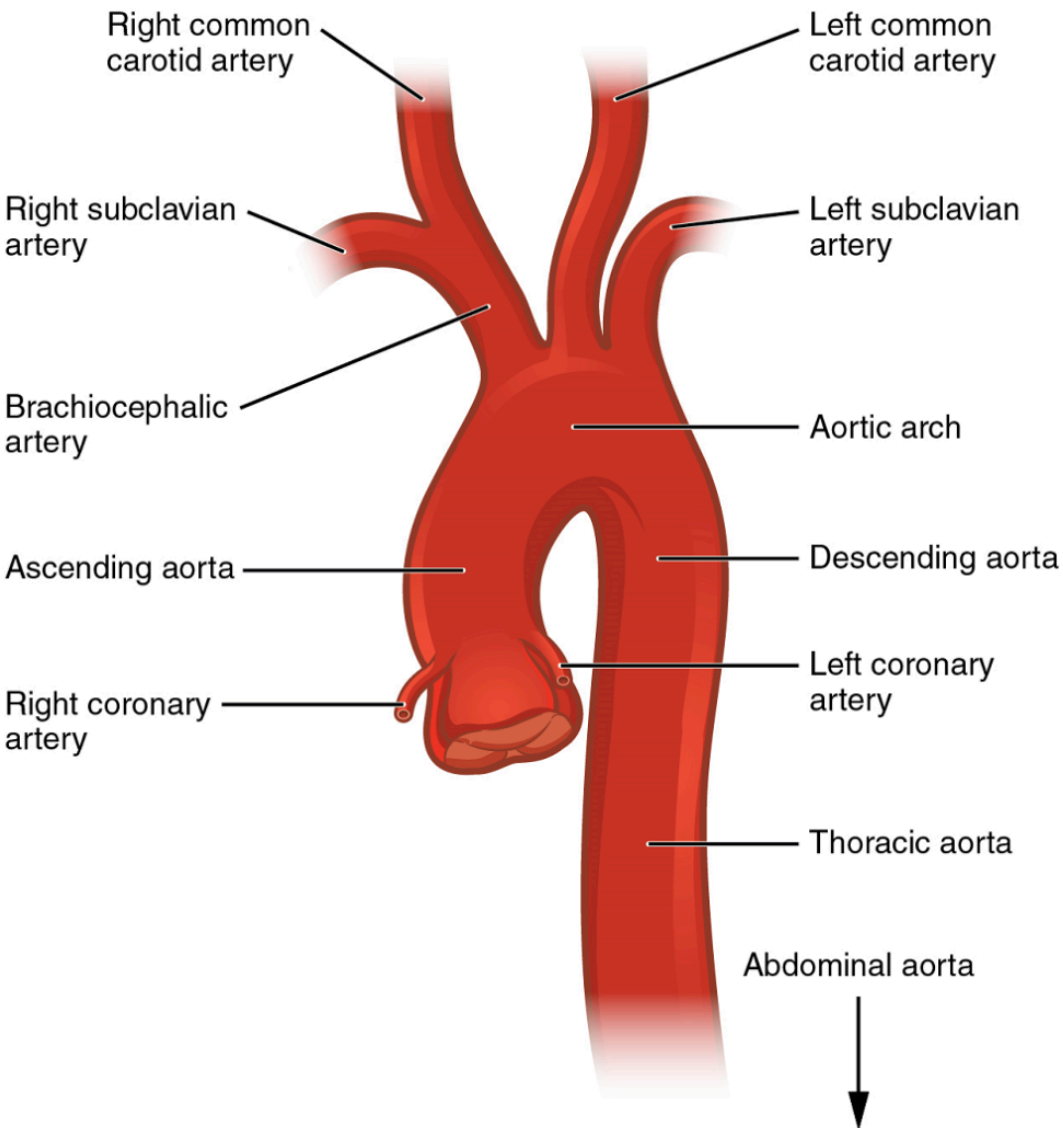


Figure 42. Aorta. *The aorta has distinct regions, including the ascending aorta, aortic arch, and the descending aorta, which includes the thoracic and abdominal regions.*

Table 11: Components of the aorta

Vessel	Description
Aorta	Largest artery in the body; originates from the left ventricle and descends to the abdominal region then bifurcates into the left and right common iliac arteries at the level of the fourth lumbar vertebra
Ascending aorta	Initial portion of the aorta; rises superiorly from the left ventricle for a distance of approximately 5 cm
Aortic arch	Graceful arc to the left that connects the ascending aorta to the descending aorta; ends at the intervertebral disk between the fourth and fifth thoracic vertebrae
Descending aorta	Continues inferiorly from the end of the aortic arch; subdivided into the thoracic aorta and the abdominal aorta
Thoracic aorta	Portion of the descending aorta superior to the aortic hiatus
Abdominal aorta	Portion of the aorta inferior to the aortic hiatus; ends at its bifurcation into the left common iliac artery and the right common iliac artery

Thoracic Aorta and Major Branches: The thoracic aorta begins at the level of vertebra T5 and continues through

to the diaphragm at the level of T12, initially traveling within the mediastinum to the left of the vertebral column. As it passes through the thoracic region, the thoracic aorta gives rise to several branches (Figure 43).

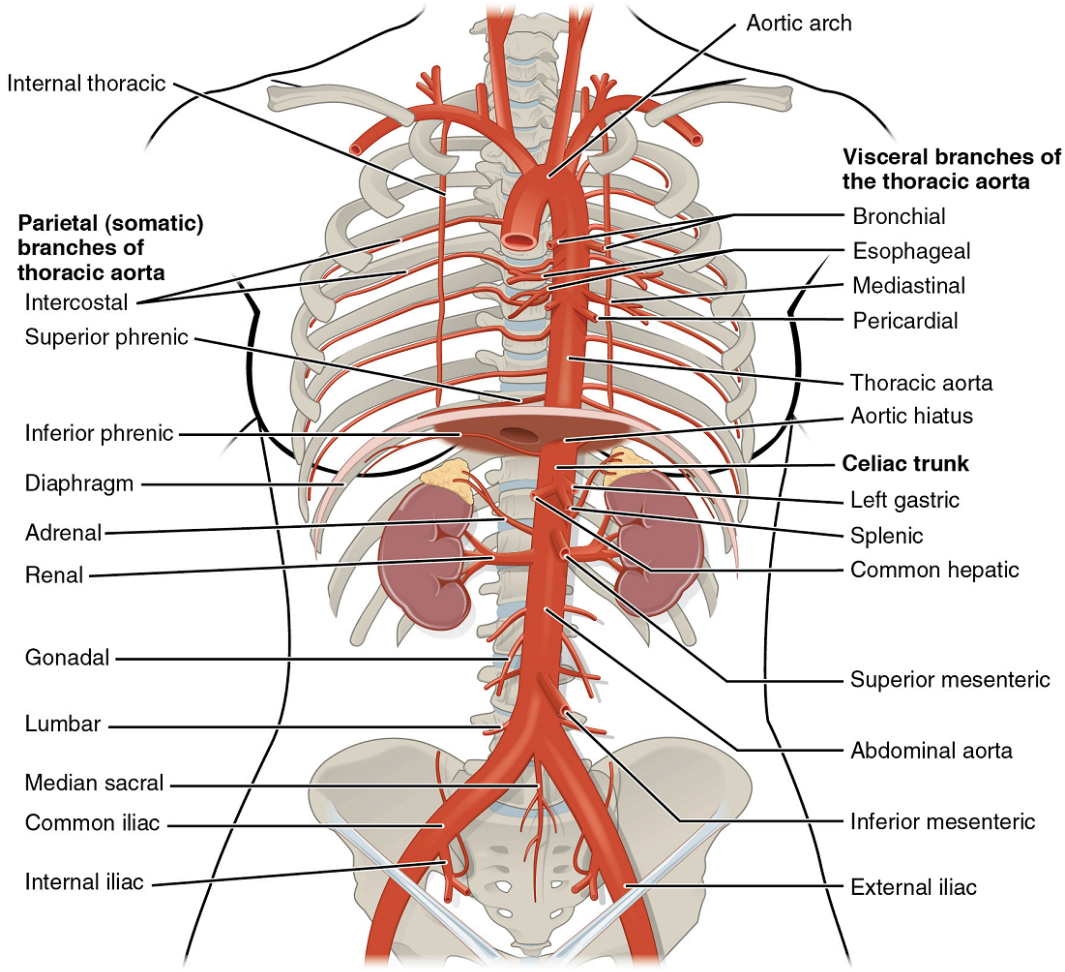


Figure 43. Arteries of the Thoracic and Abdominal Regions. *The thoracic aorta gives rise to the arteries of the visceral and parietal branches.*

Abdominal Aorta and Major Branches: After crossing through the diaphragm, the thoracic aorta is called the abdominal aorta. This vessel remains to the left of the vertebral column and is embedded in adipose tissue behind the peritoneal cavity. It formally ends at approximately the level of vertebra L4, where it bifurcates to form the two (left and right) **common iliac arteries.** Before this division, the abdominal aorta gives rise to several important branches. The common iliac arteries provide blood to the pelvic region and ultimately to the lower limbs.

Arteries Serving the Upper Limbs: As each subclavian artery exits the thorax into the axillary region, it is renamed the **axillary artery**. Although each axillary artery does branch and supply blood to the region near the head of the humerus (via the humeral circumflex arteries), the majority of the vessel continues into the upper arm, or brachium, and becomes the brachial artery.

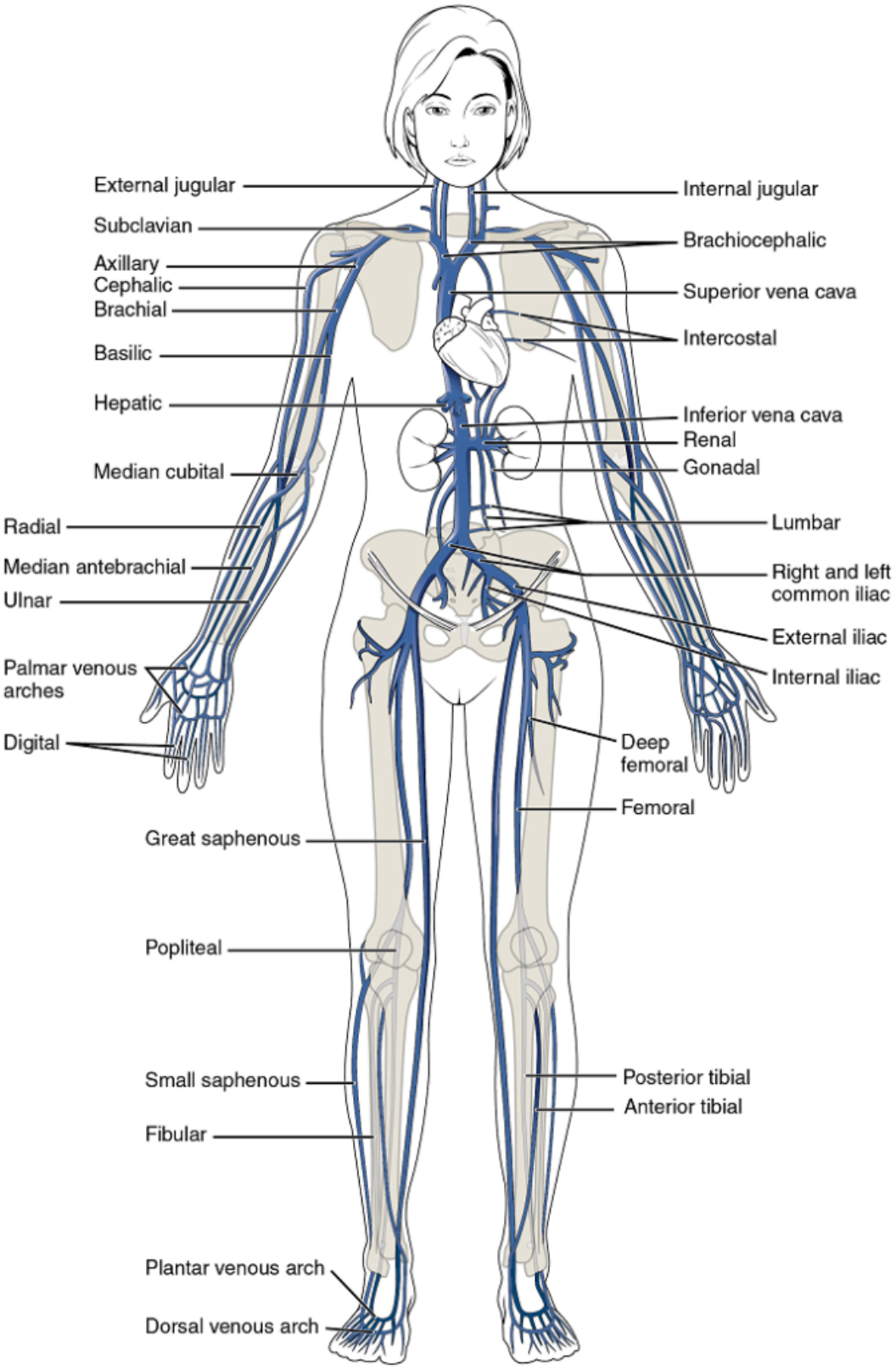


Figure 44. Major Systemic Veins of the Body. *The major systemic veins of the body are shown here in anterior view.*

Arteries Serving the Lower Limbs: Each external iliac artery exits the body cavity and enters the femoral region of the lower leg. As it passes through the body wall, it is renamed the **femoral artery**. Each femoral artery

gives rise to the genicular artery, which provides blood to the region of the knee. As each femoral artery passes posterior to the knee near the popliteal fossa, it is called the popliteal artery. Each popliteal artery branches into anterior and posterior tibial arteries.

Overview of Systemic Veins: Systemic veins return blood to the right atrium. Since the blood has already passed through the systemic capillaries, it will be relatively low in oxygen concentration (Figure 44).

The right atrium receives all of the systemic venous return. Most of the blood flows into either the **superior vena cava** or **inferior vena cava.** If you draw an imaginary line at the level of the diaphragm, systemic venous circulation from above that line will generally flow into the superior vena cava; this includes blood from the head, neck, chest, shoulders, and upper limbs. The exception to this is that most venous blood flow from the coronary veins flows directly into the coronary sinus and from there directly into the right atrium. Beneath the diaphragm, systemic venous flow enters the inferior vena cava, that is, blood from the abdominal and pelvic regions and the lower limbs.

The Superior Vena Cava: The **superior vena cava** drains most of the body superior to the diaphragm (Figure 45). On both the left and right sides, the **subclavian vein** forms when the **axillary vein** passes through the body wall from the axillary region. Each subclavian vein joins with the external and internal jugular veins from the head and neck to form the **brachiocephalic vein**.

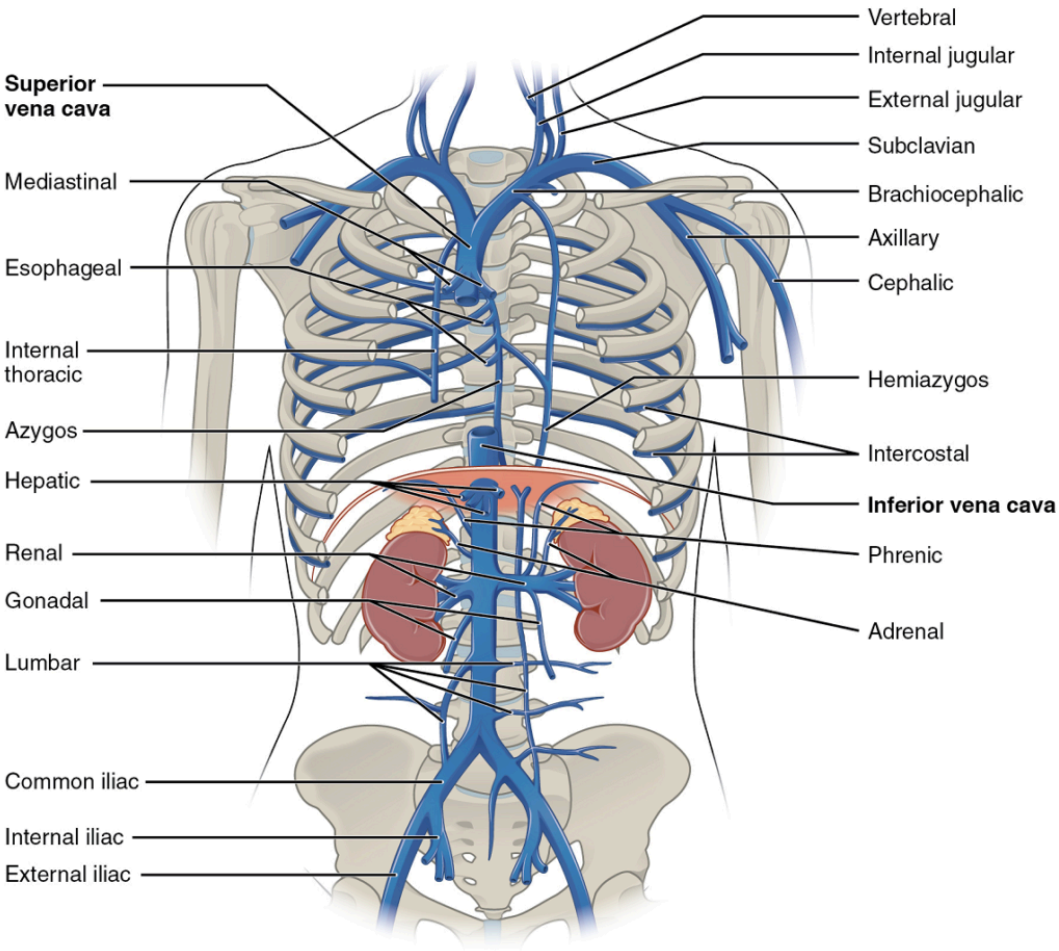


Figure 45. Veins of the Thoracic and Abdominal Regions. *Veins of the thoracic and abdominal regions drain blood from the area above the diaphragm, returning it to the right atrium via the superior vena cava.*

The Inferior Vena Cava: Most of the blood inferior to the diaphragm drains into the **inferior vena cava** before it is returned to the heart (see Figure 45). Lying just beneath the parietal peritoneum in the abdominal cavity, the inferior vena cava parallels the abdominal aorta, where it can receive blood from abdominal veins.

Veins Draining the Lower Limbs: As each **femoral vein** penetrates the body wall from the femoral portion of the upper limb, it becomes the external iliac vein, a large vein that drains blood from the leg to the common iliac vein (Figure 46). The pelvic organs and integument drain into the internal iliac vein on either side of the body, which forms from several smaller veins in the region, including the umbilical veins that run on either side of the bladder. The external and internal iliac veins combine near the inferior portion of the sacroiliac joint on either side to form the **common iliac vein**. In addition to blood supply from the external and internal iliac veins, the middle sacral vein drains the sacral region into the common iliac vein. Similar to the common iliac arteries, the two common iliac veins come together at the level of L5 to form the **inferior vena cava**.

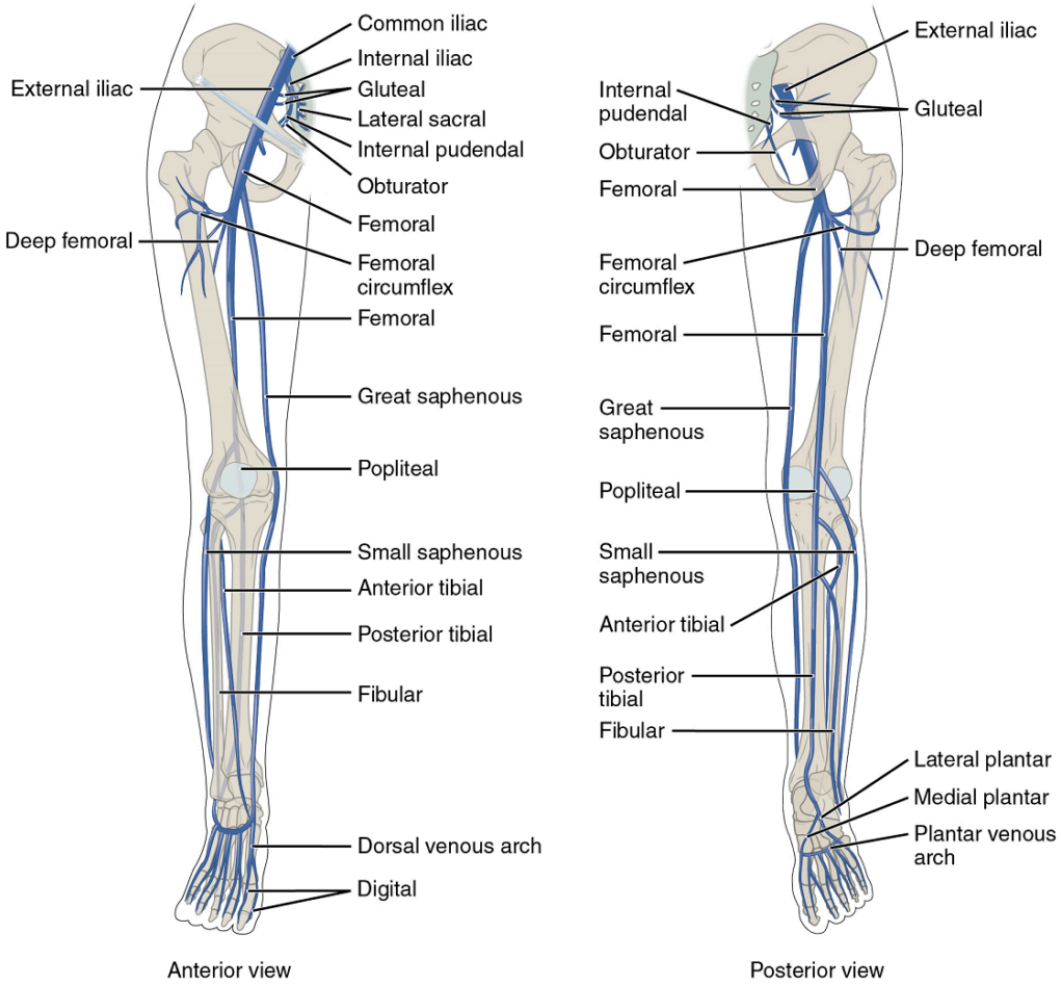


Figure 46. The Major Veins of the Lower Limbs.

Unit 3: The Lymphatic System, Resistance & Immunity

Unit outline

Part 1: Blood Cell Lineages: Leukocytes

- Characteristics of Leukocytes
- Classification of Leukocytes

Part 2: The Lymphatic and Immune Systems

- Anatomy
- Primary Lymphoid Organs and Lymphocyte Development
- The Organization of Immune Function
- Barrier Defenses and the Innate Immune Response
- The Adaptive Immune Response
- The Cellular Basis of Immunological Memory
- Active versus Passive Immunity
- Diseases Associated with Depressed or Overactive Immune Responses

Part 3: Blood Typing

- Antigens, Antibodies, and Transfusion Reactions
- The ABO and Rh Blood Groups
- Determining Blood Types

Learning Objectives

At the end of this unit, you should be able to:

I. Specify the types of leukocytes, their functions, origins and relative quantities in normal blood.

II. Describe the purpose of a differential count and how to interpret the results.

III. Describe the major functions and anatomical organization of the lymphatic system.

IV. Distinguish between the cardiovascular system and the lymphatic system.

V. Explain nonspecific (innate) resistance to disease and specify the general components of nonspecific (innate) resistance.

VI. Explain specific (adaptive) resistance to disease (immunity), and distinguish between T-cell mediated (cellular) immunity and B-cell mediated (humoral) immunity.

VII. Specify the ways in which antibodies destroy or inactivate a foreign substance in the body.

VIII. Discuss the relationship between antibodies and immunization and specify four ways of conferring immunity.

IX. Describe conditions that may result due to a compromised immune system.

X. Describe the basis of the ABO blood groups and Rh factor and explain the significance of this to transfusions and hemolytic disease of the newborn.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Specify the types of leukocytes, their functions, origins and relative quantities in normal blood.

1. List all the types of leukocytes, along with their relative frequencies (most common, least common, etc.).
2. List and describe in detail the function(s) of the five types of leukocytes. Be as specific and detailed as possible, i.e. include every subtype of leukocyte in your answer!

II. Describe the purpose of a differential count and how to interpret the results.

1. What is a differential count? What information can it provide?
2. Why are neutrophils found in high numbers in people recovering from burn injuries?
3. What could be consequences of an abnormally low neutrophil count?
4. How does one of the functions of eosinophils explain their high counts in individuals fighting a parasitic worm infestation?
5. What would the basophil count be in people experiencing allergies? Explain why.
6. Why is a low lymphocyte count observed in individuals with an active HIV infection? 9. Why is this a dangerous situation?
7. Bone marrow disorders cause a low monocyte count. Why?

III. Describe the major functions and anatomical organization of the lymphatic system.

1. For each of the following components of the lymphatic system, state its major function(s) and describe its location(s) in the human body:
    - Lymphatic vessels
    - Lymph
    - Primary lymphoid organs
    - Secondary lymphoid organs

2. Describe the anatomical relationship (i.e. where they are located relative to each other) between lymphatic vessels, lymph, and lymph nodes.

IV. Distinguish between the cardiovascular system and the lymphatic system.

1. Compare and contrast the lymphatic and cardiovascular systems by describing:
    - Any function(s) that both systems serve.
    - Differences between the functions of the two systems.
    - Similarities in the overall structure (cell types, tissue types, organs) of the two systems.
    - Differences between the overall structures found in the two systems.

V. Explain nonspecific (innate) resistance to disease and specify the general components of nonspecific (innate) resistance.

1. What are the main mechanisms that provide the human body with nonspecific resistance to infection?
2. Describe the physical and chemical mechanisms used by each of the following tissues to provide the body with a barrier to disease:
    - The skin
    - Mucous membranes
3. Compare and contrast the physical mechanisms used by the skin to provide the body with a barrier to disease with those used by mucous membranes.
4. Compare and contrast the chemical mechanisms used by the skin to provide the body with a barrier to disease with those used by mucous membranes.
5. Describe the mechanisms by which each of the following acts to provide innate defenses against disease:
    - Phagocytes
    - Natural killer cells
    - Inflammation
    - Soluble mediators
    - Fever

VI. Explain specific (adaptive) resistance to disease (immunity), and distinguish between T-cell mediated (cellular) immunity and B-cell mediated (humoral) immunity.

1. Compare and contrast B-cell mediated (humoral) immunity and T-cell mediated (cellular) immunity, including:
    - The stem cells and progenitor cells required
    - The location(s) where the cells mature
    - The mature cell types involved
    - The mechanism by which each mature cell type responds to an antigen
2. Describe the major immune functions of phagocytes.
3. Describe the immune functions of antigen-presenting cells (APCs).
4. List the specific human cell types that can perform phagocytosis, and those that can become

antigen-presenting cells. Explain any overlap between your two lists.
5. Describe the interactions that occur between the cells of the humoral immunity pathway and cells of the cell-mediated immunity pathway.
6. Describe how the human body normally produces antibodies upon exposure to an antigen.
7. Compare and contrast the innate and adaptive immune responses in terms of timing, specificity, structures and cells involved, and mechanisms of action. Include advantages and disadvantages of each.

VII. Specify the ways in which antibodies destroy or inactivate a foreign substance in the body.

1. Explain each of the 4 possible ways in which an antibody can interact with an antigen to reduce or prevent damage caused by a pathogen (including the mechanism(s) through which each way functions to destroy or inactivate a pathogen or toxin).

VIII. Discuss the relationship between antibodies and immunization and specify four ways of conferring immunity.

1. Discuss the premise of conferring immunity through vaccination.
2. Briefly describe each of the following methods of conferring immunity:
    - Live attenuated vaccines
    - Killed inactivated vaccines
    - Toxoid vaccines
    - Antibody therapy

IX. Describe conditions that may result due to a compromised immune system.

1. Describe the immunological basis of immunodeficiencies in general, and differentiate between inherited and acquired immunodeficiencies. Briefly describe one example of each (i.e., identify the deficiency for the inherited example, and the cause and eventual deficiency of the acquired example).
2. Describe the immunological basis of autoimmune diseases in general, and name and briefly describe at least two specific examples of autoimmune diseases.

X. Describe the basis of the ABO blood groups and Rh factor and explain the significance of this to transfusions and hemolytic disease of the newborn.

1. List all the antigens and antibodies present shortly after birth in each of the 8 major blood types found in humans.
2. Create a chart with all possible human blood types that shows which blood types may donate to which others, and which may not. Explain any cases where one type may receive erythrocytes from, but not donate to, another specific blood type (e.g. “blood type X may not donate to blood type Z, but blood type Z could donate to blood type X”).
3. Explain the conditions under which anti-D antibodies are produced in humans.
4. Describe how hemolytic disease of the newborn occurs and how it can be prevented.

Part 1: Blood Cell Lineages

The cells of the blood, including all those involved in the immune response, arise in the bone marrow via various differentiation pathways from hematopoietic stem cells (Figure 1). In contrast with embryonic stem cells,

hematopoietic stem cells are present throughout adulthood and allow for the continuous differentiation of blood cells to replace those lost to age or function.

Leukocytes: The leukocyte, commonly known as a white blood cell (or WBC), is a major component of the body's defenses against disease. Leukocytes protect the body against invading microorganisms as well as genetically transformed body cells that are potentially cancerous. They also clean up extracellular debris and can signal and enhance the healing and repair process.

Characteristics of Leukocytes: Although leukocytes and erythrocytes both originate from hematopoietic stem cells in the bone marrow, they are very different from each other in many significant ways. For instance, leukocytes are far less numerous than erythrocytes. Typically, there are only 5000 to 10,000 leukocytes per microliter (μl) of blood compared to the roughly 5 million erythrocytes. They are also larger than erythrocytes, possessing a nucleus and organelles while erythrocytes expel these structures early in development. Although there is just one type of erythrocyte, there are many types of leukocytes. Most of these leukocytes have a much shorter lifespan than that of erythrocytes, some as short as a few hours or even a few minutes in the case of acute infection.

One of the most distinctive characteristics of leukocytes is their movement. Whereas erythrocytes spend their days circulating within the blood vessels, leukocytes routinely leave the bloodstream to perform their defensive functions in the body's tissues. For leukocytes, the vascular network is simply a highway they travel and soon exit to reach their true destination. When they arrive, they are often given distinct names, such as macrophage or microglia, depending on their function.

Once they have exited the capillaries, some leukocytes will take up fixed positions in lymphatic tissue, bone marrow, the spleen, the thymus, or other organs. Others will move about through the tissue spaces (**diapedesis**), very much like amoebas, continuously extending their plasma membranes, sometimes wandering freely, and sometimes moving toward the direction in which they are drawn by chemical signals. This attracting of leukocytes occurs because of **positive chemotaxis** (literally "movement in response to chemicals"), a phenomenon in which injured or infected cells and nearby leukocytes emit the equivalent of a chemical "911" call, attracting more leukocytes to the site. In medicine, determining the quantity of the different leukocytes can provide pertinent clinical information. These **differential counts** of the types and percentages of leukocytes present in a sample are often key indicators in making a diagnosis and selecting a treatment.

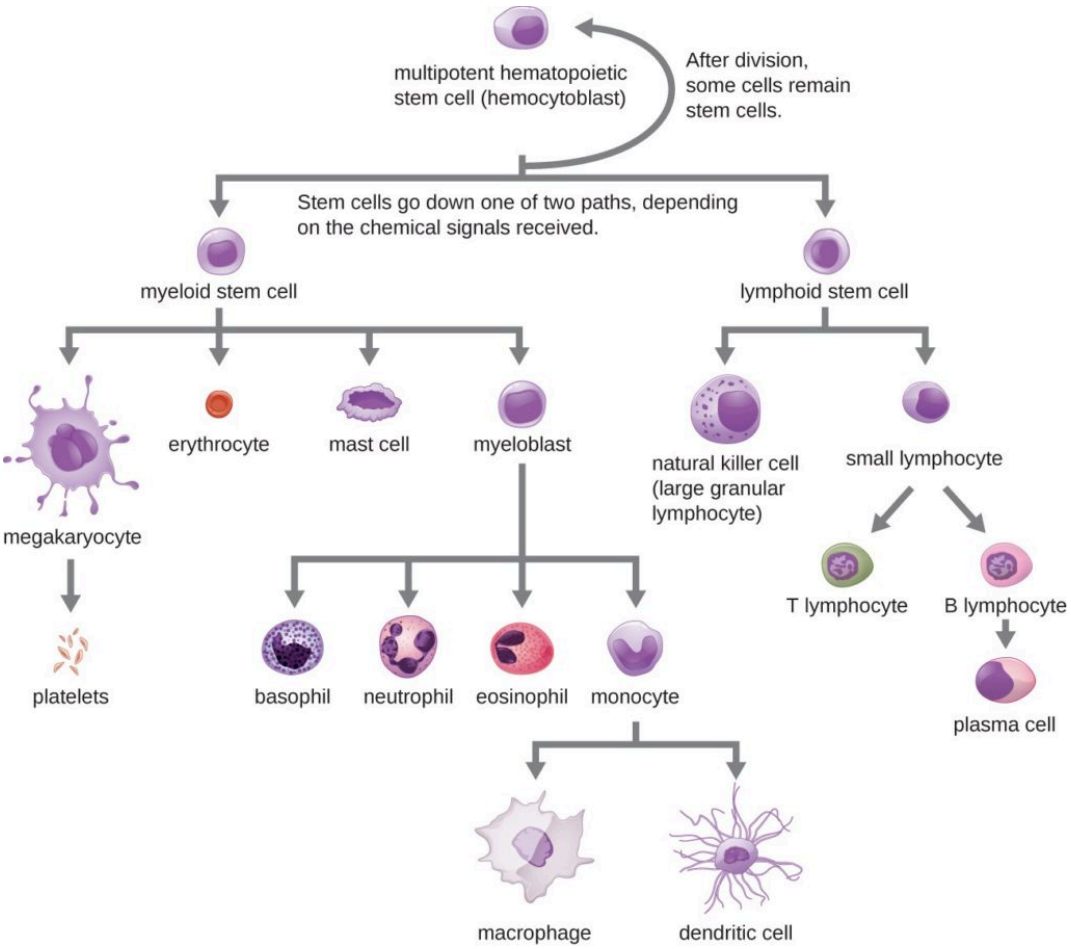


Figure 1. Hematopoietic System of the Bone Marrow. *All the cells of the immune response as well as of the blood arise by differentiation from hematopoietic stem cells. Platelets are cell fragments involved in the clotting of blood.*

Classification of Leukocytes: When scientists first began to observe stained blood slides, it quickly became evident that leukocytes could be divided into two groups, according to whether their cytoplasm contained highly visible granules:

- **Granular leukocytes** contain abundant granules within the cytoplasm. They include neutrophils, eosinophils, and basophils.
- While granules are not totally lacking in **agranular leukocytes**, they are far fewer and less obvious. Agranular leukocytes include monocytes, which mature into phagocytic macrophages, and lymphocytes, which arise from the lymphoid stem cell line.

Granular Leukocytes: We will consider the granular leukocytes in order from most common to least common. All of these are produced in the red bone marrow and have a short lifespan of hours to days. They typically have a lobed nucleus and are classified according to which type of stain best highlights their granules (Figure 2).

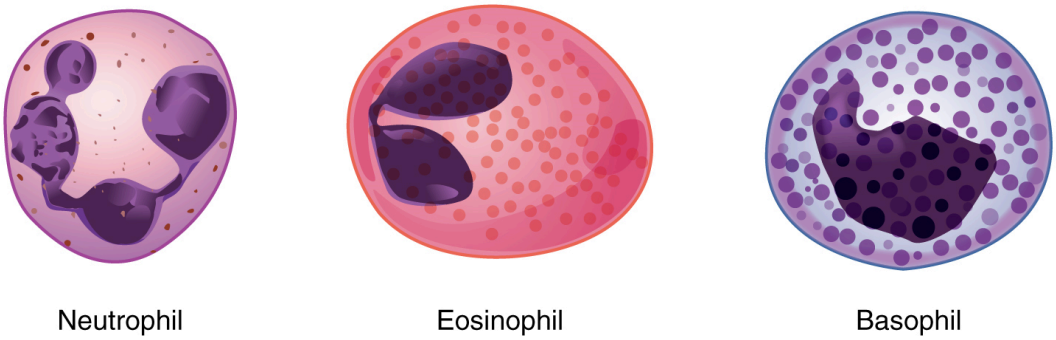


Figure 2. Granular Leukocytes. *A neutrophil has small granules that stain light lilac and a nucleus with two to five lobes. An eosinophil's granules are slightly larger and stain reddish-orange, and its nucleus is typically bilobed. A basophil has large granules that stain dark blue to purple and a bilobed nucleus.*

The most common of all the leukocytes, **neutrophils** will normally comprise 50–70 percent of total leukocyte count. They are called neutrophils because their granules show up most clearly with stains that are chemically neutral (neither acidic nor basic). The nucleus has a distinct lobed appearance and may have two to five lobes, the number increasing with the age of the cell.

Neutrophils are rapid responders to the site of infection and are efficient phagocytes with a preference for bacteria. Their granules include **lysozyme**, an enzyme capable of lysing, or breaking down, bacterial cell walls and **defensins**, proteins that bind to and puncture bacterial and fungal plasma membranes causing the cell contents to leak out. Abnormally high counts of neutrophils indicate infection and/or inflammation, particularly triggered by bacteria, but are also found in burn patients and others experiencing unusual stress. A burn injury increases the proliferation of neutrophils in order to fight off infection that can result from the destruction of the barrier of the skin. Low counts may be caused by drug toxicity and other disorders, and may increase an individual's susceptibility to infection.

Eosinophils typically represent 2–4 percent of total leukocyte count. The granules of eosinophils stain best with an acidic stain known as eosin. The granules of eosinophils include antihistamine molecules, which counteract the activities of histamines, inflammatory chemicals produced by basophils and other inflammatory cells. Some eosinophil granules contain molecules toxic to parasitic worms, which can enter the body either through the integument or when an individual consumes raw or undercooked fish or meat. Eosinophils are also capable of phagocytosis. High counts of eosinophils are typical of patients experiencing allergies, parasitic worm infestations, and some autoimmune diseases. Low counts may be due to drug toxicity and stress.

Basophils are the least common leukocyte, typically comprising less than one percent of the total leukocyte count. The granules of basophils stain best with basic (alkaline) stains. In general, basophils intensify the inflammatory response. The granules of basophils release histamines, which contribute to inflammation, and heparin, which opposes blood clotting. High counts of basophils are associated with allergies, parasitic infections, and hypothyroidism. Low counts are associated with pregnancy, stress, and hyperthyroidism.

Agranular Leukocytes: Agranular leukocytes contain smaller, less-visible granules in their cytoplasm than do granular leukocytes. The nucleus is simple in shape, sometimes with an indentation but without distinct lobes. There are two major types of agranulocytes: lymphocytes and monocytes.

Lymphocytes are the primary cells of adaptive immune responses (Table 1). They are the only formed element of blood that arises from lymphoid stem cells. Although they initially form in the bone marrow, much of their subsequent development and reproduction occurs in the lymphatic tissues. Lymphocytes are the second most common type of leukocyte, accounting for about 20–30 percent of all leukocytes, and are essential for the immune response.

Table 1: Lymphocytes

Type of Lymphocyte	Primary Function
B Lymphocyte	Generates diverse antibodies Memory for subsequent infections
T Lymphocyte	Secretes chemical messengers Cytotoxic activity Memory for subsequent infections
Natural Killer Cell	Destroys virally infected cells

Abnormally high lymphocyte counts are characteristic of viral infections as well as some types of cancer. Abnormally low lymphocyte counts are characteristic of prolonged (chronic) illness or immunosuppression, including that caused by HIV infection and drug therapies that often involve steroids.

The two basic types of lymphocytes, B cells and T cells (also called **B lymphocytes and T lymphocytes)**, are identical morphologically, with a large, often spherical, central nucleus surrounded by a thin layer of cytoplasm. They are distinguished from each other by their surface protein markers as well as by the molecules they secrete. B cells mature in red bone marrow and T cells mature in the thymus. B cells and T cells are found in many parts of the body, circulating in the bloodstream and lymph, and residing in secondary lymphoid organs, including the spleen and lymph nodes. The human body contains approximately 10^{12} lymphocytes. Both B cells and T cells play prominent roles in defending the body against specific pathogens (disease-causing microorganisms) and are involved in specific immunity.

One form of B cells, when activated, become **plasma cells.** These cells differ in morphology from standard B and T cells in that they contain a large amount of cytoplasm packed with the protein-synthesizing machinery known as rough endoplasmic reticulum. A **plasma cell** forms from a naïve B cell with the purpose of producing antibodies or immunoglobulins. An **antibody** is any of the group of proteins that binds specifically to pathogen-associated molecules known as antigens. An **antigen** is a chemical structure on the surface of a pathogen, or the soluble product of a pathogen (ie. a toxin), that binds to T or B lymphocyte receptors. Once activated by binding to antigen, B cells differentiate into plasma cells and begin producing and secreting large quantities of antigen specific antibodies. These travel through the body targeting pathogens or toxins for destruction using mechanisms that will be discussed later in the chapter. This is also referred to as **humoral (body fluid) immunity**.

The **T cell**, on the other hand, does not secrete antibody but performs a variety of functions in the adaptive immune response. Different T cell types have the ability to either secrete soluble factors that communicate with and activate other cells of the adaptive immune response or destroy cells infected with intracellular pathogens. Therefore, T cells provide **cell-mediated immunity** by physically attacking foreign or diseased cells. Both B and T cells can differentiate to **memory cells** that form after exposure to a pathogen and mount rapid responses upon subsequent exposures. Unlike other leukocytes, memory cells live for many years. The roles of T and B lymphocytes in the adaptive immune response will be discussed further on.

Another important lymphocyte is the natural killer cell, a participant in the innate immune response. A **natural killer (NK) cell** is a circulating blood cell that contains cytotoxic (cell-killing) granules in its extensive cytoplasm. It shares this mechanism with the cytotoxic T cells of the adaptive immune response. NK cells are capable of recognizing cells that do not express “self” proteins on their plasma membrane or that contain foreign or abnormal markers. These “non-self” cells include cancer cells, cells infected with a virus, and other cells with atypical surface proteins. Thus, they provide generalized, nonspecific immunity and are among the body’s first lines of defense against viruses and certain types of cancer.

Monocytes originate from myeloid stem cells. They normally represent 2–8 percent of the total leukocyte count. Macrophages are monocytes that have left the circulation and phagocytize debris, foreign pathogens, and many dead, worn out, or damaged cells, including red blood cells. Macrophages also release antimicrobial

defensins and chemotactic chemicals that attract other leukocytes to the site of an infection. Some macrophages occupy fixed locations, whereas others wander through the tissue fluid.

Abnormally high counts of monocytes are associated with certain viral or fungal infections, tuberculosis, and some forms of leukemia and other chronic diseases. Abnormally low counts are typically caused by suppression of the bone marrow due to drugs or infiltration by tumor cells.

Part 2: The Lymphatic and Immune System

In June 1981, the Centers for Disease Control and Prevention (CDC), in Atlanta, Georgia, published a report of an unusual cluster of five patients in Los Angeles, California. All five were diagnosed with a rare pneumonia caused by a fungus called *Pneumocystis jirovecii* (formerly known as *Pneumocystis carinii*).

Why was this unusual? Although commonly found in the lungs of healthy individuals, this fungus is an opportunistic pathogen that causes disease in individuals with suppressed or underdeveloped immune systems. The very young, whose immune systems have yet to mature, and the elderly, whose immune systems have declined with age, are particularly susceptible. The five patients from LA, though, were between 29 and 36 years of age and should have been in the prime of their lives, immunologically speaking. What could be going on?

A few days later, a cluster of eight cases was reported in New York City, also involving young patients, this time exhibiting a rare form of skin cancer known as Kaposi's sarcoma. This cancer of the cells that line the blood and lymphatic vessels was previously observed as a relatively innocuous disease of the elderly. The disease that doctors saw in 1981 was frighteningly more severe, with multiple, fast-growing lesions that spread to all parts of the body, including the trunk and face. Could the immune systems of these young patients have been compromised in some way? Indeed, when they were tested, they exhibited extremely low numbers of a specific type of white blood cell in their bloodstreams, indicating that they had somehow lost a major part of the immune system.

Acquired immune deficiency syndrome, or AIDS, turned out to be a new disease caused by the previously unknown human immunodeficiency virus (HIV). Although nearly 100 percent fatal in those with active HIV infections in the early years, the development of anti-HIV drugs has transformed HIV infection into a chronic, manageable disease and not the certain death sentence it once was. One positive outcome resulting from the emergence of HIV disease was that the public's attention became focused as never before on the importance of having a functional and healthy immune system.

Anatomy of the Lymphatic and Immune Systems: The **immune system** is the complex collection of cells and organs that destroys or neutralizes pathogens that would otherwise cause disease or death. The **lymphatic system** is the system of vessels, cells, and organs that carries excess fluids to the bloodstream and filters pathogens from the blood. The swelling of lymph nodes during an infection and the transport of lymphocytes via the lymphatic vessels are but two examples of the many connections between these critical organ systems.

Functions of the Lymphatic System: A major function of the lymphatic system is to drain body fluids and return them to the bloodstream. Blood pressure causes leakage of fluid from the capillaries, resulting in the accumulation of fluid in the interstitial space—that is, spaces between individual cells in the tissues. In humans, 20 liters of plasma is released into the interstitial space of the tissues each day due to capillary filtration. Once this filtrate is out of the bloodstream and in the tissue spaces, it is referred to as interstitial fluid. Of this, 17 liters is reabsorbed directly by the blood vessels. But what happens to the remaining three liters? This is where the lymphatic system comes into play. It drains the excess fluid and empties it back into the bloodstream via a series of vessels, trunks, and ducts. **Lymph** is the term used to describe interstitial fluid once it has entered the lymphatic system. When the lymphatic system is damaged in some way, such as by being blocked by cancer cells or destroyed by injury, protein-rich interstitial fluid accumulates (sometimes "backs up" from the lymph vessels) in the tissue spaces. This inappropriate accumulation of fluid referred to as lymphedema may lead to serious medical consequences.

As the vertebrate immune system evolved, the network of lymphatic vessels became convenient avenues for

transporting the cells of the immune system. Additionally, dietary lipids and fat-soluble vitamins absorbed in the gut use this system of transport.

Cells of the immune system not only use lymphatic vessels to make their way from interstitial spaces back into the circulation, but they also use **lymph nodes** as major staging areas for the development of critical immune responses. A lymph node is one of the small, bean-shaped organs located throughout the lymphatic system.

Structure of the Lymphatic System: The lymphatic vessels begin as open-ended capillaries, which feed into larger and larger lymphatic vessels, and eventually empty into the bloodstream by a series of ducts. Along the way, the lymph travels through the lymph nodes, which are commonly found near the groin, armpits, neck, chest, and abdomen. Humans have about 500–600 lymph nodes throughout the body (Figure 3).

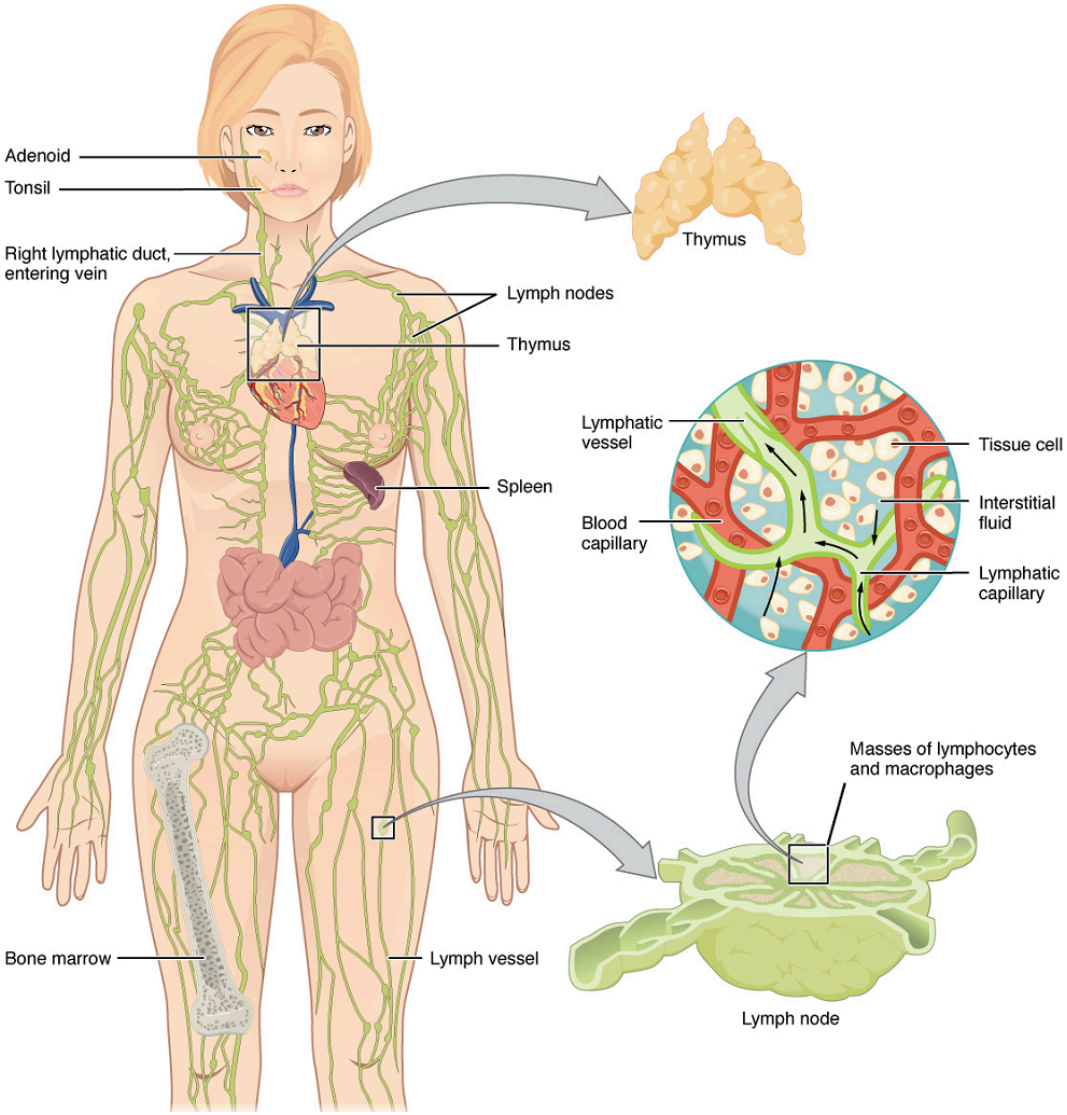


Figure 3. Anatomy of the Lymphatic System. *Lymphatic vessels in the arms and legs convey lymph to the larger lymphatic vessels in the torso.*

A major distinction between the lymphatic and cardiovascular systems in humans is that lymph is not actively pumped by the heart, but is forced through the vessels by the movements of the body, the contraction of skeletal muscles during body movements, and breathing. One-way valves (semi-lunar valves) in lymphatic vessels keep the lymph moving toward the heart. Lymph flows from the lymphatic capillaries, through

lymphatic vessels, and then re-enters the circulatory system via the lymphatic ducts located at the junction of the jugular and subclavian veins in the neck.

Lymphatic Capillaries: Lymphatic capillaries, also called the terminal lymphatics, are vessels where interstitial fluid enters the lymphatic system to become lymph fluid. Located in almost every tissue in the body, these vessels are interlaced among the arterioles and venules of the circulatory system in the soft connective tissues of the body (Figure 4). Exceptions are the central nervous system, bone marrow, bones, teeth, and the cornea of the eye, which do not contain lymph vessels.

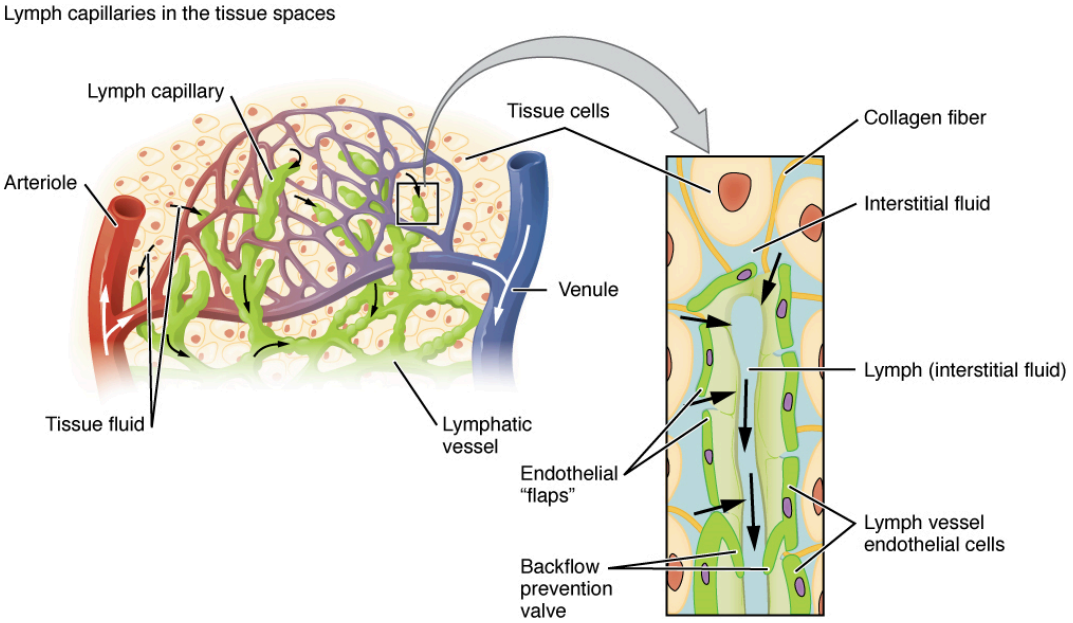


Figure 4. Lymphatic Capillaries. *Lymphatic capillaries are interlaced with the arterioles and venules of the cardiovascular system. Collagen fibers anchor a lymphatic capillary in the tissue (inset). Interstitial fluid slips through spaces between the overlapping endothelial cells that compose the lymphatic capillary.*

Larger Lymphatic Vessels, Trunks, and Ducts: The lymphatic capillaries empty into larger lymphatic vessels, which are similar to veins in terms of their three-tunic structure and the presence of valves. These one-way valves are located fairly close to one another, and each one causes a bulge in the lymphatic vessel, giving the vessels a beaded appearance (see Figure 4). The superficial and deep lymphatics eventually merge to form larger lymphatic vessels known as **lymphatic trunks**. On the right side of the body, the right sides of the head, thorax, and right upper limb drain lymph fluid into the right subclavian vein via the right lymphatic duct (Figure 5). On the left side of the body, the remaining portions of the body drain into the larger thoracic duct, which drains into the left subclavian vein.

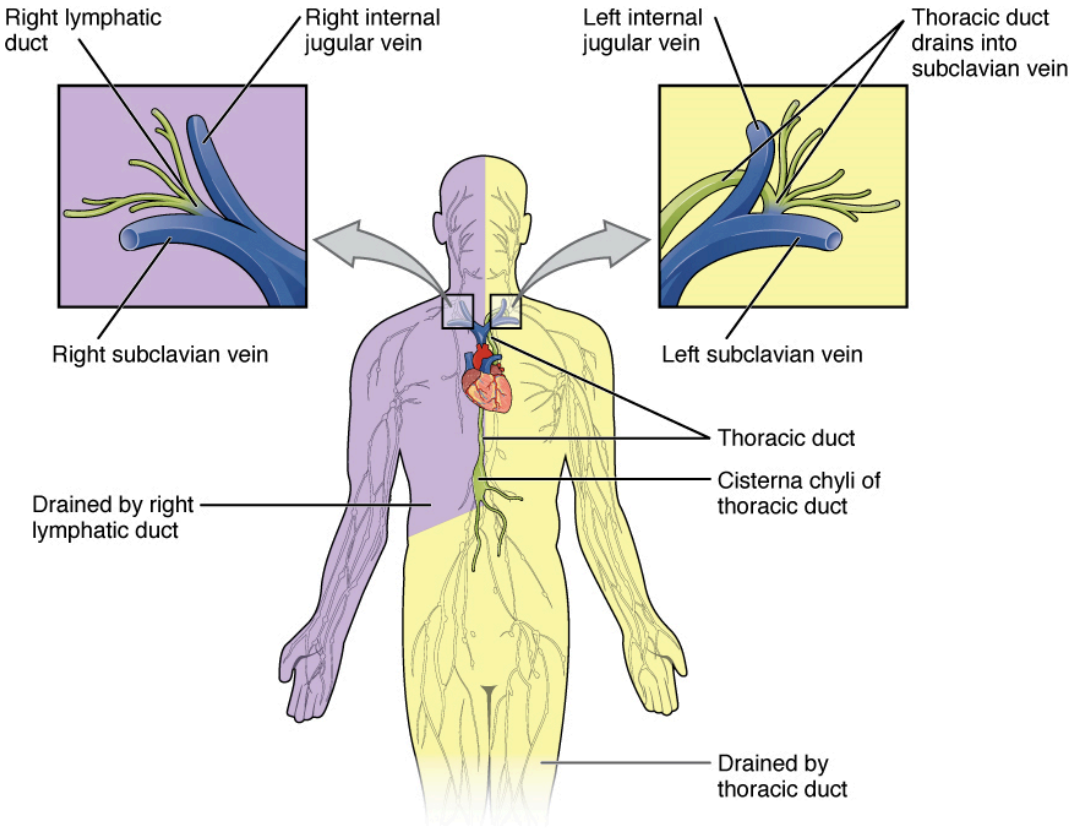


Figure 5. Major Trunks and Ducts of the Lymphatic System. *The thoracic duct drains a much larger portion of the body than does the right lymphatic duct.*

Primary Lymphoid Organs and Lymphocyte Development: Understanding the differentiation and development of B and T cells is critical to the understanding of the adaptive immune response. It is through this process that the body (ideally) learns to destroy only pathogens and leaves the body's own cells relatively intact. The primary lymphoid organs are the bone marrow and thymus gland. The lymphoid organs are where lymphocytes mature, proliferate, and are selected, which enables them to attack pathogens without harming the cells of the body.

Bone Marrow: The bone marrow is responsible for most hematopoietic functions, although the final stages of the differentiation of some cells may take place in other organs. The red bone marrow is a loose collection of cells where hematopoiesis occurs, and the yellow bone marrow is a site of energy storage, which consists largely of fat cells (Figure 6). The B cell undergoes nearly all of its development in the red bone marrow, whereas the immature T cell, called a thymocyte, leaves the bone marrow and matures largely in the thymus gland.

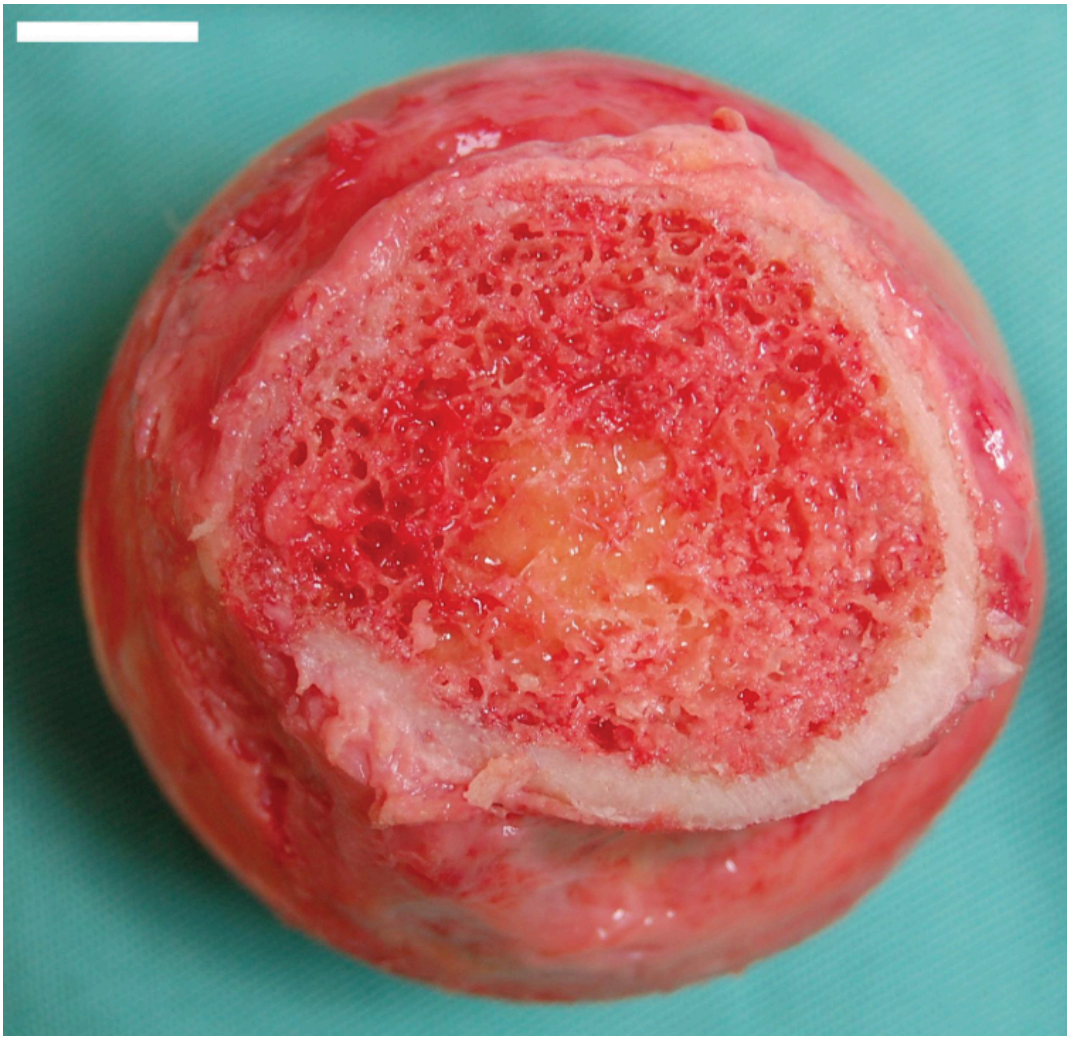


Figure 6. Bone Marrow. *Red bone marrow fills the head of the femur, and a spot of yellow bone marrow is visible in the center. The white reference bar is 1 cm.*

Thymus: The thymus gland is a bilobed organ found in the space behind the sternum and anterior to the heart, it overlies the aortic arch, superior vena cava and trachea. The organ contains large numbers of thymocytes with some epithelial cells, macrophages, and dendritic cells (two types of phagocytic cells that are derived from monocytes). As mentioned, thymocytes mature into T cells in the thymus.

Secondary Lymphoid Organs and their Roles in Active Immune Responses: Lymphocytes develop and mature in the primary lymphoid organs, but they mount immune responses from the secondary lymphoid organs. A **naïve lymphocyte** is one that has left the primary organ and entered a secondary lymphoid organ. Naïve lymphocytes are fully functional immunologically, but have yet to encounter an antigen to respond to. In addition to circulating in the blood and lymph, lymphocytes concentrate in secondary lymphoid organs, which include the lymph nodes, spleen, and lymphoid nodules such as the tonsils.

Lymph Nodes: Lymph nodes function to remove debris and pathogens from the lymph, and are thus sometimes referred to as the "filters of the lymph". Any bacteria that infect the interstitial fluid are taken up by the lymphatic capillaries and transported to a regional lymph node. Dendritic cells and macrophages within this organ internalize and kill many of the pathogens that pass through, thereby removing them from the body. The lymph node is also the site of adaptive immune responses mediated by T cells, B cells, and accessory cells of the adaptive immune system.

Spleen: In addition to the lymph nodes, the **spleen** is a major secondary lymphoid organ. It is about 12 cm

(5 in) long and is attached to the lateral border of the stomach via the gastrosplenic ligament. The spleen is sometimes called the “filter of the blood” because of its extensive vascularization and the presence of macrophages and dendritic cells that remove microbes and other materials from the blood, including dying red blood cells. The spleen also functions as the location of immune responses to blood-borne pathogens.

Tonsils: These lymphoid nodules located along the inner surface of the pharynx are important in developing immunity to oral pathogens (Figure 7). The tonsil located at the back of the throat, called the pharyngeal tonsil, is sometimes referred to as the adenoid when swollen. Such swelling is an indication of an active immune response to infection. This seems to be the major function of tonsils—to help children’s bodies recognize, destroy, and develop immunity to common environmental pathogens so that they will be protected in their later lives. Tonsils are often removed in those children who have recurring throat infections, especially those involving the palatine tonsils on either side of the throat, whose swelling may interfere with their breathing and/or swallowing.

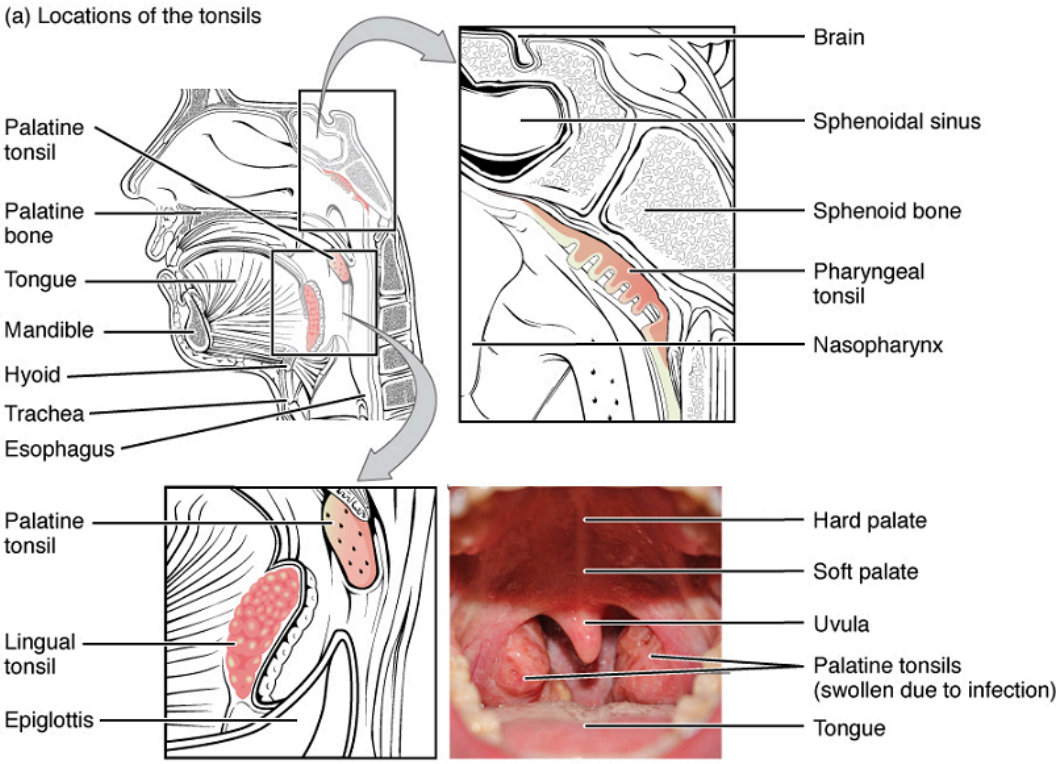
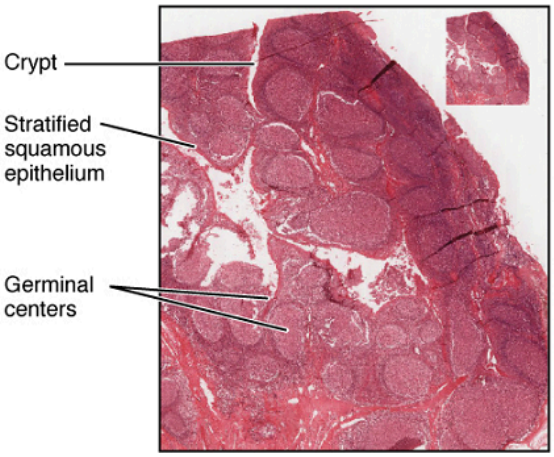


Figure 7. Locations of the Tonsils. *(a) The pharyngeal tonsil is located on the roof of the posterior superior wall of the nasopharynx. The palatine tonsils lay on each side of the pharynx.*

(b) Histology of palatine tonsil



Watch this CrashCourse video to learn more about the lymphatic system! Direct link: https://youtu.be/l7orwMgTQ5I

The Organization of Immune Function: The immune system is a collection of barriers, cells, and soluble proteins that interact and communicate with each other in extraordinarily complex ways. The modern model of immune function is organized into three phases based on the timing of their effects. The three temporal phases consist of the following:

- **Barrier defenses** such as the skin and mucous membranes, which act instantaneously to prevent pathogenic invasion into the body tissues.
- The rapid but nonspecific **innate immune response**, which consists of a variety of specialized cells and soluble factors.
- The slower but more specific and effective **adaptive immune response**, which involves many cell types and soluble factors, but is primarily controlled by white blood cells (leukocytes) known as **lymphocytes**, which help control immune responses.

Barrier Defenses and the Innate Immune Response: The immune system can be divided into two overlapping mechanisms to destroy pathogens: the innate immune response, which is relatively rapid but nonspecific and thus not always effective, and the adaptive immune response, which is slower in its development following infection, but is highly specific and effective at attacking a wide variety of pathogens (Figure 8).

Any discussion of the innate immune response usually begins with the physical barriers that prevent pathogens from entering the body, destroy them after they enter, or flush them out before they can establish themselves in the hospitable environment of the body's soft tissues. Barrier defenses are part of the body's most basic defense mechanisms. The barrier defenses are not a response to infections, but they are continuously working to protect against a broad range of pathogens.

The different modes of barrier defenses are associated with the external surfaces of the body, where pathogens may try to enter (Table 2). The primary barrier to the entrance of microorganisms into the body is the skin. Not only is the skin covered with a layer of dead, keratinized epithelium that is too dry for bacteria in which to grow, but as these cells are continuously sloughed off from the skin, they carry bacteria and other pathogens with them. Additionally, sweat and other skin secretions may lower pH, contain toxic lipids, contain antimicrobial peptides such as dermcidin, and physically wash microbes away.

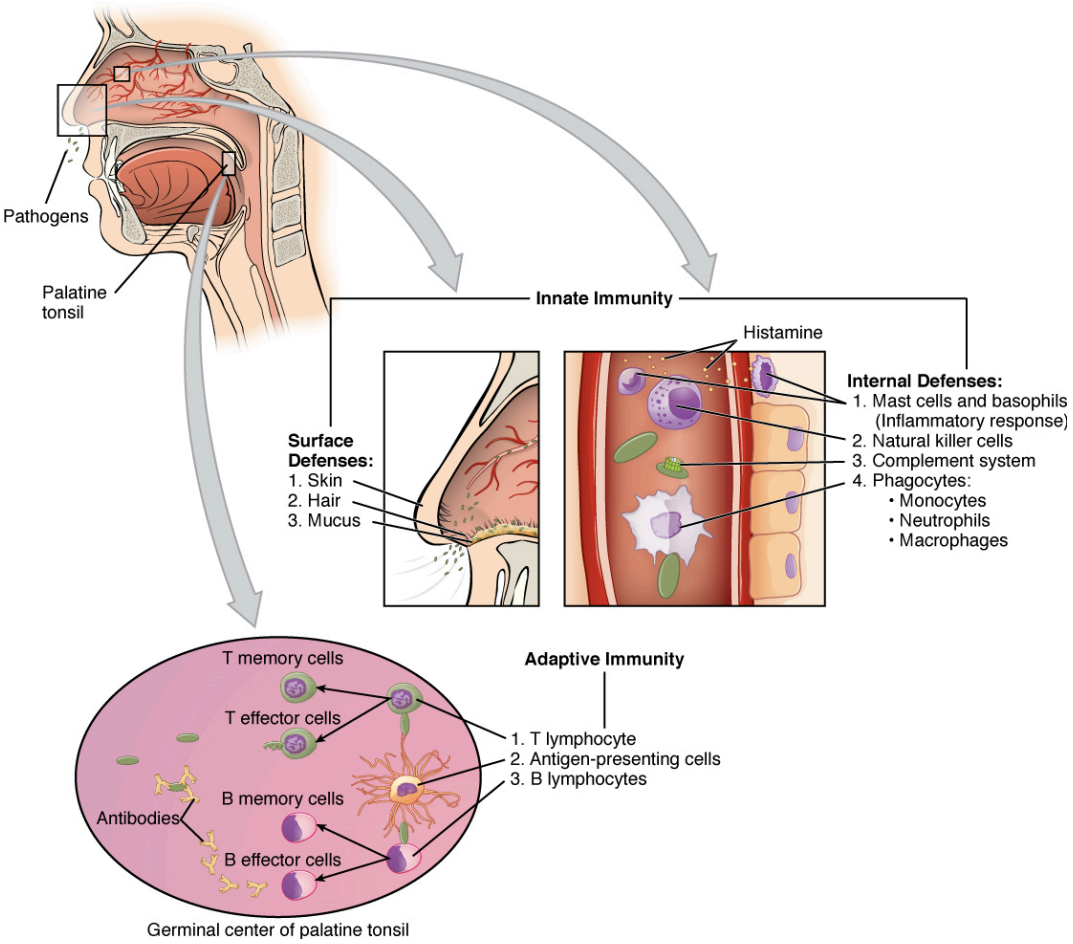


Figure 8. Cooperation between Innate and Adaptive Immune Responses. *The innate immune system enhances adaptive immune responses so they can be more effective*

Another barrier is the saliva in the mouth, which is rich in lysozyme—an enzyme that destroys bacteria by digesting their cell walls. The acidic environment of the stomach, which is fatal to many pathogens, is also a barrier. Additionally, the mucus layer of the gastrointestinal tract, respiratory tract, reproductive tract, eyes, ears, and nose traps both microbes and debris, and facilitates their removal. In the case of the upper respiratory tract, ciliated epithelial cells move potentially contaminated mucus upwards to the mouth, where it is then swallowed into the digestive tract, ending up in the harsh acidic environment of the stomach. Considering how often you breathe compared to how often you eat or perform other activities that expose you to pathogens, it is not surprising that multiple barrier mechanisms have evolved to work in concert to protect this vital area.

Cells of the Innate Immune Response: A phagocyte is a cell that is able to surround and engulf a particle or cell, a process called phagocytosis**.** The phagocytes of the immune system engulf other particles or cells, either to clean an area of debris, remove old cells, or to kill pathogenic organisms such as bacteria. The phagocytes are the body's fast acting, first line of immunological defense against organisms that have breached barrier defenses and have entered the vulnerable tissues of the body.

Table 2: Barrier Defenses

Site	Defensive structure	Protective aspect
Skin (physical structure)	Epidermal surface	Keratinized cells of surface, Langerhans cells
Skin (secretions)	Eccrine glands	Low pH, dermcidin, washing action
Oral cavity	Salivary glands	Lysozyme
Stomach	Gastric juice	Low pH
Mucous membranes	Mucosal epithelium	Layered cells
Mucous membranes (secretions)	Cells producing mucus	Traps pathogens, dust, debris, etc.; washing action; defensins and lysozyme
Skin and mucosal surfaces	Normal flora (nonpathogenic bacteria)	Compete with pathogenic microbes

Phagocytes: Macrophages and Neutrophils:

Many of the cells of the immune system have a phagocytic ability, at least at some point during their life cycles. Phagocytosis is an important and effective mechanism of destroying pathogens during innate immune responses. The phagocyte takes the organism inside itself as a phagosome, which subsequently fuses with a lysosome and its digestive enzymes, forming a phagolysosome, and thus effectively killing many pathogens. On the other hand, some bacteria including *Mycobacterium tuberculosis*, the pathogen causing tuberculosis, may be resistant to these enzymes and are therefore much more difficult to clear from the body. Macrophages, neutrophils, and dendritic cells are the major phagocytes of the immune system.

A **macrophage** is an irregularly shaped phagocyte that is amoeboid in nature and is the most versatile of the phagocytes in the body. Macrophages move through tissues and squeeze through capillary walls using pseudopodia. They not only participate in innate immune responses but have also evolved to cooperate with lymphocytes as part of the adaptive immune response. Macrophages exist in many tissues of the body, either freely roaming through connective tissues or fixed to reticular fibers within specific tissues such as lymph nodes. When pathogens breach the body's barrier defenses, macrophages are the first line of defense (Table 3). They are called different names, depending on the tissue: Kupffer cells in the liver, histiocytes in connective tissue, microglia in the brain, and alveolar macrophages in the lungs.

A **neutrophil** is a phagocytic cell that is attracted via chemotaxis from the bloodstream to infected tissues. Whereas macrophages act like sentries, always on guard against infection, neutrophils can be thought of as military reinforcements that are called into a battle to hasten the destruction of the enemy. Neutrophils are usually thought of as the primary pathogen-killing cell of the inflammatory process of the innate immune response.

A **monocyte** is a circulating precursor cell that differentiates into either a macrophage or dendritic cell. Monocytes can be rapidly attracted to areas of infection by signal molecules of inflammation.

Table 3: Phagocytic Cells of the Innate Immune System

Cell	Cell type	Primary location	Function in the innate immune response
Macrophage	Agranulocyte	Body cavities/organs	Phagocytosis
Dendritic cell	Agranulocyte	Skin and mucous membranes	Phagocytosis
Neutrophil	Granulocyte	Blood	Phagocytosis
Monocyte	Agranulocyte	Blood	Precursor of macrophages and dendritic cells

Natural Killer Cells: NK cells, as mentioned previously, are a type of lymphocyte that play an important role in the innate immune response. They have the ability to induce apoptosis, that is, programmed cell death, in cells infected with intracellular pathogens such as obligate intracellular bacteria (for example, *Mycobacteria*) and

viruses. NK cells recognize these cells by mechanisms that are still not well understood, but that presumably involve their surface receptors. NK cells can induce apoptosis, in which a cascade of events inside the cell causes its own death. In addition, NK cells secrete chemicals which enhance inflammation.

Should the cells of the innate immune system come into contact with a species of pathogen they recognize, the cell will bind to the pathogen and initiate phagocytosis (or cellular apoptosis in the case of an intracellular pathogen) in an effort to destroy the offending microbe.

Soluble Mediators of the Innate Immune Response: These are soluble factors secreted during innate or early induced responses, and later during adaptive immune responses. Examples include signaling molecules such as **cytokines** or **chemokines** used to recruit and activate immunological cells. Proteins involved in the **complement system** are also important mediators of the immune response. The complement system is a series of signaling cascades with functions such as labeling pathogens for phagocytosis (opsonization) or killing pathogens by directly damaging the plasma membrane.

Inflammatory Response: Everyone has experienced inflammation at some point in their lives. Stub a toe, cut a finger, or perform any activity that causes tissue damage and inflammation will result, with its four characteristics: heat, redness, pain, and swelling ("loss of function" is sometimes mentioned as a fifth characteristic). It is important to note that inflammation does not have to be initiated by an infection, but can also be caused by tissue injuries. The release of damaged cellular contents into the site of injury is enough to stimulate the response, even in the absence of breaks in physical barriers that would allow pathogens to enter (by hitting your thumb with a hammer, for example). The inflammatory reaction brings in phagocytic cells to the damaged area to clear cellular debris and to set the stage for wound repair (Figure 9).

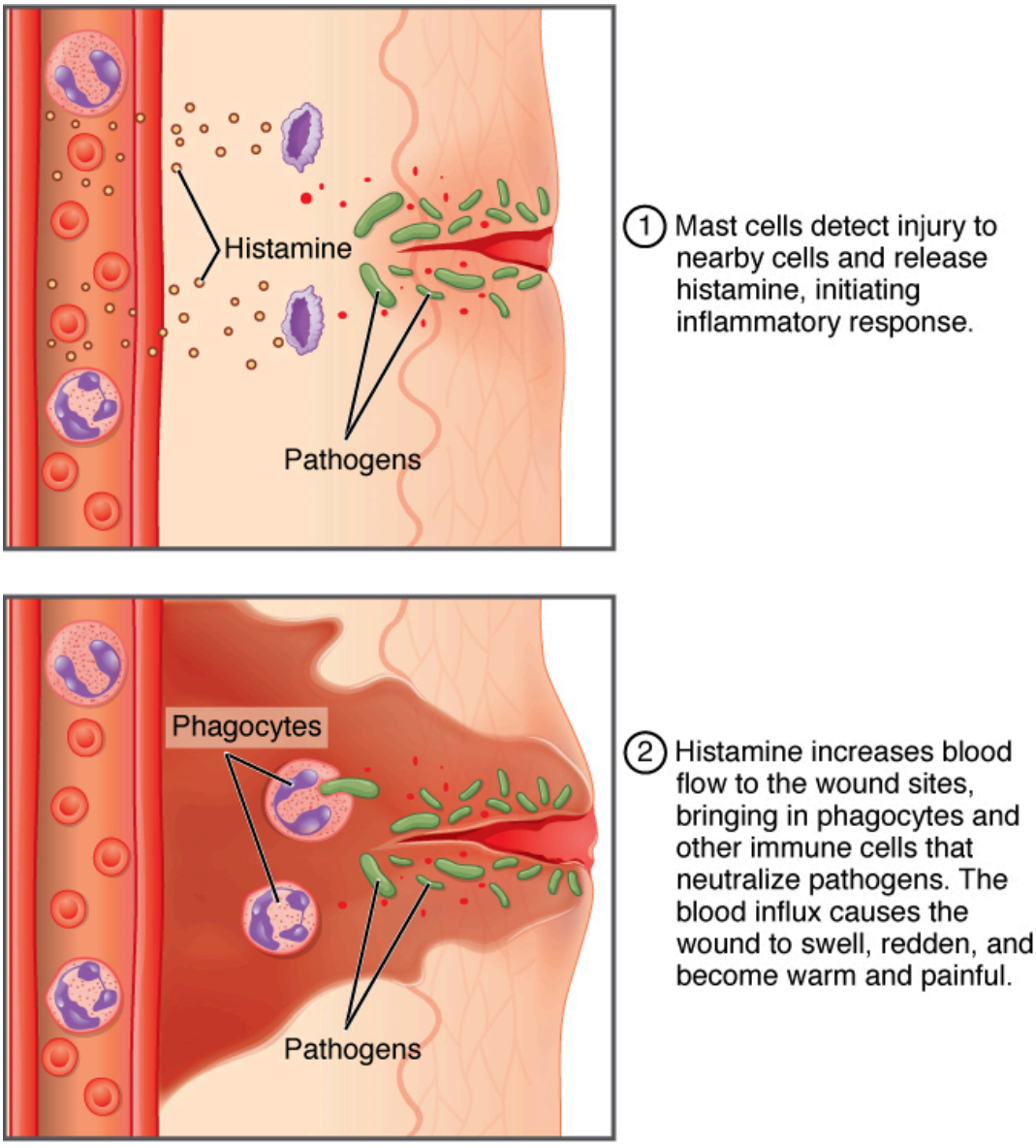


Figure 9. Inflammatory Response.

This reaction also brings in the cells of the innate immune system, allowing them to get rid of the sources of a possible infection or injury. Inflammation is part of a very basic form of immune response. The process not only brings fluid and cells into the site of damage to destroy the pathogen and remove it and any debris from the site, but it also helps to isolate the site, limiting the spread of the pathogen. **Acute inflammation** is a short-term inflammatory response to an insult to the body. If the cause of the inflammation is not resolved, however, it can lead to chronic inflammation, which is associated with major tissue destruction and fibrosis. **Chronic inflammation** is ongoing inflammation. It can be caused by foreign bodies, persistent pathogens, and autoimmune diseases such as rheumatoid arthritis.

There are four important parts to the inflammatory response:

- *Tissue Injury.* The released contents of injured cells stimulate the release of **mast cell** granules and their potent inflammatory mediators. These mediators are involved in vasodilation, increasing permeability of local capillaries, and recruitment of phagocytes.
- *Vasodilation.* Many inflammatory mediators such as histamine are vasodilators that increase the diameters

of local capillaries. This causes increased blood flow and is responsible for the heat and redness of inflamed tissue. It allows greater access of the immune components of blood to the site of inflammation.

- *Increased Vascular Permeability.* At the same time, inflammatory mediators increase the permeability of the local vasculature, causing leakage of fluid into the interstitial space, resulting in the swelling, or edema, associated with inflammation. This allows immune cells and mediators to exit the blood stream and enter the site of infection or injury.
- *Recruitment of Phagocytes.* Inflammatory mediators also attract neutrophils from the blood to the site of infection by chemotaxis. Following an early neutrophil infiltration stimulated by macrophage signals, more macrophages are recruited to clean up the debris remaining at the site. When local infections are severe, neutrophils are attracted to the sites of infections in large numbers, and as they phagocytose the pathogens and subsequently die, their accumulated cellular remains are visible as pus at the infection site.

Overall, inflammation is valuable for many reasons. Not only are the pathogens killed and debris removed, but the increase in vascular permeability encourages the entry of clotting factors, the first step towards wound repair. Inflammation also facilitates the transport of antigen to lymph nodes by macrophages or dendritic cells for the development of the adaptive immune response.

Fever: The mechanisms of inflammation described so far are primarily local. Another inflammatory response that is systemic in nature is that of fever. Fever is defined as an increase in the set-point of the body's thermostat, with the result that homeostatic mechanisms raise the temperature of the body above the normal of about 37°C.

The increase in temperature has several effects that are beneficial to the body's defense. These include increasing the activity of the immune system (e.g., enhancing the efficiency of white blood cells). Fever also results in an increase in the production of the iron-binding protein transferrin which reduces the availability of iron in the blood, which in turn can reduce the rate of growth of microbes. The beneficial effects of such an increase in body temperature disappear, however, should the value go over 41°C, as human proteins begin denaturing.

Watch this CrashCourse video to learn more about the immune system! Direct link: https://youtu.be/GIJK3dwCWCw

The Adaptive Immune Response

T lymphocytes and their Functional Types: Innate immune responses (and early induced responses) are in many cases ineffective at completely controlling pathogen growth. However, they slow pathogen growth and allow time for the adaptive immune response to strengthen, and either control or eliminate the pathogen. The innate immune system also sends signals to the cells of the adaptive immune system, guiding them in how to attack the pathogen. Thus, the innate and adaptive mechanisms are two important arms of the immune response.

The Benefits of the Adaptive Immune Response: The specificity of the adaptive immune response—its ability to specifically recognize and make a response against a wide variety of pathogens—is its great strength.

Antigens, the small chemical groups often associated with pathogens and their products, are recognized by receptors on the surfaces of B and T lymphocytes. The adaptive immune response to these antigens is so versatile that it can respond to nearly any pathogen. This increase in specificity comes because the adaptive immune response has a unique way to develop as many as 10^{11}, or 100 trillion, different receptors to recognize nearly every conceivable pathogen. Immunological memory is another benefit of adaptive immunity and this will be described further.

Primary Disease and Immunological Memory: The immune system's first exposure to a pathogen is called a **primary adaptive response**. Symptoms of a first infection, called primary disease, are always relatively severe because it takes time for an initial adaptive immune response to a pathogen to become effective.

Upon re-exposure to the same pathogen, a secondary adaptive immune response is generated, which is stronger and faster that the primary response. The **secondary adaptive response** often eliminates a pathogen before it can cause significant tissue damage or any symptoms. Without symptoms, there is no disease, and the individual is not even aware of the infection. This secondary response is the basis of **immunological memory**, which protects us from getting diseases repeatedly from the same pathogen. By this mechanism, an individual's exposure to pathogens early in life spares the person from these diseases later in life.

Self-Recognition: Another important feature of the adaptive immune response is its ability to distinguish between self-antigens, those that are normally present in the body, and foreign antigens, those that might be on a potential pathogen. As T and B cells mature, there are mechanisms in place that prevent them from recognizing self-antigen, preventing a damaging immune response against the body. These mechanisms are not 100 percent effective, however, and their breakdown leads to autoimmune diseases, which will be discussed later.

T Cell-Mediated Immune Responses: The primary cells that control the adaptive immune response are the lymphocytes, the T and B cells. T cells are particularly important, as they not only control a multitude of immune responses directly, but also control B cell immune responses in many cases as well. Thus, many of the decisions about how to attack a pathogen are made at the T cell level, and knowledge of their functional types is crucial to understanding the functioning and regulation of adaptive immune responses as a whole.

Antigens: Antigens on pathogens are usually large and complex, and can be either carbohydrate or protein based. Each T cell produces only one type of receptor and thus is specific for a single particular antigen. It is the interaction of the shape of the antigen and the complementary shape of the antigen-binding receptor that accounts for the chemical basis of specificity.

Antigen-presenting Cells: Antigen-presenting cells represent an important link between the innate and adaptive immune response. These stimulators of the adaptive response include macrophages, dendritic cells, and B cells. Macrophages stimulate T cells to release cytokines that enhance phagocytosis by macrophages. Dendritic cells also kill pathogens by phagocytosis, but both have the additional function of bringing antigens to regional draining lymph nodes. The lymph nodes are the locations in which most T cell responses against pathogens of the interstitial tissues are mounted. B cells may also present antigens to T cells, which are necessary for certain types of antibody responses, to be covered later in this chapter.

Mechanisms of T Cell-mediated Immune Responses: Mature T cells become activated by recognizing foreign antigen on an antigen presenting cell and begin dividing rapidly by mitosis. This proliferation of T cells is called **clonal expansion** and is necessary to make the immune response strong enough to effectively control a pathogen. Only those clones of lymphocytes whose receptors are activated by the antigen are stimulated to proliferate. Once activated, the selected clones increase in number and make many copies of each cell type, each clone with its unique receptor. By the time this process is complete, the body will have large numbers of specific lymphocytes available to fight the infection (see Figure 10).

The Cellular Basis of Immunological Memory: As already discussed, one of the major features of an adaptive immune response is the development of immunological memory.

During a primary adaptive immune response, both **memory T cells** and **effector T cells** are generated. Memory T cells are long-lived and can even persist for a lifetime. Memory cells are primed to act rapidly. Thus,

any subsequent exposure to the pathogen will elicit a very rapid T cell response. This rapid, secondary adaptive response generates large numbers of effector T cells so fast that the pathogen is often overwhelmed before it can cause any symptoms of disease. This is what is meant by immunity to a disease. The same pattern of primary and secondary immune responses occurs in B cells and the antibody response, as will be discussed later in the chapter.

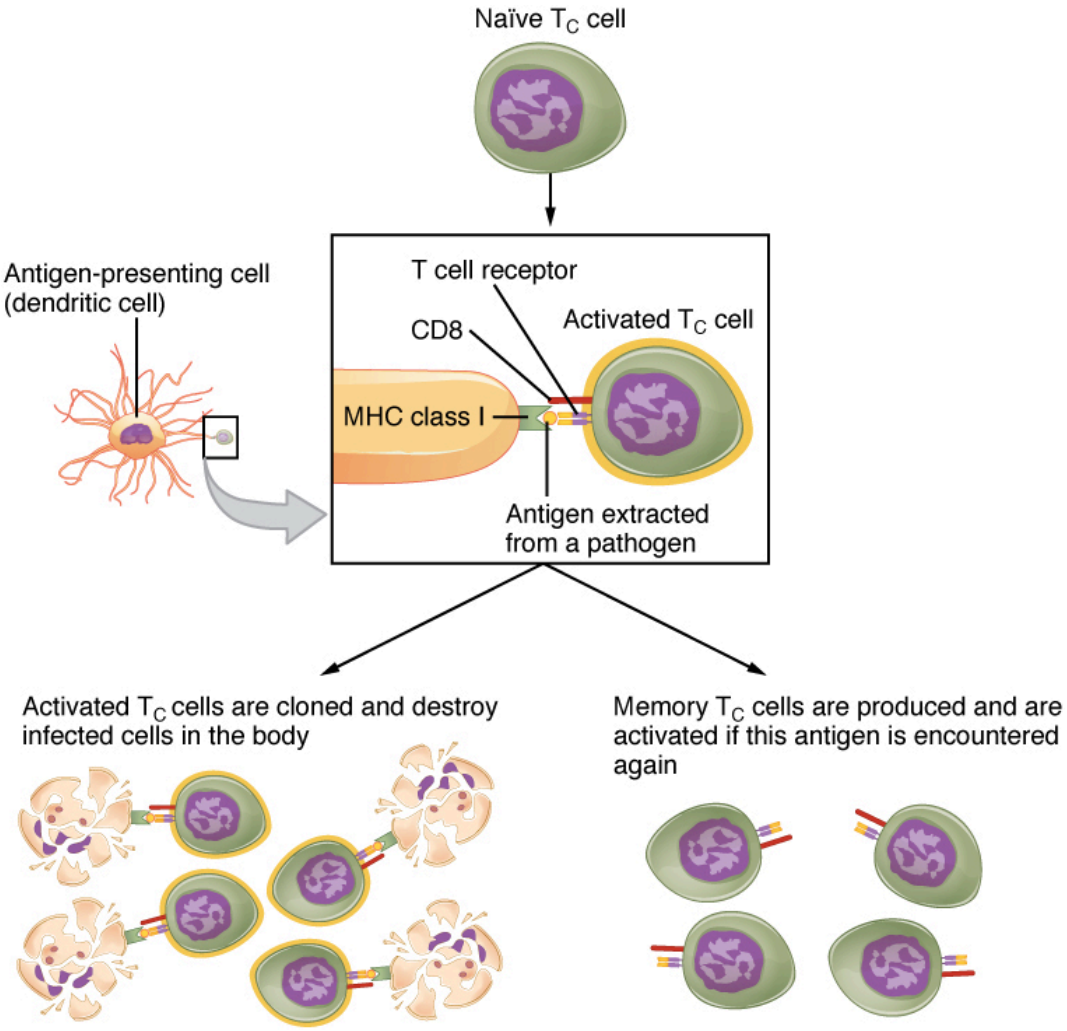


Figure 10. Clonal Selection and Expansion of T Lymphocytes. *Stem cells differentiate into T cells with specific receptors, the cells are called clones. The clones with receptors specific for antigens on the pathogen are selected for and expanded on encounter with the antigen.*

T Cell Types and their Functions: T cells can contain cell adhesion molecules that keep the T cell in close contact with the antigen-presenting cell by directly binding to the antigen-presenting receptor on its membrane. These markers are either CD4 or CD8 molecules (CD refers to Cluster of Differentiation) (Figure 11).

Although the correlation is not absolute, CD4-bearing T cells are associated with helper functions and CD8-bearing T cells are associated with cytotoxicity. These functional distinctions based on CD4 and CD8 markers are useful in defining the function of each type.

Helper T Cells and their Cytokines: Helper T cells (Th), bearing the CD4 molecule, function by secreting cytokines that act to enhance other immune responses. There are two classes of Th cells, and they act on different components of the immune response. These cells are not distinguished by their surface molecules but by the characteristic set of cytokines they secrete.

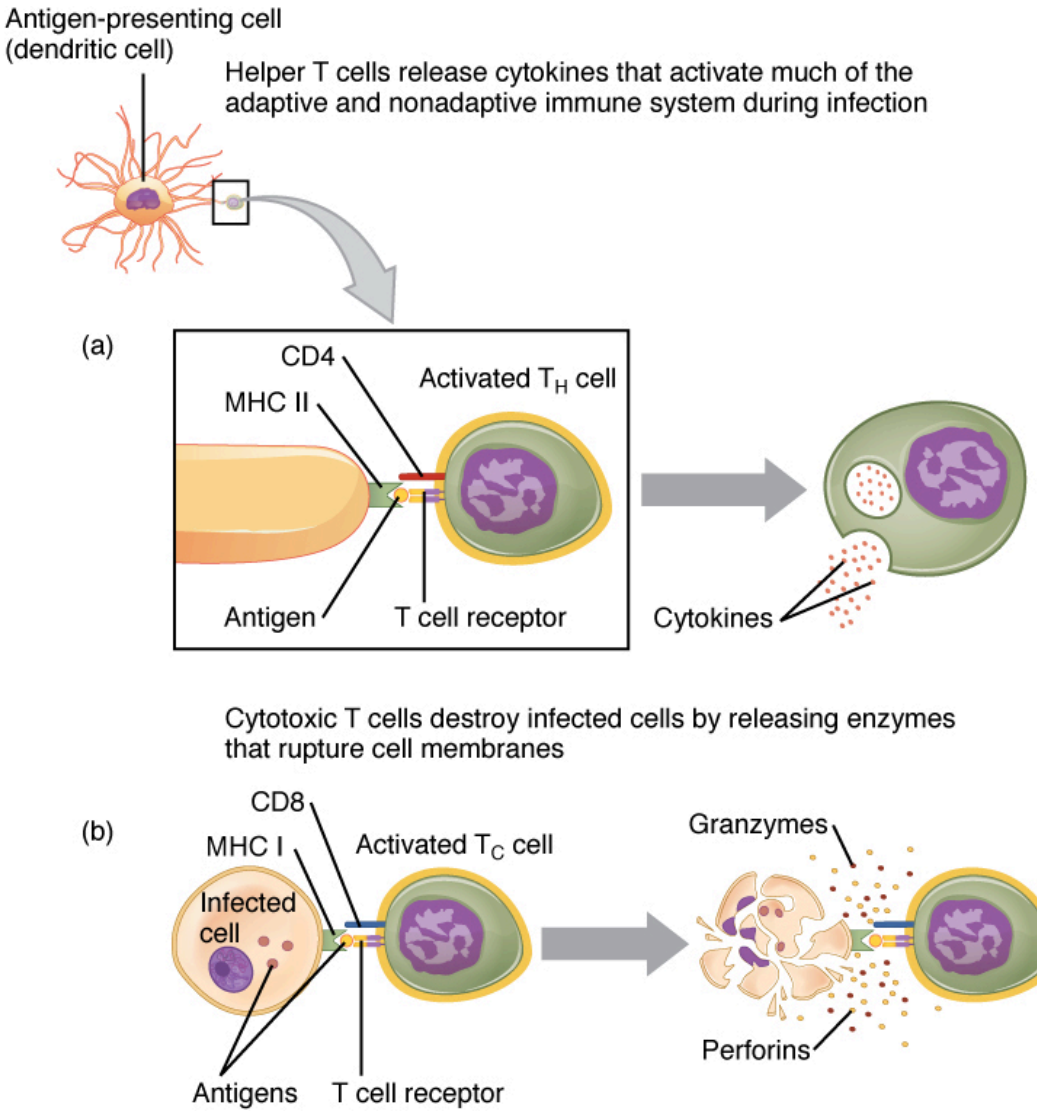


Figure 11. Antigen Presentation. *(a) CD4 is associated with helper and regulatory T cells. An extracellular antigen is processed and presented in the binding cleft of a class II MHC molecule, and this interaction is strengthened by the CD4 molecule. (b) CD8 is associated with cytotoxic T cells. An antigen of an intracellular pathogen is presented by a class I MHC molecule, and CD8 interacts with it.*

Th1 cells are a type of helper T cell that secretes cytokines that regulate the immunological activity and development of a variety of cells, including macrophages and other types of T cells.

Th2 cells, on the other hand, are cytokine-secreting cells that act on B cells to drive their differentiation into plasma cells that make antibody.

Cytotoxic T cells: Cytotoxic T cells (Tc) are T cells that kill target specific cells by inducing apoptosis using the same mechanism as NK cells. In addition, as long as the antigen is recognized by the cell, each Tc cell can kill more than one target cell, making them especially effective. While they are active in all pathogenic infections, Tc cells are so important in the antiviral immune response that some speculate that this was the main reason the adaptive immune response evolved in the first place.

B-lymphocytes and Antibodies: Antibodies were the first component of the adaptive immune response to be characterized by scientists working on the immune system. It was already known that individuals who survived a bacterial infection were immune to re-infection with the same pathogen. Early microbiologists took serum from an immune patient and mixed it with a fresh culture of the same type of bacteria, then observed the bacteria under a microscope. The bacteria became clumped in a process called agglutination. When a different bacterial species was used, the agglutination did not happen. Thus, there was something in the serum of

immune individuals that could specifically bind to and agglutinate bacteria. Scientists now know the cause of the agglutination is an antibody molecule, also called an **immunoglobulin**.

What is an antibody? An antibody protein is essentially a secreted form of a B cell receptor. (In fact, surface immunoglobulin is another name for the B cell receptor.) Not surprisingly, the same genes encode both the secreted antibodies and the surface immunoglobulins. One minor difference in the way these proteins are synthesized distinguishes a naïve B cell with antibody on its surface from an antibody-secreting plasma cell with no antibodies on its surface. The antibodies of the plasma cell have the exact same antigen-binding site and specificity as their B cell precursors.

There are five different classes of antibody (also called immunoglobulin Ig) found in humans: IgM, IgD, IgG, IgA, and IgE. Each of these has specific functions in the immune response, so by learning about them, researchers can learn about the great variety of antibody functions critical to many adaptive immune responses.

B Cell Differentiation and Activation: B cells differentiate in the bone marrow. During the process of maturation, up to 100 trillion different clones of B cells are generated, which is similar to the diversity of antigen receptors seen in T cells.

After B cells are activated by their binding to antigen, they differentiate into plasma cells. Plasma cells often leave the secondary lymphoid organs, where the response is generated, and migrate back to the bone marrow, where the whole differentiation process from a lymphoid progenitor cell started. After secreting antibodies for a specific period, plasma cells die, as most of their energy is devoted to making antibodies and not to maintaining themselves. Thus, plasma cells are said to be terminally differentiated.

The final B cell of interest is the memory B cell, which results from the clonal expansion of an activated B cell. Memory B cells function in a way similar to memory T cells. They lead to a stronger and faster secondary response when compared to the primary response, as illustrated below.

Antibody Structure: Antibodies are glycoproteins consisting of two types of polypeptide chains with attached carbohydrates. The **heavy chain** and the **light chain** are the two polypeptides that form the antibody, two of each are required to form a generic antibody structure. The main differences between the classes of antibodies are in the differences between their heavy chains, but as you shall see, the light chains have an important role, forming part of the antigen-binding site on the antibody molecules (Figure 12).

Functions of Antibodies: In general, antibodies have two basic functions. They can act as the B cell antigen receptor or they can be secreted, circulate, and bind to a pathogen, often labeling it for identification by other forms of the immune response.

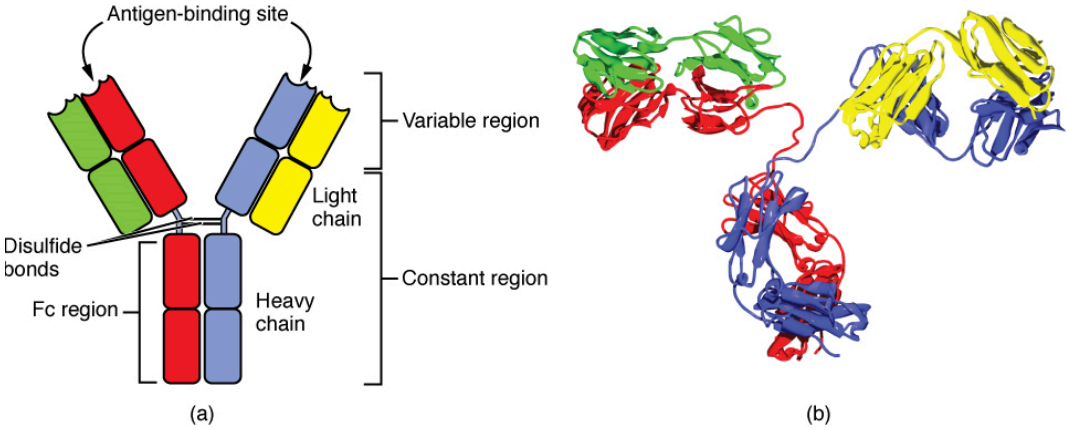


Figure 12. Antibody Structure. *The typical four chain structure of a generic antibody (a) and the corresponding three-dimensional structure of the antibody IgG2 (b). (credit b: modification of work by Tim Vickers)*

Effects of Antibody-Antigen Binding: Antibodies that bind to antigens can lead to a number of different outcomes, depending on the nature of the antigen and the structure of the antibody. In a process called **neutralization**, antibodies bind to antigens on the surface of some viruses, or to toxins secreted by bacteria,

in a way that prevents them from negatively affecting body cells. The antibodies neutralize the pathogen or toxin by physically covering up the dangerous parts so it cannot damage body cells. Antibodies have at least two antigen-binding sites and therefore they can bind to antigen on the surface of two or more cells or to multiple molecules of a soluble antigen or toxin, clumping whole cells together in a process known as **agglutination**, or causing soluble antigen molecules to clump together and **precipitate** out of solution. Neutralization, agglutination, and precipitation of antigens all enhance the likelihood that phagocytotic cells will engulf the antigen (or antigen-bearing cell).

An antibody bound to an antigen molecule on the surface of a pathogen can enhance the **phagocytosis** of the pathogen. It can also fix and activate the **complement system**, a series of signaling cascades which lead to an enhancement of phagocytosis, a local inflammatory response, and lysis of the pathogen.

Clonal Selection of B Cells: Clonal selection and expansion work much the same way in B cells as in T cells. Only B cells with appropriate antigen specificity are selected for and expanded (Figure 13). Eventually, the plasma cells secrete antibodies with antigenic specificity identical to those that were on the surfaces of the selected B cells. Notice in the figure that both plasma cells and memory B cells are generated simultaneously.

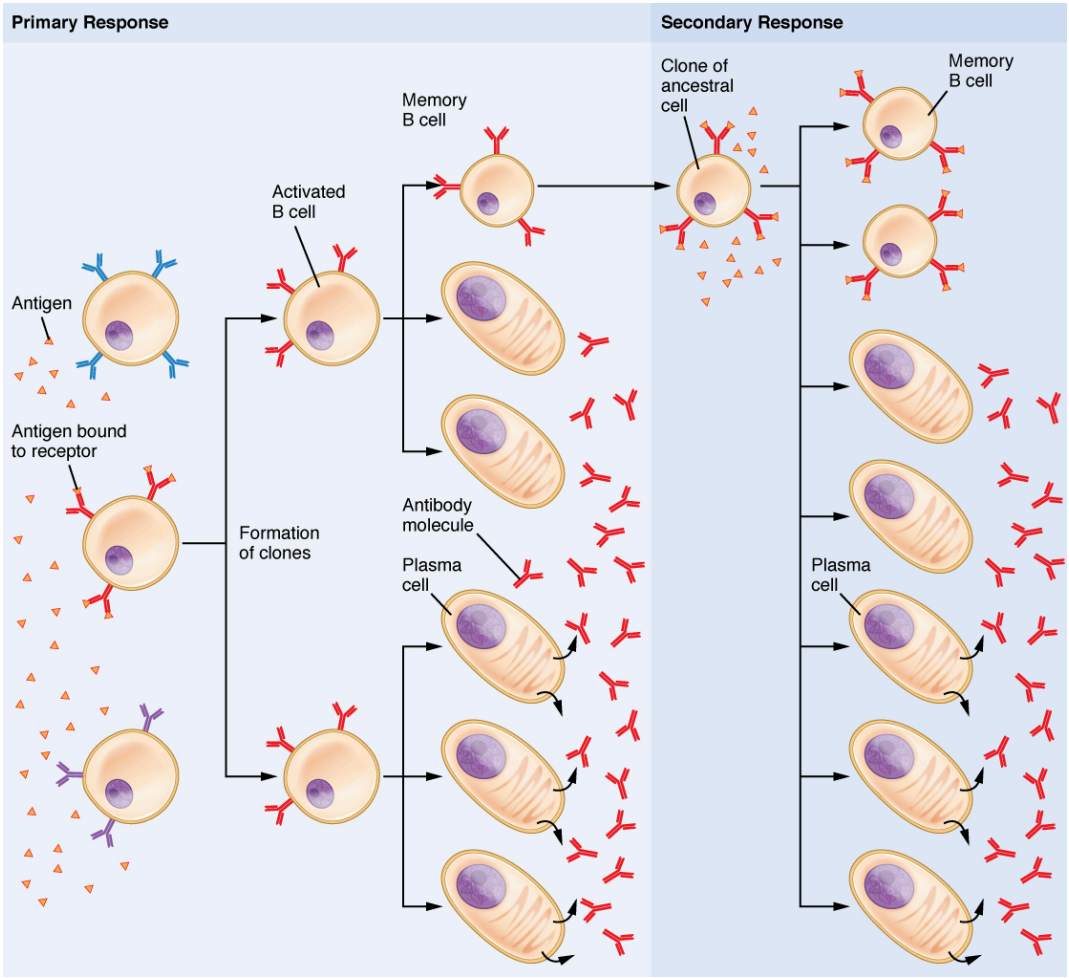


Figure 13. Clonal Selection of B Cells. *During a primary B cell immune response, both antibody-secreting plasma cells and memory B cells are produced. These memory cells lead to the differentiation of more plasma cells and memory B cells during secondary responses.*

Primary versus Secondary B Cell Responses: Primary and secondary responses as they relate to T cells were discussed earlier. This section will look at these responses with B cells and antibody production. Because antibodies are easily measured in blood samples, their concentrations are easy to follow and graph (Figure 14). As you will see from the figure, the primary response to an antigen (representing a pathogen) is delayed by several days. This is the time it takes for the B cell clones to expand and differentiate into plasma cells. The level

of antibody produced is low, but it is sufficient for immune protection. The second time a person encounters the same antigen, there is no time delay, and the amount of antibody made is much higher. Thus, the secondary antibody response overwhelms the pathogens quickly and, in most situations, no symptoms are felt. When a different antigen is used, another primary response is made with its low antibody levels and time delay.

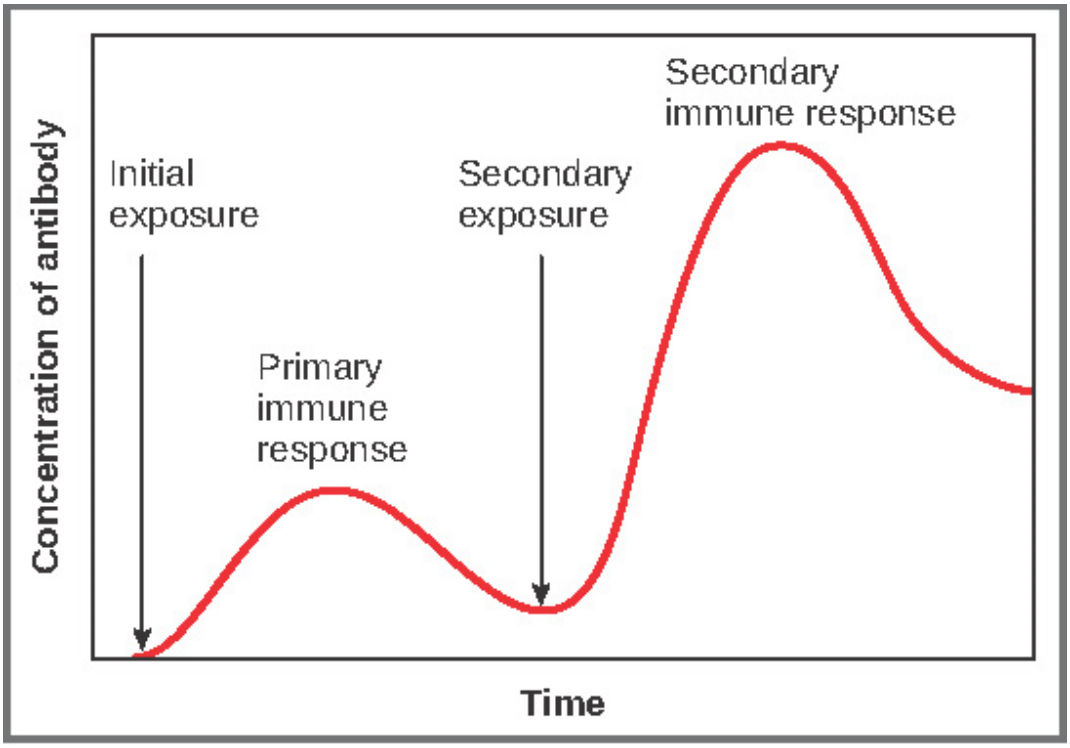


Figure 14. Primary and Secondary Antibody Responses. *Antigen A is given once to generate a primary response and later to generate a secondary response. When a different antigen is given for the first time, a new primary response is made.*

Active versus Passive Immunity: Immunity to pathogens, and the ability to control pathogen growth so that damage to the tissues of the body is limited, can be acquired by (1) the active development of an immune response in the infected individual or (2) the passive transfer of immune components from an immune individual to a nonimmune one. Both active and passive immunity have examples in the natural world and as part of medicine.

Active immunity is the resistance to pathogens acquired during an adaptive immune response within an individual (Figure 15). Naturally acquired active immunity, the response to a pathogen, is the focus of this section. Artificially acquired active immunity involves the use of vaccines. A vaccine is a killed or weakened (attenuated) pathogen or its components that, when administered to a healthy individual, leads to the development of immunological memory (via a weakened primary immune response) without causing much in the way of symptoms. **Killed vaccines** or **inactivated vaccines** consist of pathogens that have been killed and so are no longer viable, but still retain antigens that can be recognized and used to mount an immune response. **Live attenuated vaccines** are generally used when the pathogen involved does not trigger an immune response when introduced in a killed or inactivated state. The virus that causes measles, for example, when introduced in an inactivated form does not confer lasting immunity to measles. However, this same virus can be left viable but modified to render it incapable of producing the symptoms of measles, while still triggering an appropriate immune response. The pathogens contained in a live attenuated vaccine remain viable, but have been rendered harmless or less virulent. **Toxoid vaccines** include a toxin molecule that has been modified to be harmless but still elicits an immune response against the toxin. The tetanus vaccine for example contains a modified version of the toxin tetanospasmin that is normally released by the bacterium *Clostridium*

tetani. This vaccine triggers the production of anti-tetanospasmin antibodies that confer immunity to the live bacterium's harmful effects.

A person can also acquire protection from specific pathogens through the administration of pre-formed antibodies**,** known as **passive immunity**. Naturally acquired passive immunity is represented by the transfer of antibodies in breast milk or through the placenta that give new born babies protection against some pathogens as they are developing their own immune response. In medicine, artificially acquired passive immunity usually involves injections of immunoglobulins taken from animals previously exposed to a specific pathogen. This treatment is a fast-acting method of temporarily protecting an individual who was possibly exposed to a pathogen. The downside to this treatment is the lack of the development of immunological memory. Once the antibodies are transferred, they are effective for only a limited time before they degrade, so multiple injections may be necessary. An example is the treatment of suspected rabies with postexposure prophylaxis that includes immunoglobulin injections.

From the above, it is readily apparent that with the use of vaccines, one can avoid the damage from disease that results from the first exposure to the pathogen, yet reap the benefits of protection from immunological memory. The advent of vaccines was one of the major medical advances of the twentieth century and led to the eradication of smallpox and the control of many infectious diseases, including polio, measles, and whooping cough.

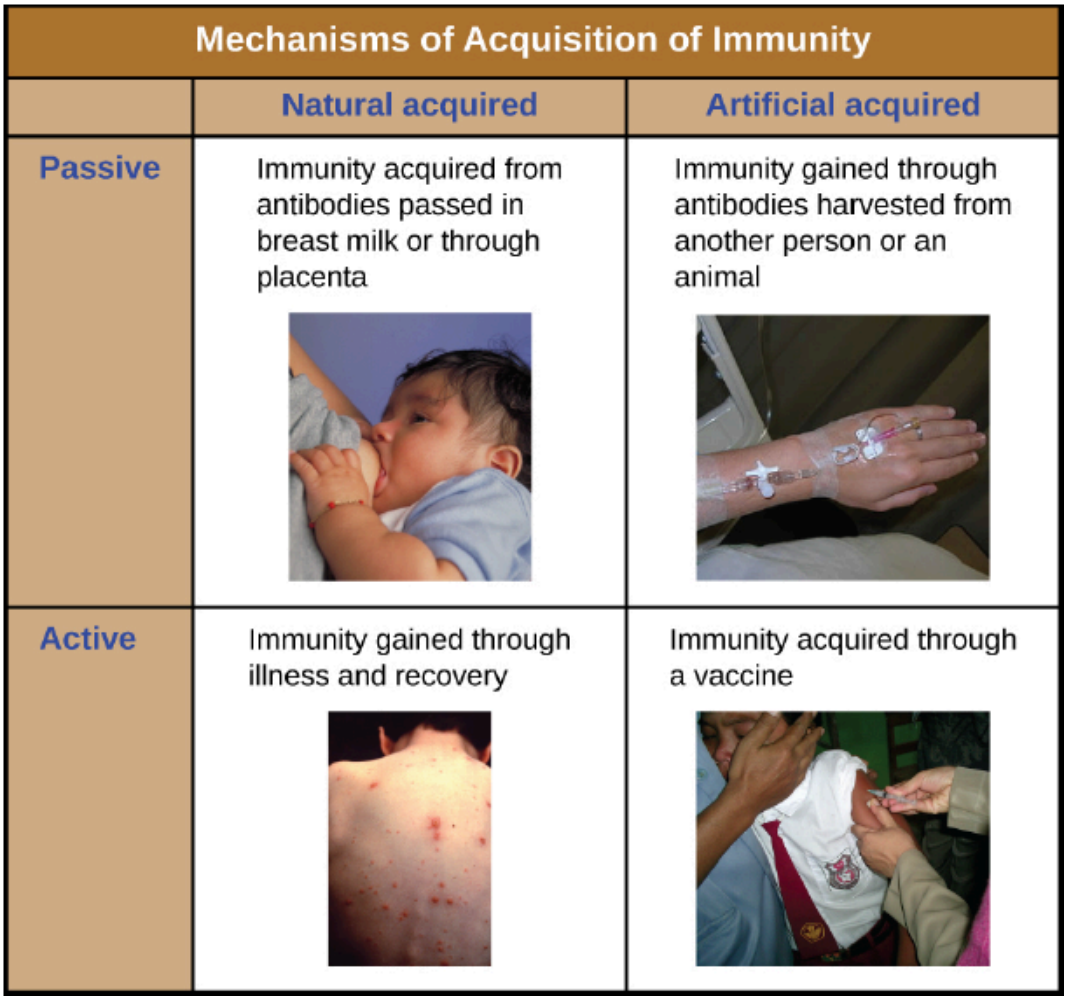


Figure 15. Classification of Acquired Immunity. *(credit top left photo: modification of work by USDA; credit top right photo: modification of work by "Michaelberry"/Wikimedia; credit bottom left photo: modification of work by Centers for Disease Control and Prevention; credit bottom right photo: modification of work by Friskila Silitonga, Indonesia, Centers for Disease Control and Prevention)*

Watch this Crash Course video for an overview of the adaptive immune response! Direct link: https://youtu.be/2DFN4IBZ3rI

Diseases Associated with Depressed or Overactive Immune Responses: This section is about how the immune system goes wrong. When it goes haywire, and becomes too weak or too strong, it leads to a state of disease. The factors that maintain immunological homeostasis are complex and incompletely understood.

Immunodeficiencies: As you have seen, the immune system is quite complex. It has many pathways using many cell types and signals. Because it is so complex, there are many ways for it to go wrong, and in the case of immunodeficiencies, become weakened. Inherited immunodeficiencies arise from gene mutations that affect specific components of the immune response. There are also acquired immunodeficiencies that result from causes other than inheritance with potentially devastating effects on the immune system, such as infection with HIV.

Inherited Immunodeficiencies: A list of all inherited immunodeficiencies is well beyond the scope of this book. The list is almost as long as the list of cells, proteins, and signaling molecules of the immune system itself. Some deficiencies, such as those for complement, cause only a higher susceptibility to some Gram-negative bacteria. Others are more severe in their consequences. Certainly, the most serious of the inherited immunodeficiencies is **severe combined immunodeficiency disease (SCID)**. This disease is complex because it is caused by many different genetic defects. What groups them together is the fact that both the B cell and T cell arms of the adaptive immune response are affected. Children with this disease usually die of opportunistic infections within their first year of life unless they receive a bone marrow transplant.

Human Immunodeficiency Virus/AIDS: Although many viruses cause suppression of the immune system, only one wipes it out completely, and that is the previously mentioned HIV. The virus is transmitted through semen, vaginal fluids, and blood. There are sometimes, but not always, flu-like symptoms in the first 1 to 2 weeks after infection. Following this time (with no medical intervention), the levels of CD4+ cells, especially helper T cells, decline steadily, eventually producing an acquired immunodeficiency syndrome (AIDS), until at some point, the immune response is so weak that opportunistic disease and eventually death result. CD4 is the receptor that HIV uses to get inside T cells and reproduce. Given that CD4+ helper T cells play an important role in other in T cell immune responses and antibody responses, it should be no surprise that both types of cellular and humoral immune responses are eventually seriously compromised.

Treatment for the disease consists of drugs that target virally encoded proteins that are necessary for viral replication but are absent from normal human cells. By targeting the virus itself and sparing the cells, this approach has been successful in significantly prolonging the lives of HIV-positive individuals. On the other hand, an HIV vaccine has been 30 years in development and is still years away. Because the virus mutates rapidly to evade the immune system, scientists have been looking for parts of the virus that do not change and thus would be good targets for a vaccine candidate.

Autoimmune Responses: The worst cases of the immune system over-reacting are autoimmune diseases. Somehow, tolerance breaks down and the immune systems in individuals with these diseases begin to attack their own bodies, causing significant damage. The trigger for these diseases is, more often than not, unknown, and the treatments are usually based on resolving the symptoms using immunosuppressive and anti-inflammatory drugs such as steroids. These diseases can be localized and crippling, as in rheumatoid arthritis,

or diffuse in the body with multiple symptoms that differ in different individuals, as is the case with systemic lupus erythematosus (Figure 16).

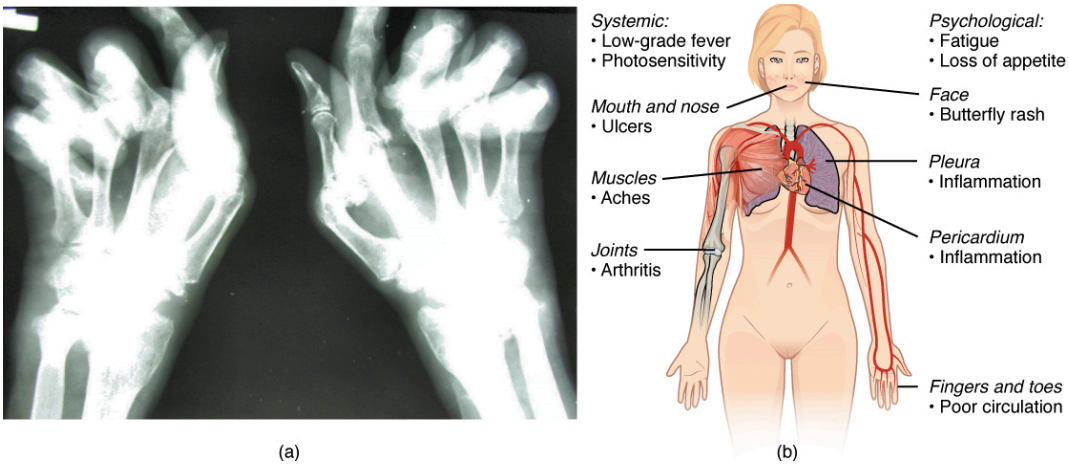


Figure 16. Autoimmune Disorders: Rheumatoid Arthritis and Lupus. *(a) Extensive damage to the right hand of a rheumatoid arthritis sufferer is shown in the x-ray. (b) The diagram shows a variety of possible symptoms of systemic lupus erythematosus.*

Part 3: Blood Typing

Blood transfusions in humans were risky procedures until the discovery of the major human blood groups by Karl Landsteiner, an Austrian biologist and physician, in 1900. Until that point, physicians did not understand that death sometimes followed blood transfusions, when the type of donor blood infused into the patient was incompatible with the patient's own blood. Blood groups are determined by the presence or absence of specific marker molecules on the plasma membranes of erythrocytes. With their discovery, it became possible for the first time to match patient-donor blood types and prevent transfusion reactions and deaths.

Antigens, Antibodies, and Transfusion Reactions: Antigens are substances that the body does not recognize as belonging to the "self" and therefore trigger a defensive response from the leukocytes of the immune system. Here, we will focus on the role of immunity in blood transfusion reactions. Following an infusion of incompatible blood, erythrocytes with foreign antigens appear in the bloodstream and trigger an immune response. Antibodies produced by the plasma cells, attach to the antigens on the plasma membranes of the infused erythrocytes and cause them to adhere to one another.

- As explained before, because the arms of the Y-shaped antibodies attach randomly to more than one non-self erythrocyte surface, they form clumps of erythrocytes (**agglutination**).
- The clumps of erythrocytes block small blood vessels throughout the body, depriving tissues of oxygen and nutrients.
- As the erythrocyte clumps are degraded, in a process called **hemolysis**, their hemoglobin is released into the bloodstream. This hemoglobin travels to the kidneys, which are responsible for filtration of the blood. However, the load of hemoglobin released can easily overwhelm the kidney's capacity to clear it, and the patient can quickly develop kidney failure.

More than 50 antigens have been identified on erythrocyte membranes, but the most significant in terms of their potential harm to patients are classified in two groups: the ABO blood group and the Rh blood group.

The ABO Blood Group: Although the **ABO blood group** name consists of three letters, ABO blood typing designates the presence or absence of just two antigens, A and B. Both are glycoproteins. People who have A antigens on their erythrocyte membrane surfaces are designated blood type A, and those whose erythrocytes have B antigens are blood type B. People can also have both A and B antigens on their erythrocytes, in which case they are blood type AB. People with neither A nor B antigens are designated blood type O. ABO blood types are genetically determined.

Normally the body must be exposed to a foreign antigen before an antibody can be produced. This is not the case for the ABO blood group. Individuals with type A blood—without any prior exposure to incompatible blood—have pre-formed antibodies to the B antigen circulating in their blood plasma. These antibodies, referred to as anti-B antibodies, will cause agglutination and hemolysis if they ever encounter erythrocytes with B antigens. Similarly, an individual with type B blood has pre-formed anti-A antibodies. Individuals with type AB blood, which has both antigens, do not have pre-formed antibodies to either of these. People with type O blood lack antigens A and B on their erythrocytes, but both pre-formed anti-A and anti-B antibodies circulate in their blood plasma.

Rh Blood Groups: The **Rh blood group** is classified according to the presence or absence of a second erythrocyte antigen identified as Rh. (It was first discovered in a type of primate known as a rhesus macaque, which is often used in research, because its blood is similar to that of humans.) Although dozens of Rh antigens have been identified, only one, designated D, is clinically important. Those who have the Rh D antigen present on their erythrocytes—about 85 percent of Americans—are described as Rh positive (Rh+) and those who lack it are Rh negative (Rh−). Note that the Rh group is distinct from the ABO group, so any individual, no matter their ABO blood type, may have or lack this Rh antigen. When identifying a patient's blood type, the Rh group is designated by adding the word positive or negative to the ABO type. For example, A positive (A+) means ABO group A blood with the Rh antigen present, and AB negative (AB−) means ABO group AB blood without the Rh antigen.

In contrast to the ABO group antibodies, which are preformed, antibodies to the Rh antigen are produced only in Rh− individuals after exposure to the antigen. This process, called sensitization, occurs following a transfusion with Rh-incompatible blood or, more commonly, with the birth of an Rh+ baby to an Rh− mother. Problems are rare in a first pregnancy, since the baby's Rh+ cells rarely cross the placenta (the organ of gas and nutrient exchange between the baby and the mother). However, during or immediately after birth, the Rh− mother can be exposed to the baby's Rh+ cells (Figure 17). Research has shown that this occurs in about 13–14 percent of such pregnancies. After exposure, the mother's immune system begins to generate anti-Rh antibodies. If the mother should then conceive another Rh+ baby, the Rh antibodies she has produced can cross the placenta into the fetal bloodstream and destroy the fetal RBCs. This condition, known as **hemolytic disease of the newborn (HDN**) or erythroblastosis fetalis, may cause anemia in mild cases, but the agglutination and hemolysis can be so severe that without treatment the fetus may die in the womb or shortly after birth.

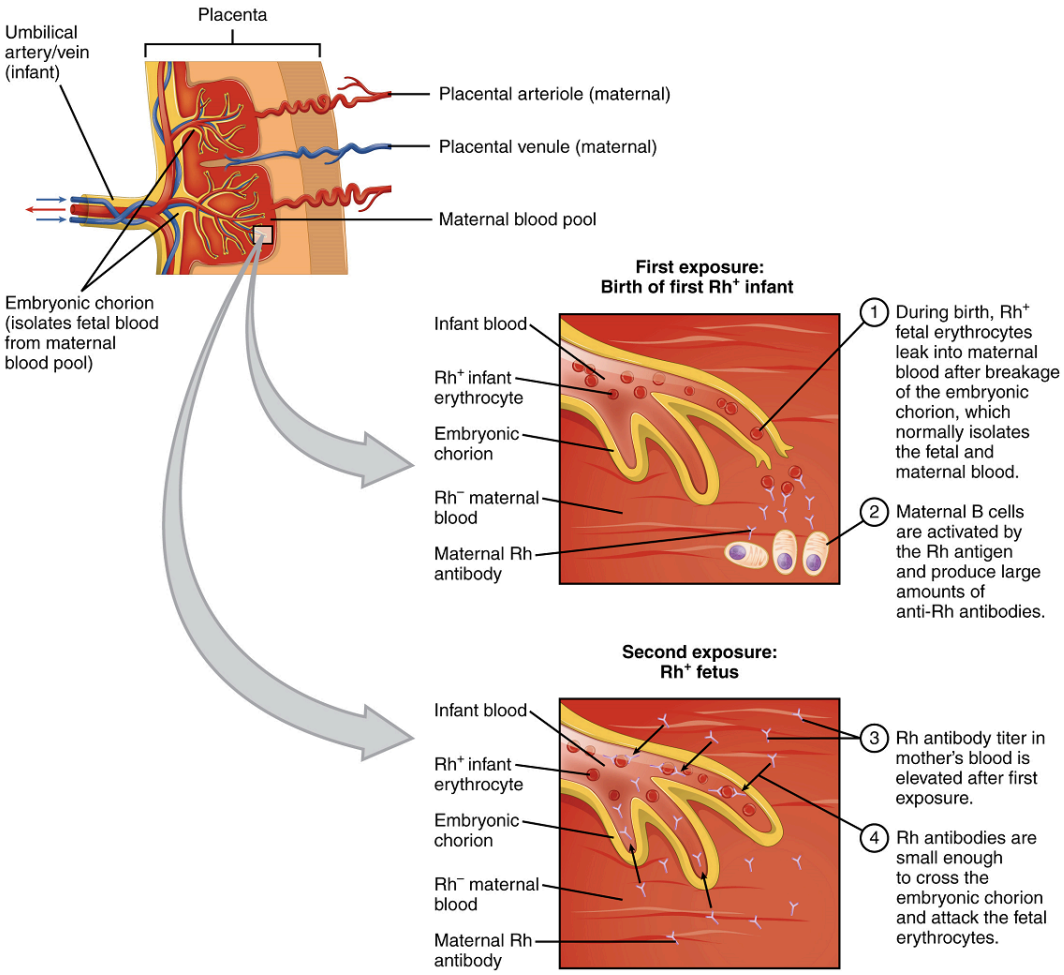


Figure 17. Erythroblastosis Fetalis. *The first exposure of an Rh− mother to Rh+ erythrocytes during pregnancy induces sensitization. Anti-Rh antibodies begin to circulate in the mother's bloodstream. A second exposure occurs with a subsequent pregnancy with an Rh+ fetus in the uterus. Maternal anti-Rh antibodies may cross the placenta and enter the fetal bloodstream, causing agglutination and hemolysis of fetal erythrocytes.*

A drug known as RhoGAM, short for Rh immune globulin, can temporarily prevent the development of Rh antibodies in the Rh− mother, thereby averting this potentially serious disease for the fetus. RhoGAM antibodies destroy any fetal Rh+ erythrocytes that may cross the placental barrier. RhoGAM is normally administered to Rh− mothers during weeks 26–28 of pregnancy and within 72 hours following birth. It has proven remarkably effective in decreasing the incidence of HDN. Earlier we noted that the incidence of HDN in an Rh+ subsequent pregnancy to an Rh− mother is about 13–14 percent without preventive treatment. Since the introduction of RhoGAM in 1968, the incidence has dropped to about 0.1 percent in the United States.

Determining ABO Blood Types: Clinicians are able to determine a patient's blood type quickly and easily using commercially prepared antibodies. An unknown blood sample is allocated into separate wells. Into one well a small amount of anti-A antibody is added, and to another a small amount of anti-B antibody. If the antigen is present, the antibodies will cause visible agglutination of the cells (Figure 18). The blood should also be tested with Rh antibodies.

ABO Transfusion Protocols: To avoid transfusion reactions, it is best to transfuse only matching blood types; that is, a type B+ recipient should ideally receive blood only from a type B+ donor and so on. That said, in emergency situations, when acute hemorrhage threatens the patient's life, there may not be time for cross matching to identify blood type. In these cases, blood from a **universal donor**—an individual with type O− blood—may be transfused. Recall that type O erythrocytes do not display A or B antigens. Thus, anti-A or anti-B antibodies that might be circulating in the patient's blood plasma will not encounter any erythrocyte surface antigens on the donated blood and therefore will not be provoked into a response.

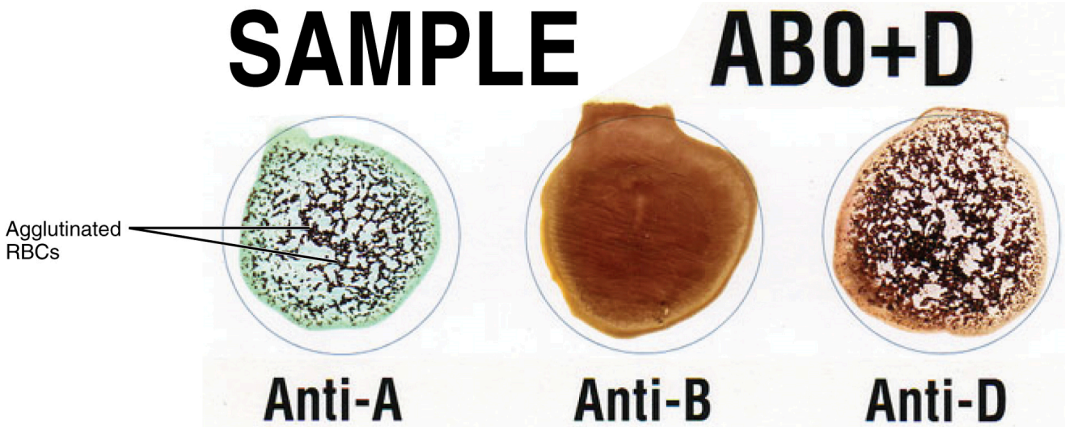


Figure 18. Cross Matching Blood Types. *This sample of a commercially produced "bedside" card enables quick typing of both a recipient's and donor's blood before transfusion. The card contains three reaction sites or wells. One is coated with an anti-A antibody, one with an anti-B antibody, and one with an anti-D antibody (tests for the presence of Rh factor D). Mixing a drop of blood and saline into each well enables the blood to interact with a preparation of type-specific antibodies, also called anti-sera. Agglutination of RBCs in a given site indicates a positive identification of the blood antigens, in this case A and Rh antigens for blood type A+. For the purpose of transfusion, the donor's and recipient's blood types must match.*

A patient with blood type AB+ is known as the universal recipient. This patient can theoretically receive any type of blood, because the patient's own blood—having both A and B antigens on the erythrocyte surface—does not produce anti-A or anti-B antibodies. In addition, an Rh+ patient can receive both Rh+ and Rh− blood. Figure 19 summarizes the blood types and compatibilities.

Blood Type

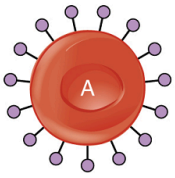
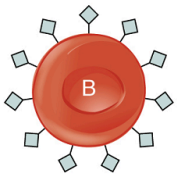
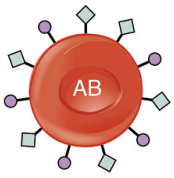
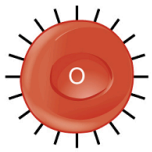
	A	B	AB	O
Red Blood Cell Type				
Antibodies in Plasma	 Anti-B	 Anti-A	None	 Anti-A and Anti-B
Antigens in Red blood Cell	 A antigen	 B antigen	 A and B antigens	None
Blood Types Compatible in an Emergency	A, O	B, O	A, B, AB, O (AB^+ is the universal recipient)	O (O is the universal donor)

Figure 19. ABO Blood Group. *This chart summarizes the characteristics of the blood types in the ABO blood group. See the text for more on the concept of a universal donor or recipient.*

Unit 4: The Respiratory System

Unit outline

Part 1: Anatomy of the Respiratory System

- Conducting zone
- Respiratory zone
- The Lungs

Part 2: The Process of Breathing – Pulmonary Ventilation

- Pulmonary ventilation
- Nervous control of ventilation
- Factors that affect the rate and depth of ventilation
- Respiratory volumes and capacities

Part 3: Gas Exchange

- External respiration
- Internal respiration

Part 4: Transport of Gases

- Oxygen transport in blood
- Factors affecting oxygen dissociation from hemoglobin
- Carbon dioxide transport in the blood

Part 5: Modifications in Respiratory Functions

- Hyperventilation
- Hypoxia

Learning Objectives

At the end of this unit, you should be able to:

I. Describe the location and function(s) of the major components of the human respiratory system.

II. Explain the mechanism of ventilation (inspiration and expiration) in humans, including the roles of the structures involved.

III. Describe the nervous control of breathing.

IV. Describe how carbon dioxide, oxygen and hydrogen ions control the rate of breathing.

V. Interpret a spirogram and define the respiratory volumes and capacities a spirogram depicts.

VI. Explain the basic principle governing the reciprocal exchange of gases between the alveoli and the blood, and between the blood and individual cells.

VII. Describe the mechanisms by which oxygen and carbon dioxide are transported in the blood.

VIII. Define hyperventilation and specify some of its causes and physiological consequences.

IX. Define hypoxia and specify some of its causes and physiological consequences.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Describe the location and function(s) of the major components of the human respiratory system.

1. Draw a series of fully-annotated diagrams showing the structure of the respiratory system. Include the location, general structure and function of all the following components:
    - External nares
    - Nasal cavity
    - Internal nares
    - Oral cavity
    - Pharynx
    - Epiglottis
    - Glottis
    - Larynx
    - Trachea
    - Bronchus
    - Bronchiole
    - Alveolus
    - Visceral pleura
    - Parietal pleura
    - Pleural cavity

II. Explain the mechanism of ventilation (inspiration and expiration) in humans, including the roles of the structures involved.

1. Describe the mechanism of pulmonary ventilation in humans, including a detailed description of the processes of quiet inspiration (inhalation) and quiet expiration (exhalation). Refer in your explanation to:

- The skeletal muscles required for quiet inhalation
- The two passive processes required for quiet exhalation
- Changes in the volume of the thoracic cavity at each step
- Changes in the intra-alveolar pressure at each step
- The direction of air travel at each step

III. Describe the nervous control of breathing.

1. Describe how breathing is modulated by the nervous system, making reference to all of the following structures:

    - Diaphragm
    - External intercostals
    - Internal intercostals
    - Pons
    - Medulla oblongata
    - Pontine respiratory group
    - Apneustic center
    - Pneumotaxic center
    - Ventral respiratory group
    - Dorsal respiratory group
    - Cerebral cortex
    - Peripheral chemoreceptors
    - Central chemoreceptors

IV. Describe how carbon dioxide, oxygen and hydrogen ions control the rate of breathing.

1. Describe the mechanisms by which each of the following influence the rate of breathing:

    - Partial pressure (or concentration) of carbon dioxide gas in the blood
    - Partial pressure (or concentration of oxygen gas in the blood
    - Concentration of hydrogen ions (protons) in the blood

V. Interpret a spirogram and define the respiratory volumes and capacities a spirogram depicts.

1. What are spirograms used for?
2. Sketch a spirogram showing several normal breaths followed by one deep inspiration and one forced exhalation.
3. Identify the following measurements on a spirogram, and describe what each measurement represents physiologically:

    - Tidal volume
    - Vital capacity
    - Residual volume

4. Can “dead space volume” be measured from a spirogram? Briefly justify your answer.

VI. Explain the basic principle governing the reciprocal exchange of gases between the alveoli and the blood, and between the blood and individual cells.

1. Define and clearly distinguish oxygenated and deoxygenated blood.
2. In the systemic circulation, which type of blood vessels normally contain 'deoxygenated blood'?
3. In the systemic circulation, which type of blood vessels normally contain 'oxygenated blood'?
4. In the pulmonary circulation, name the blood vessels which normally contain 'deoxygenated blood'.
5. In the pulmonary circulation, name the blood vessels which normally contain 'oxygenated blood'?
6. By what transport mechanism do oxygen (O_2) and carbon dioxide (CO_2) gas move across plasma membranes? Explain why this particular transport mechanism is used by oxygen and carbon dioxide gas.
7. What condition drives movement of gases in a certain direction during gas exchange?
8. What is the difference in the driving force that promotes pulmonary ventilation compared to the driving force which promotes gas exchange?
9. Explain why significant gas exchange in the lungs can only occur in the alveoli, and not in the bronchi or bronchioles.

VII. Describe the mechanisms by which oxygen and carbon dioxide are transported in the blood.

1. Describe the two mechanisms by which oxygen gas is transported within the blood, and the relative quantities transported by each mechanism.
2. Describe the three mechanisms by which carbon dioxide is transported within the blood, and the relative quantities transported by each mechanism.
3. Specify the components of blood that contain:
    - Dissolved oxygen gas
    - Dissolved carbon dioxide gas
    - Oxygen molecules bound to hemoglobin
    - Carbon dioxide bound to hemoglobin
    - Bicarbonate ions
4. For each of the following, describe how it influences the amount of oxygen bound to hemoglobin, and describe how it relates to the ability of hemoglobin to pick up or release oxygen at appropriate locations in the human body:
    - Blood pH
    - Partial pressure (or concentration) of carbon dioxide in blood
    - Blood temperature

VIII. Define hyperventilation and specify some of its causes and physiological consequences.

1. Clearly and precisely define "hyperventilation".
2. Describe and explain the effects of hyperventilation on:
    - Blood oxygen content
    - Blood carbon dioxide content
    - Blood pH
    - Blood pressure
    - Brain (neuron) function

3. State the two major causes of hyperventilation and provide at least two specific examples of conditions that could underlie each.

IX. Define hypoxia and specify some of its causes and physiological consequences.

1. Clearly and precisely define “hypoxia”.
2. For each of the following provide a specific example of a condition or circumstance that would cause it and explain how it causes “hypoxia”.
    - A deficiency in atmospheric oxygen
    - Ventilatory deficiency
    - Diffusion deficiency in the lungs
    - Deficiency of hemoglobin (anemic hypoxia)
    - Circulatory deficiency (ischemic hypoxia)
    - Edema
3. State three physiological consequences of hypoxia.

The Respiratory System: Hold your breath. Really! See how long you can hold your breath as you continue reading…How long can you do it? Chances are you are feeling uncomfortable already. A typical human cannot survive without breathing for more than 3 minutes, and even if you wanted to hold your breath longer, your autonomic nervous system would take control. This is because most cells in the body run the oxidative stages of cellular respiration, the process by which energy is produced in the form of adenosine triphosphate (ATP). For oxidative phosphorylation to occur, oxygen (O_2) is used as a reactant and carbon dioxide (CO_2) is released as a waste product. You may be surprised to learn that although oxygen is a critical need for cells, it is actually the accumulation of carbon dioxide that primarily drives your need to breathe. Carbon dioxide is exhaled and oxygen is inhaled through the respiratory system, which involves muscles to move air into and out of the lungs, passageways through which air moves, and microscopic gas exchange surfaces covered by capillaries. The circulatory system transports gases from the lungs to tissues throughout the body and vice versa.

Part 1: Anatomy of the Respiratory System

The major organs of the respiratory system (Figure 1) function primarily to provide oxygen to body tissues for cellular respiration, remove the waste product carbon dioxide, and help to maintain acid-base balance. Functionally, the respiratory system can be divided into a conducting zone and a respiratory zone. The conducting zone of the respiratory system includes the organs and structures not directly involved in gas exchange. The gas exchange occurs in the respiratory zone.

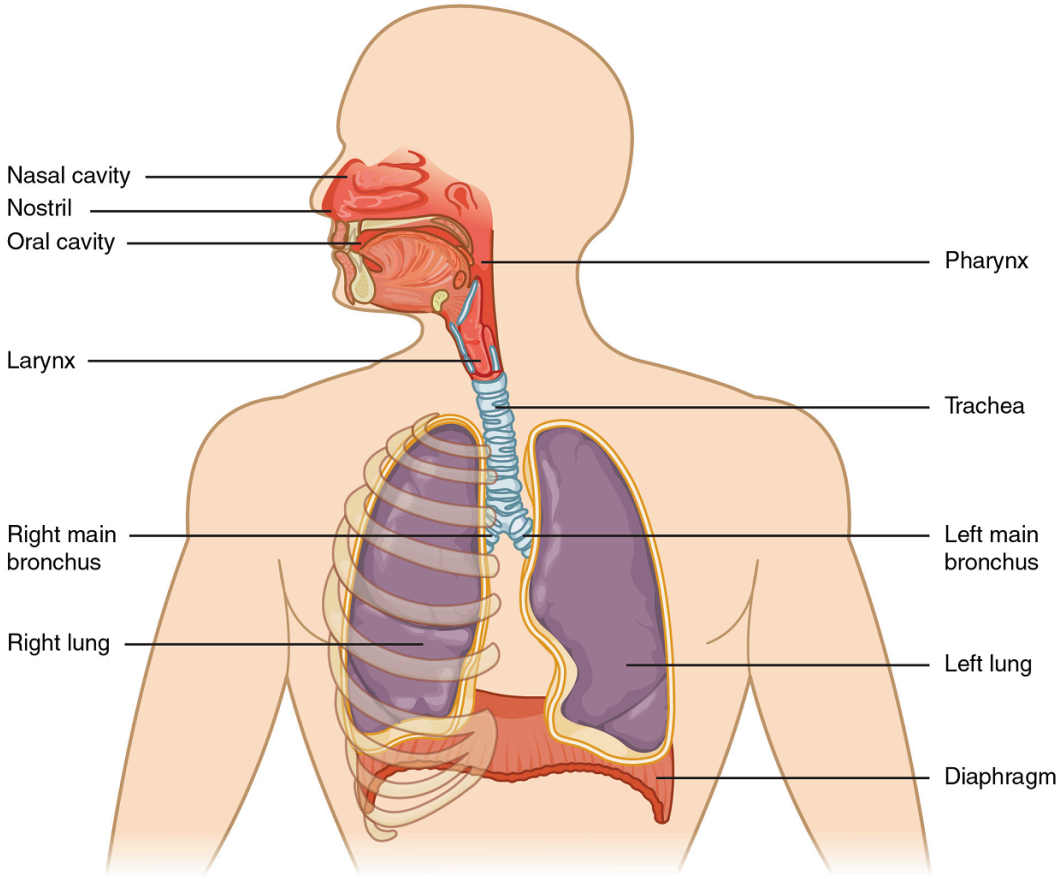


Figure 1. Major Respiratory Structures. *The major respiratory structures span the nasal cavity to the diaphragm.*

Conducting Zone: The major functions of the conducting zone are to provide a route for incoming and outgoing air, remove debris and pathogens from the incoming air, and warm and humidify the incoming air. In addition, the epithelium of the nasal passages, is essential to sensing odors

The Nose and its Adjacent Structures: The major entrance and exit for the respiratory system is through the nose; some air may also enter through the oral cavity. Large particles, such as dirt, are removed from air by hairs as it enters the **external nares** (nostrils) of the nose and flows into the **nasal cavity** (which is separated into left and right sections by the nasal septum). Each lateral wall of the nasal cavity has three conchae (Figure 2) or bony projections, which serve to increase the surface area of the nasal cavity and slow the flow of air in the nose to allow it to be cleaned, warmed and moistened. The nasal cavity is lined by **respiratory epithelium** (ciliated columnar epithelium) and goblet cells (Figure 3). The goblet cells produce mucus to trap debris and the cilia of the respiratory epithelium beat to clear the mucus from the nasal cavity to the throat to be swallowed. For further protection, goblet cells secrete antibacterial substances. Deep in the nasal cavity, an olfactory epithelium is used to detect odors (as described in the special sense unit of BIOL 1103/1109). Air exits the nasal cavities via the **internal nares** and moves into the **pharynx**.

Pharynx: The **pharynx** is a tube formed by skeletal muscle and lined by mucous membrane that is continuous with the epithelium of the nasal cavities. The pharynx is divided into three major regions: the nasopharynx, the oropharynx, and the laryngopharynx (Figure 3).

The **nasopharynx** is flanked by anteriorly by the conchae of the nasal cavity, and it serves only as an airway. The **oropharynx** and **laryngopharynx** are passageways for both air and food. At the inferior end of the laryngopharynx, the digestive and respiratory systems diverge. Anteriorly, the laryngopharynx opens into the larynx, whereas posteriorly, it enters the esophagus.

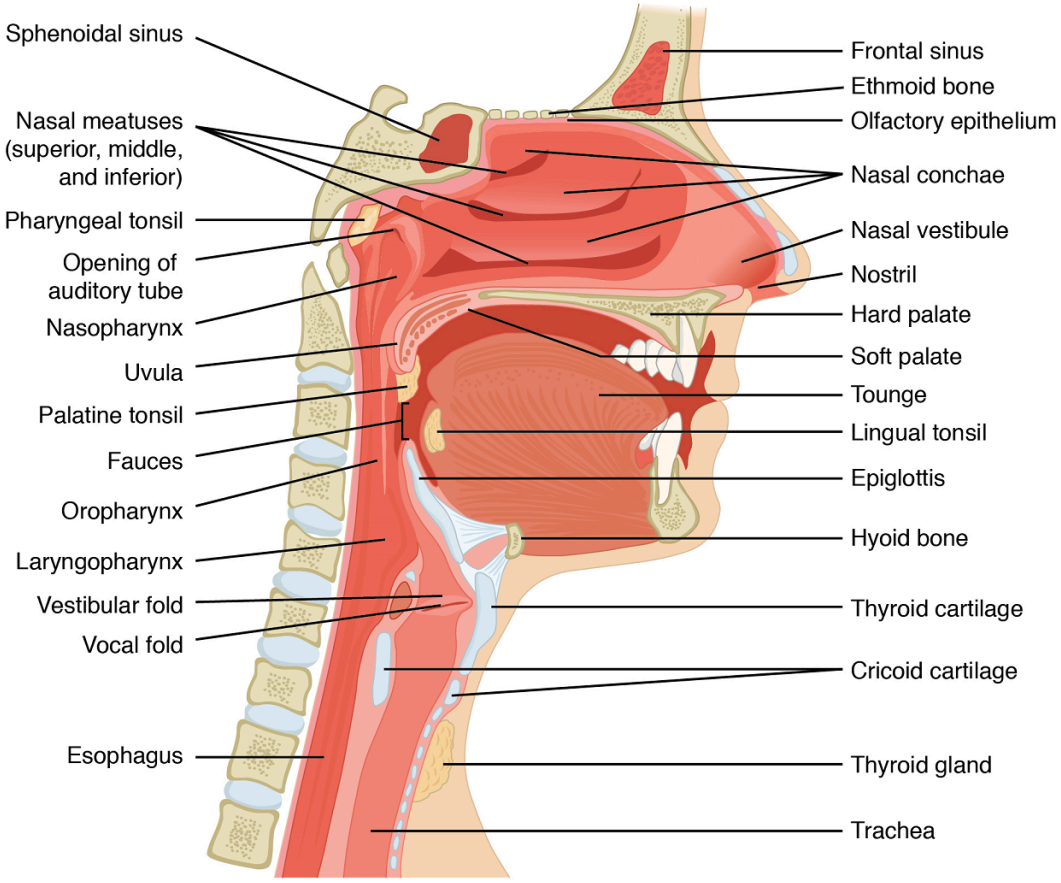


Figure 2. Upper Airway.

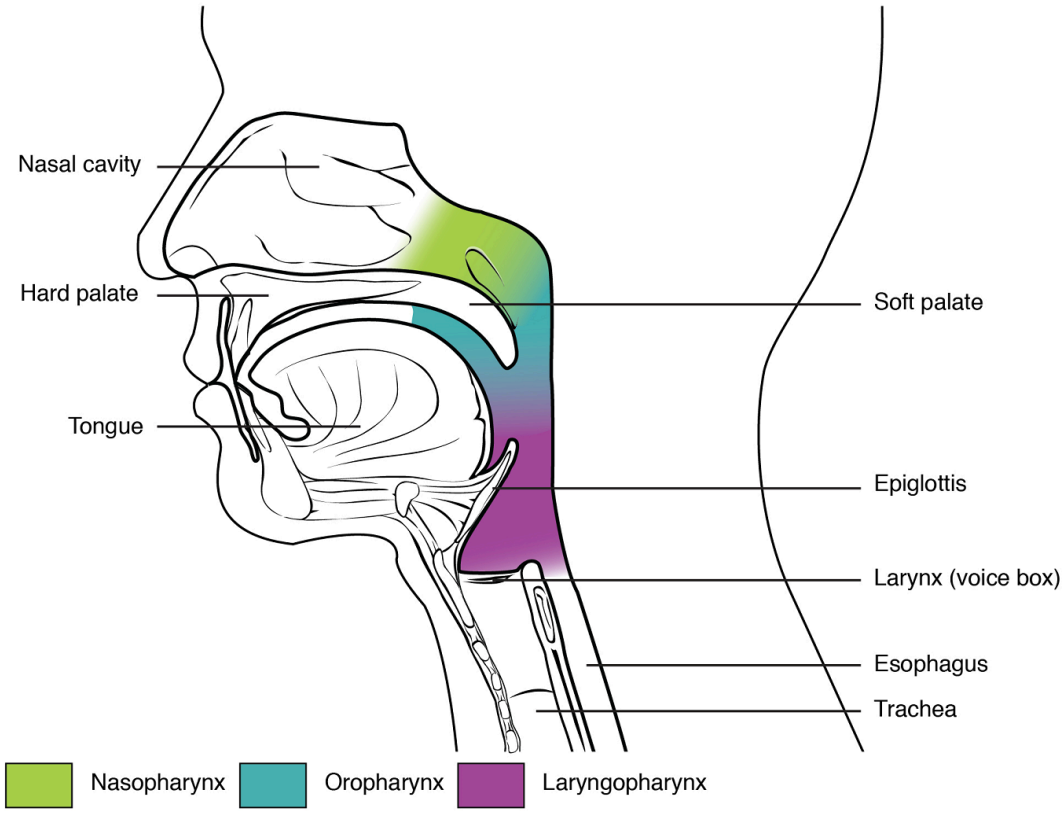


Figure 3. Pharynx. *The pharynx is divided into three regions: the nasopharynx, the oropharynx, and the laryngopharynx.*

Larynx: The **larynx** is a cartilaginous structure inferior to the laryngopharynx that connects the pharynx to the trachea and helps regulate the volume of air that enters and leaves the lungs. The structure of the larynx is formed by several pieces of cartilage (Figure 4). Three large cartilage pieces—the thyroid cartilage (anterior; which contains the Adam's apple or laryngeal prominence), epiglottis (superior), and cricoid cartilage (inferior)—form the major structure of the larynx. Other smaller, paired pieces of cartilage help move the vocal cords for speech.

The **epiglottis** is a very flexible piece of elastic cartilage that covers the opening of the trachea (see Figure 2). When in the "closed" position, during swallowing, the unattached end of the epiglottis rests on the glottis preventing food and beverages from entering the trachea (as will be described in the digestive system section). The glottis is composed of the vestibular folds, the true vocal cords, and the space between these folds (Figure 5).

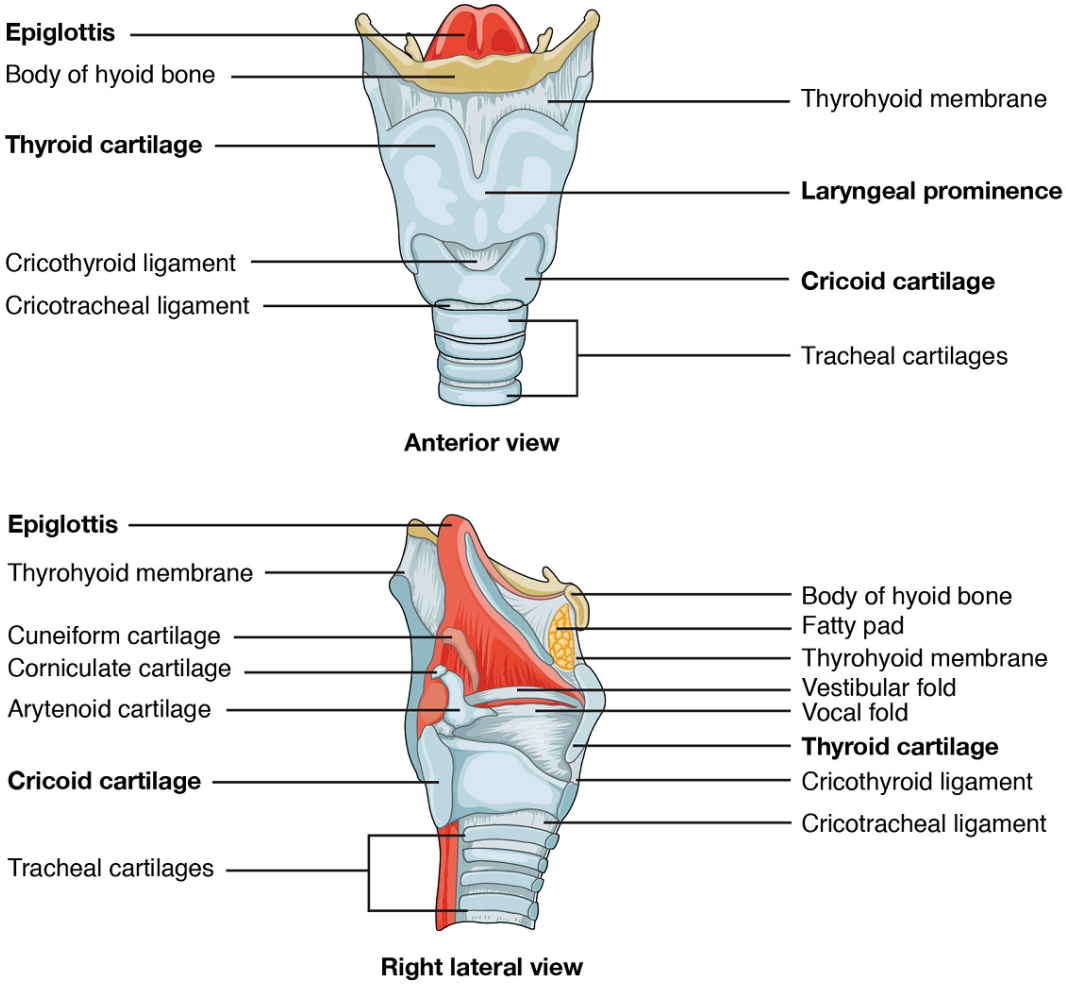


Figure 4. Larynx. *The larynx extends from the laryngopharynx to the trachea.*

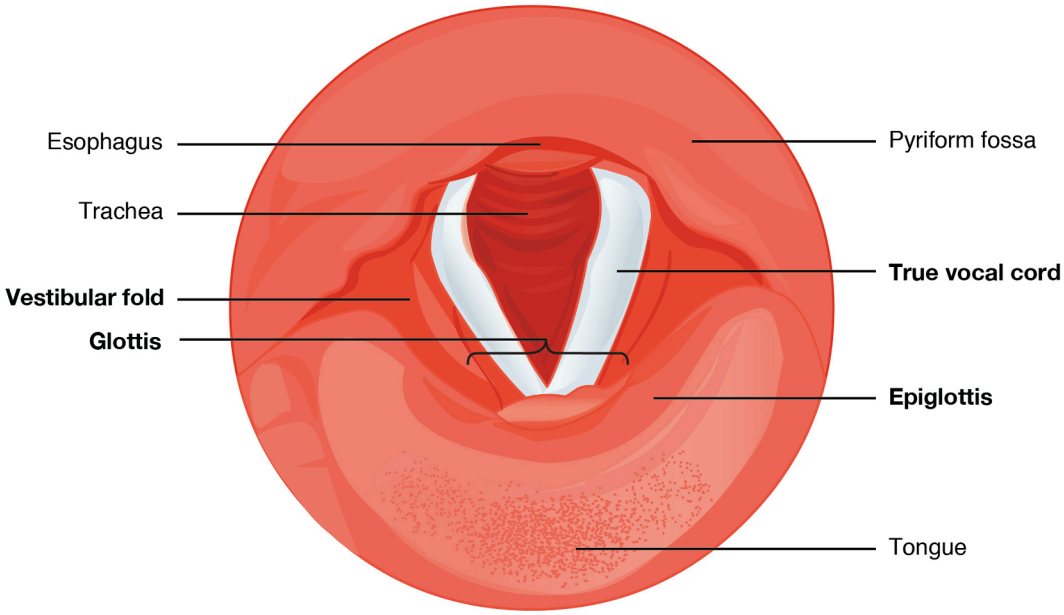


Figure 5. Vocal Cords. *The true vocal cords and vestibular folds of the larynx are viewed inferiorly from the laryngopharynx.*

Trachea: The **trachea** (windpipe) extends from the larynx toward the lungs (Figure 6a) and is formed by 16 to 20 stacked, C-shaped pieces of hyaline cartilage connected by dense connective tissue. The rings of cartilage provide structural support and prevent the trachea from collapsing.

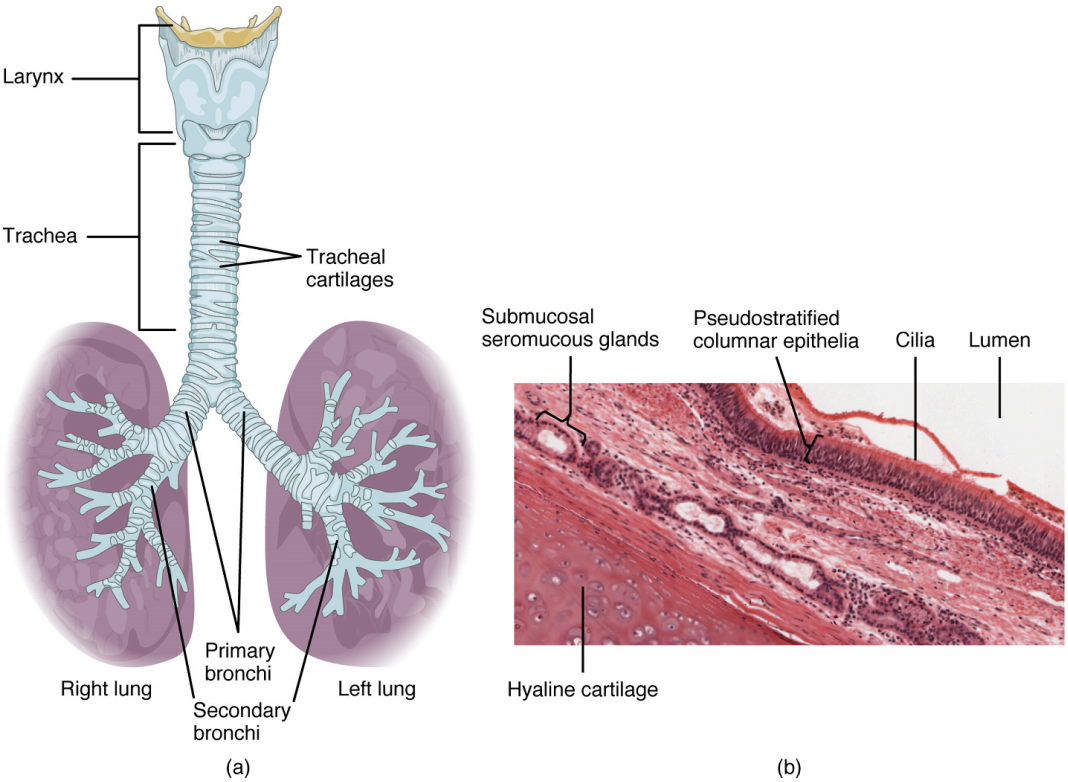


Figure 6. Trachea. *(a) The tracheal tube is formed by stacked, C-shaped pieces of hyaline cartilage. (b) The layer visible in this cross-section of tracheal wall tissue between the hyaline cartilage and the lumen of the trachea is the mucosa, which is composed of pseudostratified ciliated columnar epithelium that contains goblet cells. LM × 1220. (Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

Bronchial Tree: The trachea branches into the right and left primary **bronchi** (Figure 6b). Rings of cartilage, similar to those of the trachea, support the structure of the bronchi and prevent their collapse. The primary

bronchi enter the lungs and each primary bronchus branches into a secondary bronchus and then into a tertiary bronchus and so on. These multiple-branched bronchi are referred to as the bronchial tree. The main function of the bronchi, like other conducting zone structures, is to provide a passageway for air to move into and out of each lung. In addition, the mucous membrane traps debris and pathogens.

Bronchioles branch from the tertiary bronchi. Bronchioles, which are about 1 mm in diameter, further branch until they become the tiny **terminal bronchioles**, which lead to the structures of gas exchange (Figure 7). There are more than 1000 terminal bronchioles in each lung. The muscular walls of the bronchioles do not contain cartilage like those of the bronchi. This muscular wall can change the size of the tubing to increase or decrease airflow through the tube, for example during exercise.

Respiratory Zone: In contrast to the conducting zone, the respiratory zone includes structures that are directly involved in gas exchange of carbon dioxide and oxygen. The respiratory zone begins where the terminal bronchioles join a **respiratory bronchiole**, the smallest type of bronchiole (Figure 7), which then leads to an **alveolar duct**, opening into a cluster of alveoli (called an alveolar sac).

Alveoli: An **alveolus** is one of the many small, grape-like sacs in an alveolar sac responsible for gas exchange. An alveolus is approximately 200 µm in diameter with elastic walls that allow the alveolus to stretch during air intake, which greatly increases the surface area available for gas exchange. Alveoli are connected to their neighbors by alveolar pores, which help maintain equal air pressure throughout the alveoli and lung (Figure 8).

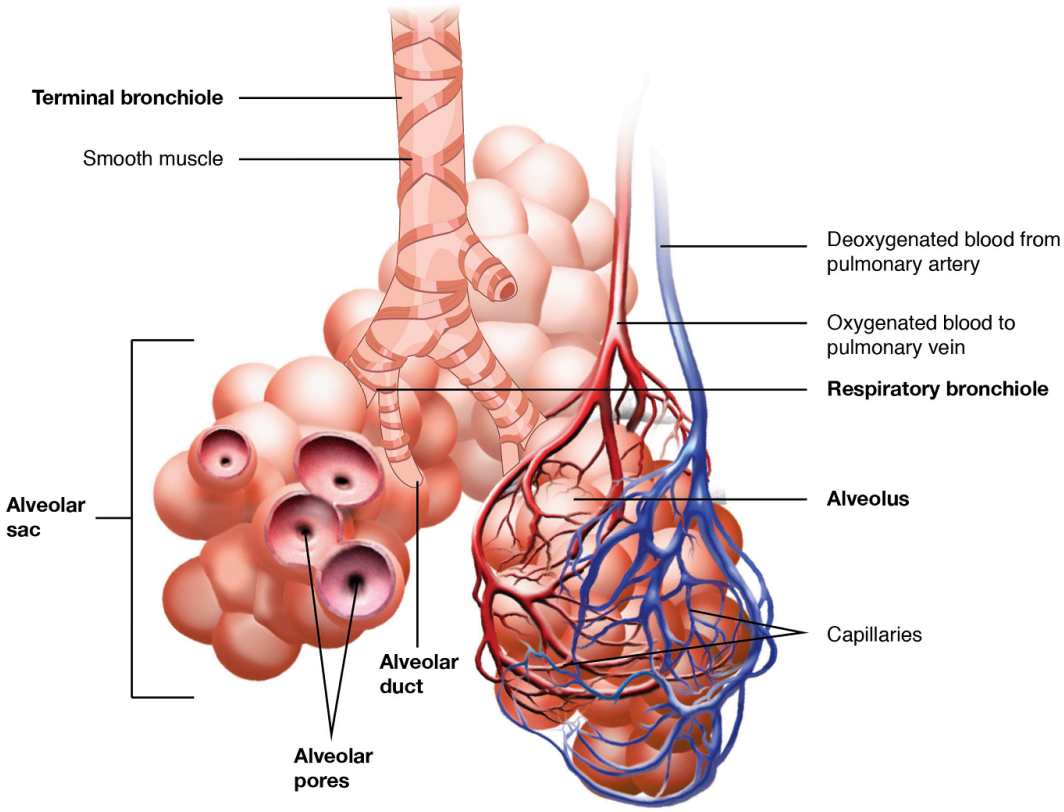


Figure 7. Respiratory Zone. *Bronchioles lead to alveolar sacs in the respiratory zone, where gas exchange occurs.*

The alveolar wall consists mostly of simple squamous epithelial cells. These cells are about 25 nm thick and are highly permeable to gases.

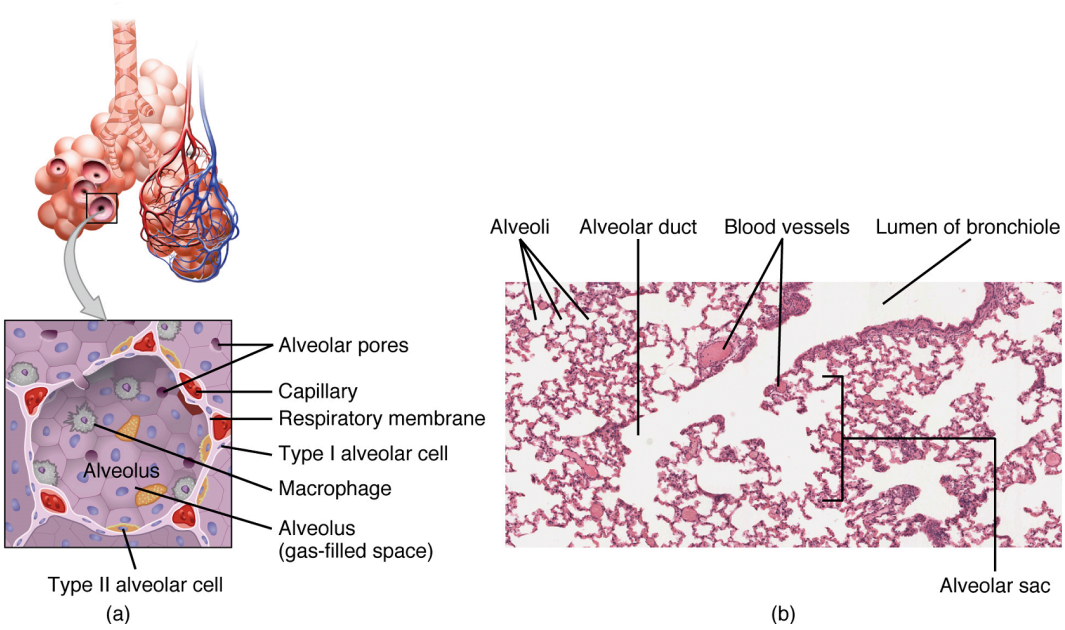


Figure 8. Structures of the Respiratory Zone. *(a) The alveolus is responsible for gas exchange. (b) A micrograph shows the alveolar structures within lung tissue. LM × 178. (Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

The simple squamous epithelium formed by type I alveolar cells is attached to a thin, elastic basement membrane. This epithelium is extremely thin and borders the endothelial membrane of capillaries, which is also composed of simple squamous epithelium. Taken together, the alveoli and capillary membranes form a **respiratory membrane** that is approximately 0.5 mm thick. The respiratory membrane allows gases to cross by simple diffusion, allowing oxygen to be picked up by the blood for transport and carbon dioxide to be released into the air of the alveoli. The respiratory membrane covers a surface area—about 70 square meters.

Watch this Crash Course video for an overview of the respiratory system! Direct link: https://youtu.be/bHZsvBdUC2I

The Lungs: The major organs of the respiratory system; each lung houses structures of both the conducting and respiratory zones starting at the primary bronchi and ending at the alveoli.

Gross Anatomy of the Lungs: The lungs are pyramid-shaped, paired organs that are connected to the trachea by the right and left bronchi; on the inferior surface, the lungs are bordered by the diaphragm. The diaphragm is the flat, dome-shaped muscle located at the base of the lungs and thoracic cavity. The lungs are enclosed by the pleurae. The right lung is shorter and wider than the left lung, and the left lung occupies a smaller volume than the right to allow space for the heart (Figure 9). Each lung is composed of smaller units called lobes. The right lung consists of three lobes: the superior, middle, and inferior lobes. The left lung consists of two lobes: the superior and inferior lobes.

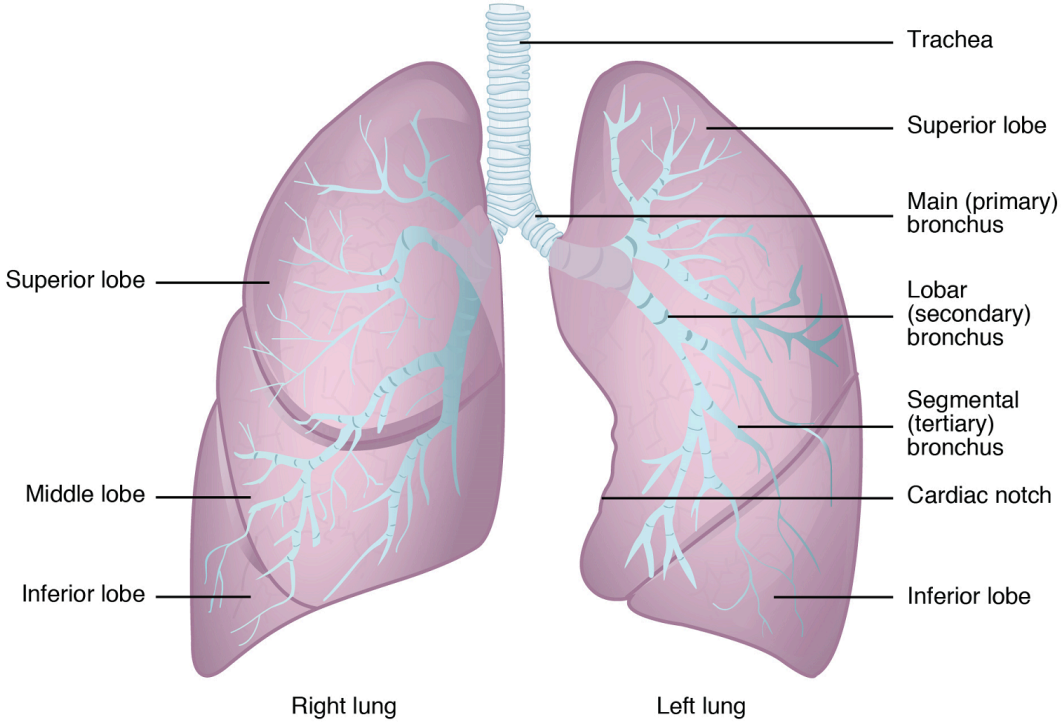


Figure 9. Gross Anatomy of the Lungs.

Pleura of the Lungs: Each lung is enclosed within a cavity that is surrounded by a serous membrane called pleura (plural: pleurae). The right and left pleurae enclose the right and left lungs, respectively. The pleurae consist of two layers: the **visceral pleura** is the layer that is superficial to the lungs (Figure 10). In contrast, the **parietal pleura** is the outer layer that connects to the thoracic wall and the diaphragm. The visceral and parietal pleurae connect to each other at the hilum (near the branching of the trachea to the primary bronchi). The **pleural cavity** is the space between the visceral and parietal layers.

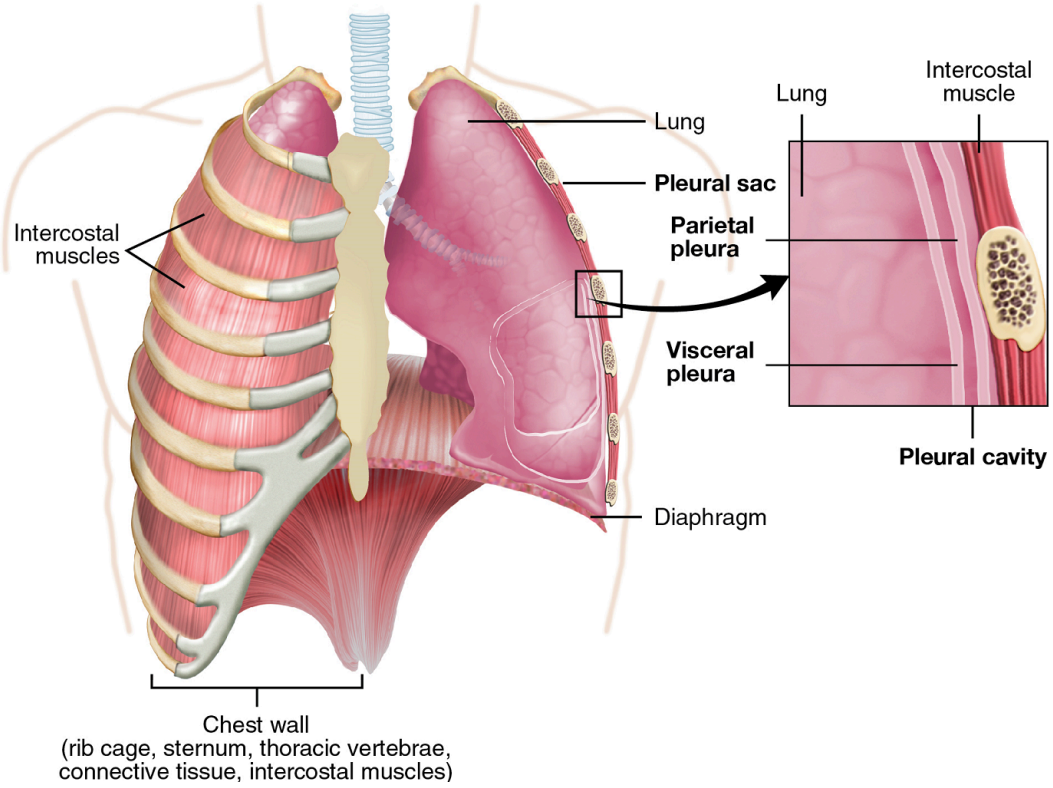


Figure 10. Parietal and Visceral Pleurae of the Lungs.

The pleurae perform two major functions: to produce pleural fluid and to create cavities that separate the major organs. **Pleural fluid** is secreted by cells from both pleural layers into the pleural cavity and acts to lubricate their surfaces. This lubrication reduces friction between the two layers to prevent damage during breathing, and creates surface tension that helps maintain the position of the lungs against the thoracic wall. This adhesive characteristic of the pleural fluid causes the lungs to enlarge when the thoracic wall expands during ventilation, allowing the lungs to fill with air. The division that the pleurae create between major organs prevents interference due to the movement of lungs during breathing and also helps to prevent the spread of infection.

Part 2: The Process of Breathing

Pulmonary ventilation is the act of breathing: the movement of air into the lungs (inspiration or inhalation) and movement of air out of the lungs (expiration or exhalation. The major mechanisms that drive pulmonary ventilation are atmospheric pressure (P_{atm}); the air pressure within the alveoli, called intra-alveolar pressure (P_{alv}); and the pressure within the pleural cavity, called intrapleural pressure (P_{ip}).

Atmospheric pressure is the amount of force that is exerted by gases in the air surrounding any given surface, such as the body. Atmospheric pressure at sea level is 760 mmHg ("millimetres of mercury"). Typically, for respiration, other pressure values are discussed in relation to atmospheric pressure. Therefore, negative pressure is pressure lower than the atmospheric pressure, whereas positive pressure is pressure that is greater than the atmospheric pressure. A pressure that is equal to the atmospheric pressure is expressed as zero.

Intra-alveolar pressure is the pressure of the air within the alveoli, which changes during the different phases of breathing (Figure 12). Because the alveoli are connected to the atmosphere via the tubing of the airways, the intra-alveolar pressure always equalizes with the atmospheric pressure.

Intrapleural pressure is the pressure of the air within the pleural cavity, between the visceral and parietal pleurae. Similar to intra-alveolar pressure, intrapleural pressure also changes during the different phases of

breathing. The intrapleural pressure always remains lower than, or negative to, the intra-alveolar pressure (by approximately 4 mmHg).

Understanding Boyle's Law of gases helps to explain the movement of air in and out of the lungs during pulmonary ventilation. **Boyle's law** describes the relationship between volume and pressure in a gas in a confined space, at a constant temperature. Boyle discovered that the pressure of a gas is inversely proportional to its volume: If volume increases, pressure decreases. Likewise, if volume decreases, pressure increases; pressure and volume are inversely related (see Figure 11).

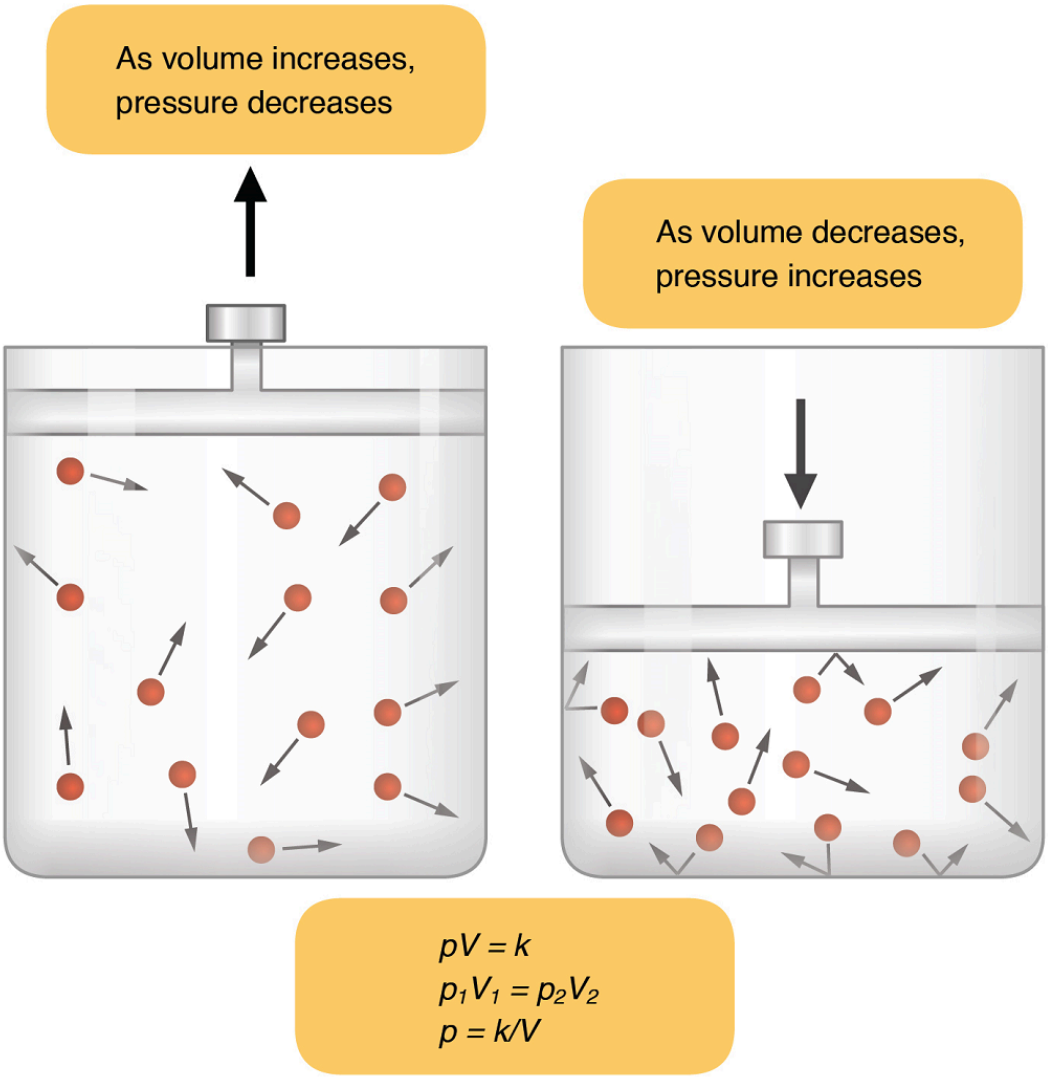


Figure 11. Boyle's Law. *In a gas, pressure increases as volume decreases.*

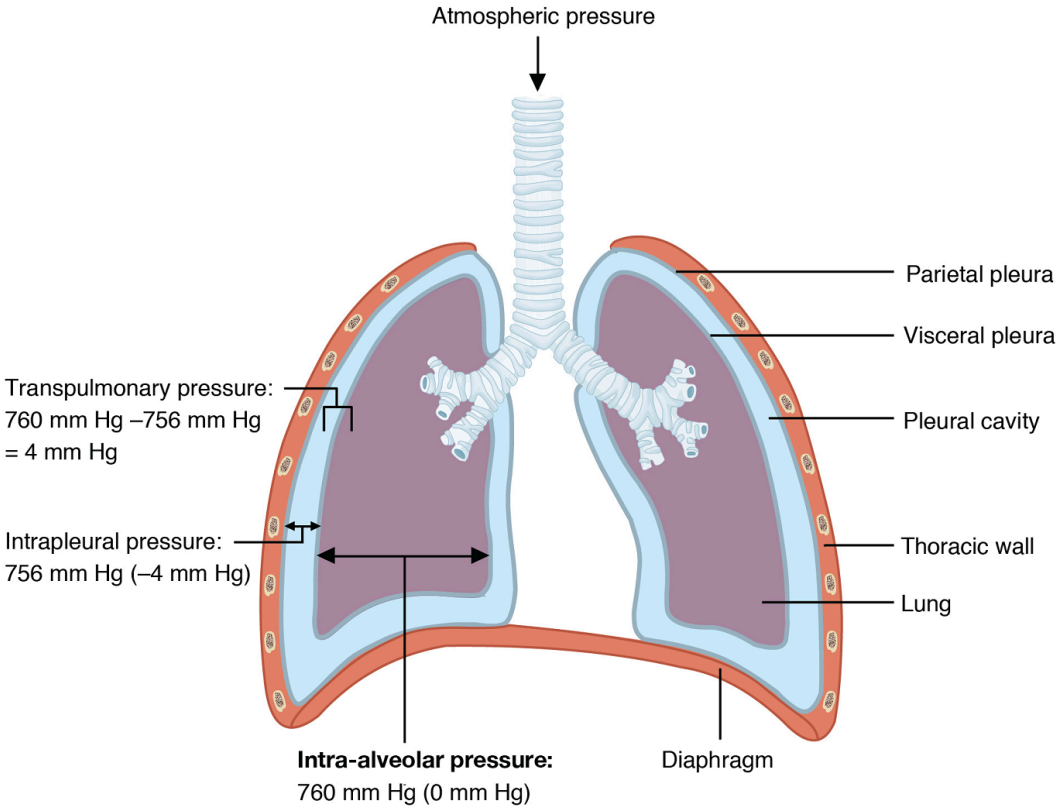


Figure 12. Intrapulmonary and Intrapleural Pressure Relationships. *Intrapleural pressure is normally always slightly lower than atmospheric pressure, but intra-alveolar pressure changes during the different phases of the respiratory cycle. When intra-alveolar pressure is made higher or lower than atmospheric pressure, air moves out of or into the lungs until intra-alveolar pressure and atmospheric pressure are the same (as shown). (Pressures shown in brackets are relative to a normal atmospheric pressure of 760 mmHg.)*

Pulmonary Ventilation: Pulmonary ventilation is mainly dependent on the contraction and relaxation of the diaphragm and the external intercostal muscles which change the intra-alveolar pressure and air flows down a pressure gradient.

A **respiratory cycle** is one sequence of inspiration and expiration (Figure 13). In general, two muscle groups are used during normal quiet inspiration: the diaphragm and the external intercostal muscles. Additional muscles can be used if a bigger breath is required. When the diaphragm contracts, it moves inferiorly toward the abdominal cavity, creating a larger thoracic cavity and more space for the lungs.

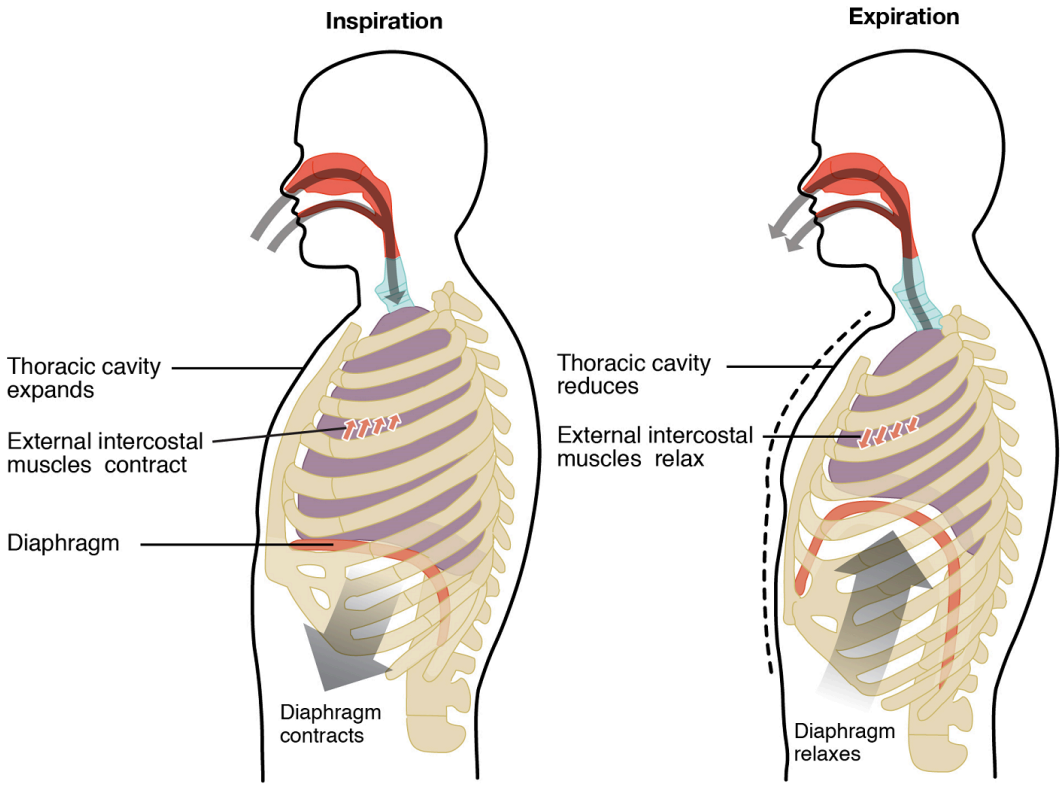


Figure 13. Inspiration and Expiration. *Inspiration and expiration occur due to the expansion and contraction of the thoracic cavity, respectively.*

Contraction of the external intercostal muscles moves the ribs upward and outward, causing the rib cage to expand, which increases the volume of the thoracic cavity. Due to the adhesive force of the pleural fluid, the expansion of the thoracic cavity forces the lungs to passively stretch and expand as well. This increase in volume leads to a decrease in intra-alveolar pressure (as stated in Boyle's Law of Gases), creating a pressure lower than atmospheric pressure. As a result, a pressure gradient is created that drives air into the lungs.

The process of normal expiration is passive, meaning that energy is not required to push air out of the lungs. The diaphragm and external intercostal muscles relax following inspiration and the stretched elastic tissue of the lungs passively recoils. In turn, the thoracic cavity and lungs decrease in volume, causing an increase in intra-alveolar pressure. The intra-alveolar pressure rises above atmospheric pressure, creating a pressure gradient that causes air to leave the lungs.

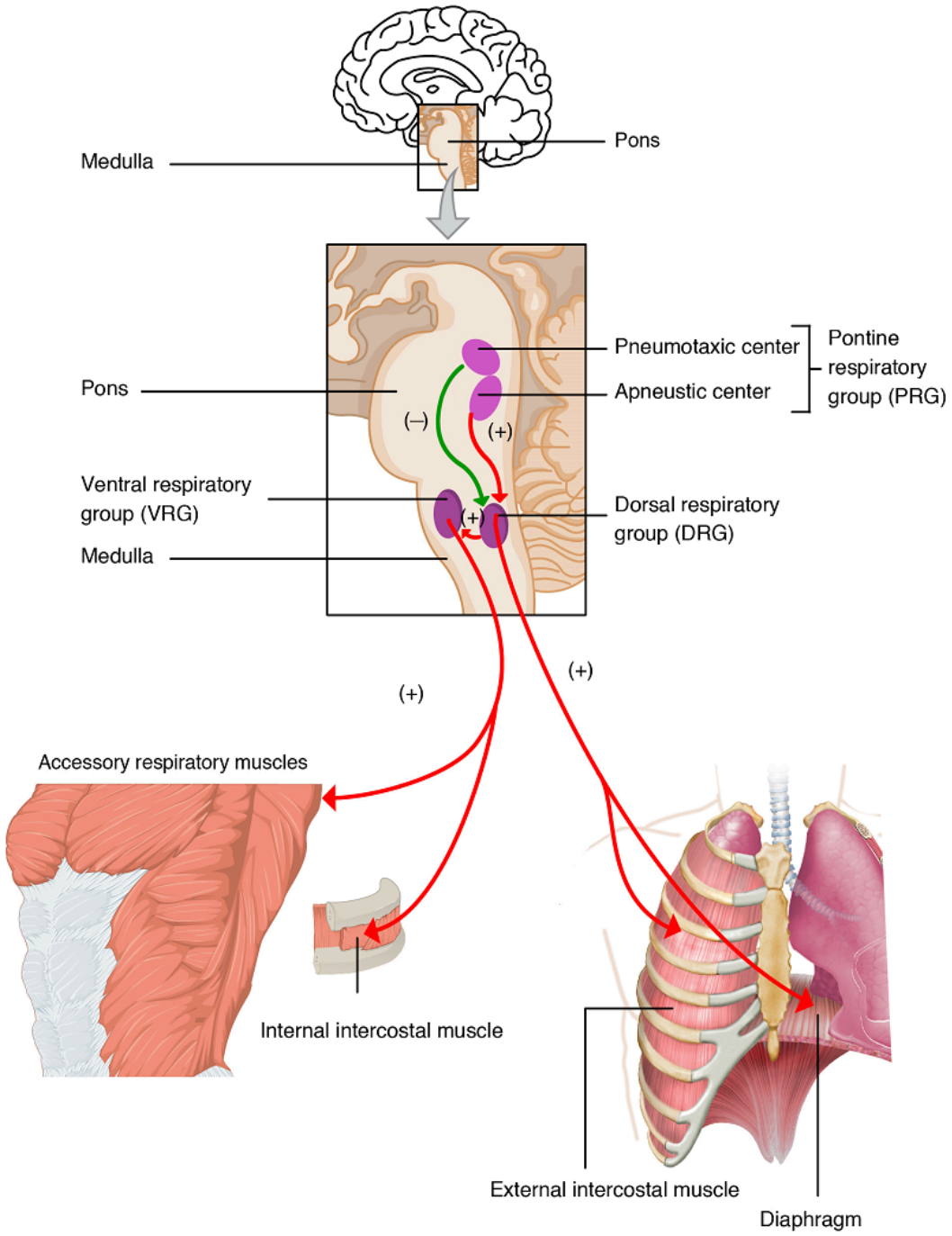
Watch this Crash Course video to learn more about the breathing process! Direct link: https://youtu.be/bHZsvBdUC2I

Nervous Control of Ventilation: Breathing usually occurs without thought, although at times you can

consciously control it, such as when you swim under water, sing a song, or blow bubbles. The control of ventilation is a complex interplay of multiple regions in the brain, primarily in the brainstem, that signal the muscles used in pulmonary ventilation to contract. The result is typically a rhythmic, consistent ventilation rate that provides the body with sufficient amounts of oxygen, while adequately removing carbon dioxide.

The **respiratory rate** is the total number of breaths, or respiratory cycles, that occur each minute. Respiratory rate can be an important indicator of disease, as the rate may increase or decrease during an illness or in a disease condition. The respiratory rate is controlled by the respiratory center located within the medulla oblongata in the brain, which responds primarily to changes in carbon dioxide, oxygen, and pH levels in the blood. The normal respiratory rate of a child decreases from birth to adolescence. A child under 1 year of age has a normal respiratory rate between 30 and 60 breaths per minute, but by the time a child is about 10 years old, the normal rate is closer to 18 to 30. By adolescence, the normal respiratory rate is similar to that of adults, 12 to 18 breaths per minute.

Figure 14. Respiratory Centers of the Brain



The medulla oblongata includes two main populations of neurons controlling breathing: the **dorsal respiratory group (DRG)** and the **ventral respiratory group (VRG)** (Figure 14). These medullary nuclei maintain a constant breathing rhythm by alternatively stimulating the diaphragm and external intercostal muscles to contract, resulting in inspiration, and ceasing this stimulation, resulting in relaxation of these muscle and thus expiration.

Activity of the dorsal and ventral respiratory groups is modulated by input from peripheral stretch and chemoreceptors, and by input from the pontine respiratory group. Neurons in the ventral respiratory group also receive input from the dorsal respiratory group, and can also stimulate the accessory muscles involved in forced breathing to contract, resulting in forced inspiration or forced expiration.

The second respiratory center of the brain is located within the pons, called the pontine respiratory group, and consists of the apneustic and pneumotaxic centers.

The **apneustic center** is a double cluster of neuronal cell bodies that stimulate neurons in the DRG, controlling the depth of inspiration, particularly for deep breathing. The **pneumotaxic center** is a network of neurons that inhibits the activity of neurons in the DRG, allowing relaxation after inspiration, and thus controlling the overall rate.

Factors That Affect the Rate and Depth of Respiration: The respiratory rate and the depth of inspiration are regulated by the medulla oblongata and pons; however, these regions of the brain do so in response to systemic stimuli, mainly carbon dioxide (CO_2), oxygen (O_2) and hydrogen ions (H^+) in the blood. Multiple systemic factors are involved in stimulating the brain to produce pulmonary ventilation.

The major factor that stimulates the medulla oblongata and pons to increase respiration is surprisingly not oxygen concentration, but rather the concentration (or partial pressure) of carbon dioxide (CO_2) in the blood.

Concentration changes in the blood in certain substances, such as carbon dioxide, oxygen, or hydrogen ions (H^+) stimulate chemoreceptors, which in turn signal the respiratory centers of the brain. There are **central chemoreceptors** located in the brain or brainstem, which are sensitive to are sensitive to hydrogen ions and **peripheral chemoreceptors** located in two blood vessels: the carotid arteries and aortic arch, which are sensitive to changes in arterial carbon dioxide, oxygen, and hydrogen ion concentrations.

- **Central chemoreceptors:** As the concentration of carbon dioxide in the blood increases, it readily diffuses across the blood-brain barrier, where it collects in the extracellular fluid. As will be explained in more detail later, carbon dioxide is converted into carbonic acid, which leads to an increased concentration of hydrogen ions (increased acidity or decreased pH). The increase in hydrogen ions in the brain triggers the **central chemoreceptors** to stimulate the respiratory centers to initiate contraction of the diaphragm and intercostal muscles. As a result, the rate and depth of respiration increase, allowing more carbon dioxide to be expelled, which promotes a reduction in the concentration of carbon dioxide in the blood, and thus hydrogen ions. In contrast, a low concentration of carbon dioxide in the blood causes a low concentration of hydrogen ions in the brain, leading to a decrease in the rate and depth of pulmonary ventilation, producing shallow, slow breathing. Peripheral chemoreceptors are sensitive to arterial carbon dioxide but are less active in stimulating respiratory rate then the central H^+ chemoreceptors.
- **Peripheral chemoreceptors:** Increased hydrogen ion concentration in the blood can be due to increasing carbon dioxide, as mentioned above, but can also be due to other metabolic activities, such as lactic acid accumulation after strenuous exercise or ketoacidosis. **Peripheral chemoreceptors** of the aortic arch and carotid arteries sense arterial levels of hydrogen ions. When peripheral chemoreceptors sense decreasing, or more acidic, pH levels, they stimulate the respiratory centre and cause an increase in ventilation to remove carbon dioxide from the blood at a quicker rate. Removal of carbon dioxide from the blood helps to reduce hydrogen ions, thus increasing systemic pH.
- **Blood oxygen** (O_2) concentration (or partial pressure of oxygen) is also important in influencing respiratory rate. In addition to their role outlined above, the **peripheral chemoreceptors** are responsible for sensing large changes in blood oxygen concentration. If blood oxygen concentration falls very low (about 60 mmHg or less) then peripheral chemoreceptors signal the respiratory centre which causes an increase in respiratory activity. The chemoreceptors are only able to sense dissolved oxygen molecules, not the oxygen that is bound to hemoglobin. As will be described in the next section, the majority of oxygen is bound by hemoglobin; when dissolved levels of oxygen drop, hemoglobin releases oxygen. Therefore, a large drop in blood oxygen concentration is required to stimulate the chemoreceptors of the aortic arch and carotid arteries.
- **Higher brain centers** influence the regulation of breathing by interacting with the respiratory centers. Most of the time, the brainstem regulates breathing involuntarily but we do have some voluntary, conscious control over breathing rate, for example during singing or holding our breath. This is due to signals from the

cerebral cortex to the respiratory centers. The hypothalamus and other regions associated with the limbic system are involved in regulating respiration in response to **emotions, pain**, and **temperature**. For example, an increase in body temperature or feeling excited (the fight-or-flight response) will result in an increase in respiratory rate.

Respiratory Volumes and Capacities: Respiratory volume is the term used for various volumes of air moved by or associated with the lungs at a given point in the respiratory cycle. There are four major types of respiratory volumes: tidal, residual, inspiratory reserve, and expiratory reserve (Figure 15). **Tidal volume (TV)** is the amount of air that normally enters the lungs during quiet breathing, which is about 500 milliliters (mL). **Expiratory reserve volume (ERV)** is the amount of air you can forcefully exhale past a normal tidal expiration, up to 1200 mL for men. **Inspiratory reserve volume (IRV)** is produced by a deep inhalation, past a tidal inspiration. This is the extra volume that can be brought into the lungs during a forced inspiration. **Residual volume (RV)** is the amount of air left in the lungs if you exhale as much air as possible. This air prevents alveoli collapsing in on themselves, which makes it possible for alveoli to re-inflate during inhalation. Without residual volume, pulmonary ventilation would cease. Respiratory volume is dependent on a variety of factors, and measuring the different types of respiratory volumes by using a spirometer to generate a spirogram can provide important clues about a person's respiratory health (Figure 15).

Respiratory capacity is the combination of two or more selected volumes, which further describes the amount of air in the lungs during a given time. For example, **total lung capacity (TLC)** is the sum of all of the lung volumes (TV + ERV + IRV + RV), which represents the total amount of air a person can hold in the lungs after a forceful inhalation. TLC is about 6000 mL air for men, and about 4200 mL for women. **Vital capacity (VC)**, which is between 4000 and 5000mL, is the amount of air a person can move into or out of his or her lungs, and is the sum of all of the volumes except residual volume (TV + ERV + IRV).

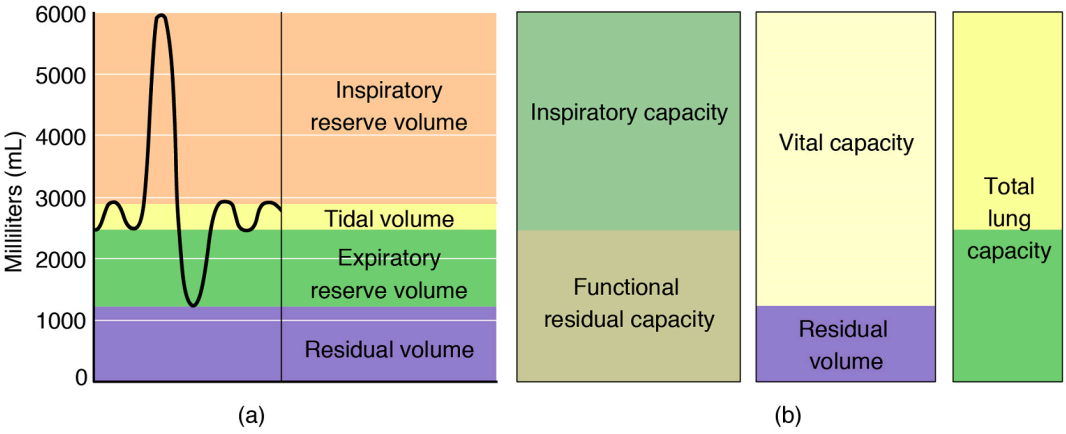


Figure 15. Respiratory Volumes and Capacities. *These two graphs show (a) respiratory volumes and (b) the combination of volumes that results in respiratory capacity.*

Inspiratory capacity (IC) is the maximum amount of air that can be inhaled past a normal tidal expiration, and is the sum of the tidal volume and inspiratory reserve volume. On the other hand, the **functional residual capacity (FRC)** is the amount of air that remains in the lung after a normal tidal expiration; it is the sum of expiratory reserve volume and residual volume (see Figure 15).

In addition to the air that creates respiratory volumes, the respiratory system also contains **anatomical dead space**, which is air that is present in the airway that never reaches the alveoli and therefore never participates in gas exchange. **Alveolar dead space** involves air found within alveoli that are unable to function, such as those affected by disease or abnormal blood flow. **Total dead space** is the anatomical dead space and alveolar dead space together, and represents all of the air in the respiratory system that is not being used in the gas exchange process.

Part 3: Gas Exchange

To understand the mechanisms of gas exchange in the lung, it is important to understand the underlying principles of gases and their behavior.

Gas Laws and Air Composition: Gas molecules exert force on the surfaces with which they are in contact; this force is called pressure. In natural systems, gases are normally present as a mixture of different types of molecules. For example, the atmosphere consists of oxygen, nitrogen, carbon dioxide, and other gaseous molecules, and this gaseous mixture exerts a certain pressure referred to as atmospheric pressure (Table 1). **Partial pressure** (P_x) is the pressure of a single type of gas in a mixture of gases (Figure 16). **Total pressure** is the sum of all the partial pressures of a gaseous mixture; this is **Dalton's law.**

Table 1: Partial Pressures of Atmospheric Gases at Sea Level

Gas	Percent of total composition	Partial pressure (mmHg)
Nitrogen (N_2)	78.6%	597.4
Oxygen (O_2)	20.9%	158.8
Water (H_2O)	0.04%	3.0
Carbon dioxide (CO_2)	0.004%	0.3
Others	0.0006%	0.5
Total	100%	760.0

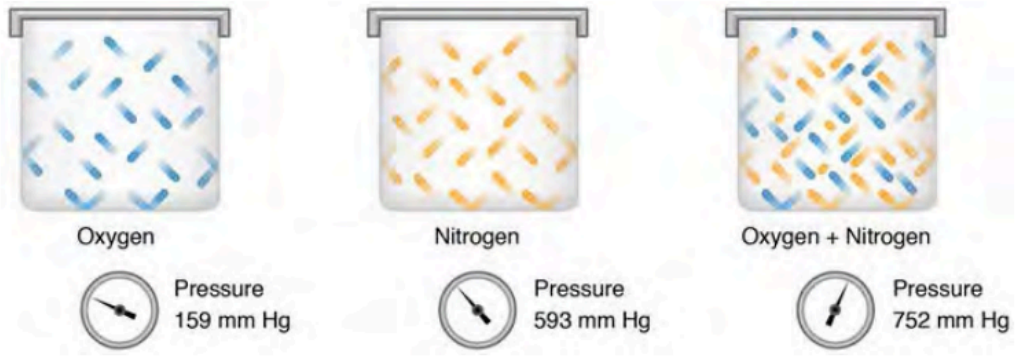


Figure 16. Partial and Total Pressures of a Gas. *Partial pressure is the force exerted by a gas. The sum of the partial pressures of all the gases in a mixture equals the total pressure.*

Partial pressure of a gas can be thought of as "concentration" of that gas and, like concentration, partial pressure is extremely important in predicting the movement of gases. A gas will move from an area where its partial pressure is higher to an area where its partial pressure is lower. In addition, the greater the partial pressure difference between the two areas, the more rapid is the movement of gases.

The gases from the atmosphere must be dissolved in water for gas exchange to take place in the alveoli. Atmospheric air dissolves in liquid as it passes through the respiratory system. The partial pressure of a gas and its concentration in a liquid, such as blood, are not identical because solubility of the specific gas plays a factor in concentration. For example, although nitrogen is present in the atmosphere, very little nitrogen dissolves into the blood, because the solubility of nitrogen in alveolar fluid and blood is very low. Oxygen and carbon dioxide are more soluble.

Gas Exchange

Gas exchange occurs at two sites in the body:

- In the lungs, at the respiratory membrane oxygen acquired from inspiration is picked up and carbon dioxide is released and removed via expiration. This exchange of gases with the external environment (atmosphere) in the lungs is known as **external respiration**.
- At the tissues, oxygen is released into cells and carbon dioxide is picked up from cells. This gas exchange within the internal environment of the tissues of body is known as **internal respiration**.

The actual exchange of these non-polar, hydrophobic gases occurs due to **simple diffusion**. Energy is not required to move oxygen or carbon dioxide across membranes. Instead, these gases follow partial pressure gradients that allow them to diffuse.

External Respiration: The pulmonary artery carries deoxygenated blood into the lungs from the heart, where it branches and eventually becomes the capillary network composed of pulmonary capillaries that wrap around the alveoli and form the respiratory membrane (Figure 17). As the blood is pumped through this capillary network, gas exchange occurs.

The anatomy of the lung maximizes the diffusion of gases: The respiratory membrane is highly permeable to gases; the respiratory and blood capillary membranes are very thin; and there is a large surface area throughout the lungs.

Although a small amount of the oxygen is able to dissolve directly into plasma from the alveoli, most of the oxygen is picked up by erythrocytes (red blood cells) and binds to a protein called hemoglobin, a process described later in this chapter. Oxygenated hemoglobin is red, causing the overall appearance of bright red oxygenated blood, which returns to the heart through the pulmonary veins. Carbon dioxide is released in the opposite direction of oxygen, from the blood to the alveoli.

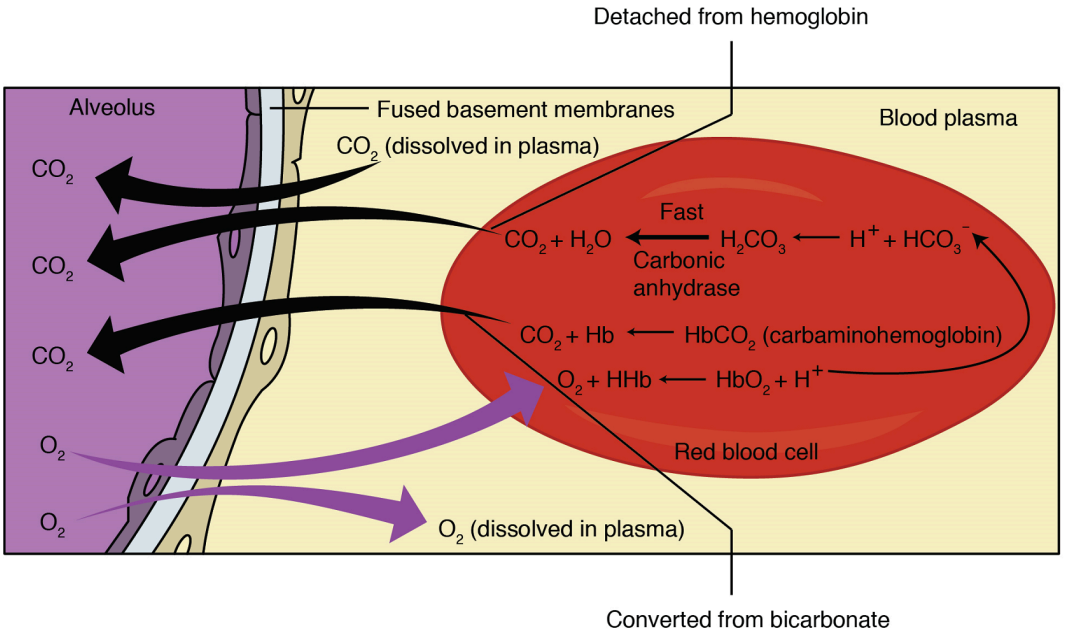


Figure 17. External Respiration. *In external respiration, oxygen diffuses across the respiratory membrane from the alveolus to the capillary, whereas carbon dioxide diffuses out of the capillary into the alveolus.*

External respiration occurs as a function of partial pressure differences in oxygen and carbon dioxide between the alveoli and the blood in the pulmonary capillaries. Although the solubility of oxygen in blood is not high, there is a drastic difference in the partial pressure of oxygen in the alveoli versus in the blood of the pulmonary

capillaries. This large difference in partial pressure creates a very strong pressure gradient that causes oxygen to rapidly cross the respiratory membrane from the alveoli into the blood.

The partial pressure of carbon dioxide is also different between the alveolar air and the blood of the capillary allowing carbon dioxide to diffuse from the blood into the alveoli.

Internal Respiration: Internal respiration is gas exchange that occurs at the level of body tissues (Figure 18). Similar to external respiration, internal respiration also occurs as simple diffusion due to a partial pressure gradient. However, the partial pressure gradients are opposite of those present at the respiratory membrane. The partial pressure of oxygen in tissues is low, because oxygen is continuously used for cellular respiration whereas the partial pressure of oxygen in the blood higher. This creates a pressure gradient that causes oxygen to dissociate from hemoglobin, diffuse out of the blood, cross the interstitial space, and enter the tissue cells. Hemoglobin that has little oxygen bound to it loses much of its brightness, so that blood returning to the heart is more burgundy in color.

Considering that cellular respiration continuously produces carbon dioxide, the partial pressure of carbon dioxide is lower in the blood than it is in the tissue, causing carbon dioxide to diffuse out of the tissue cells, cross the interstitial fluid, and enter the blood. It is then carried back to the lungs either bound to hemoglobin, dissolved in plasma, or in a converted form.

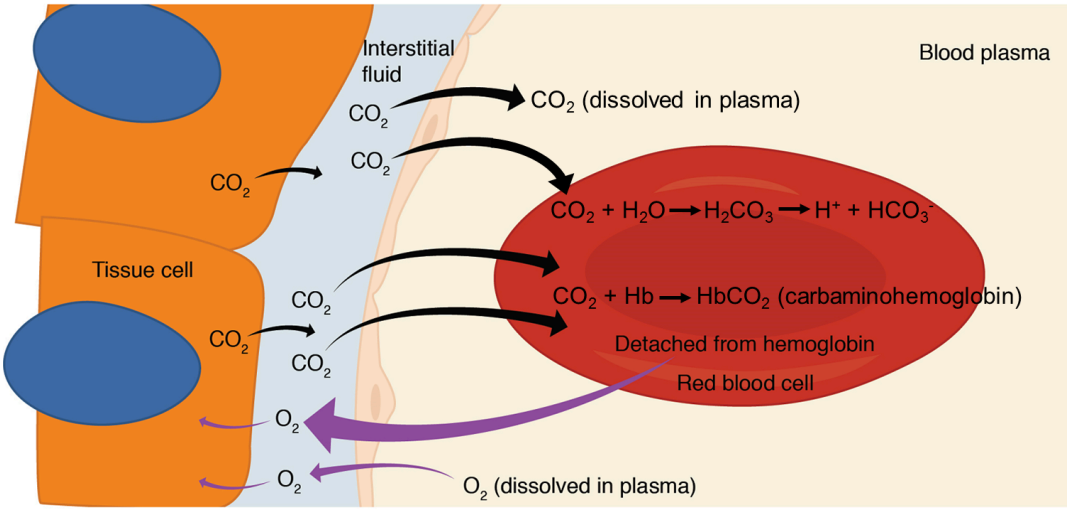


Figure 18. Internal Respiration. *Oxygen diffuses out of the capillary and into cells, whereas carbon dioxide diffuses out of cells and into the capillary.*

By the time blood returns to the heart, the partial pressure of oxygen has dropped, and the partial pressure of carbon dioxide has returned increased. The blood is then pumped back to the lungs to be oxygenated once again during external respiration.

Part 4: Transport of Gases

The function of respiration is to provide oxygen for use by body cells during cellular respiration and to eliminate carbon dioxide, a waste product of cellular respiration, from the body. In order for the exchange of oxygen and carbon dioxide to occur, both gases must be transported between the external and internal respiration sites. Although carbon dioxide is more soluble than oxygen in blood, both gases require a specialized transport system for the majority of the gas molecules to be moved between the lungs and other tissues.

Oxygen Transport in the Blood: Even though oxygen is transported via the blood, you may recall that oxygen is not very soluble in water. A small amount of oxygen does dissolve in the **blood plasma** and is transported in the bloodstream, but it is only about 1.5% of the total amount. The majority of oxygen molecules are carried from the lungs to the body's tissues by a specialized transport system, which relies on the erythrocyte—the red blood cell. Erythrocytes contain a metalloprotein, **hemoglobin**, which serves to bind oxygen molecules to the erythrocyte (Figure 19). Hemoglobin is composed of four subunits. Each of the four subunits that make up

hemoglobin is arranged in a ring-like fashion, with an iron atom covalently bound to the heme in the center of each subunit

Heme is the portion of hemoglobin that binds oxygen. Therefore, one hemoglobin molecule is capable of carrying up to four molecules of oxygen. As oxygen (O_2) diffuses across the respiratory membrane from the alveolus to the capillary, it also diffuses into the red blood cell and is bound by hemoglobin (Hb). The following reversible chemical reaction describes the production of the final product, **oxyhemoglobin** ($Hb–O_2$), which is formed when oxygen binds to hemoglobin. Oxyhemoglobin is a bright red-colored molecule that contributes to the bright red color of oxygenated blood.

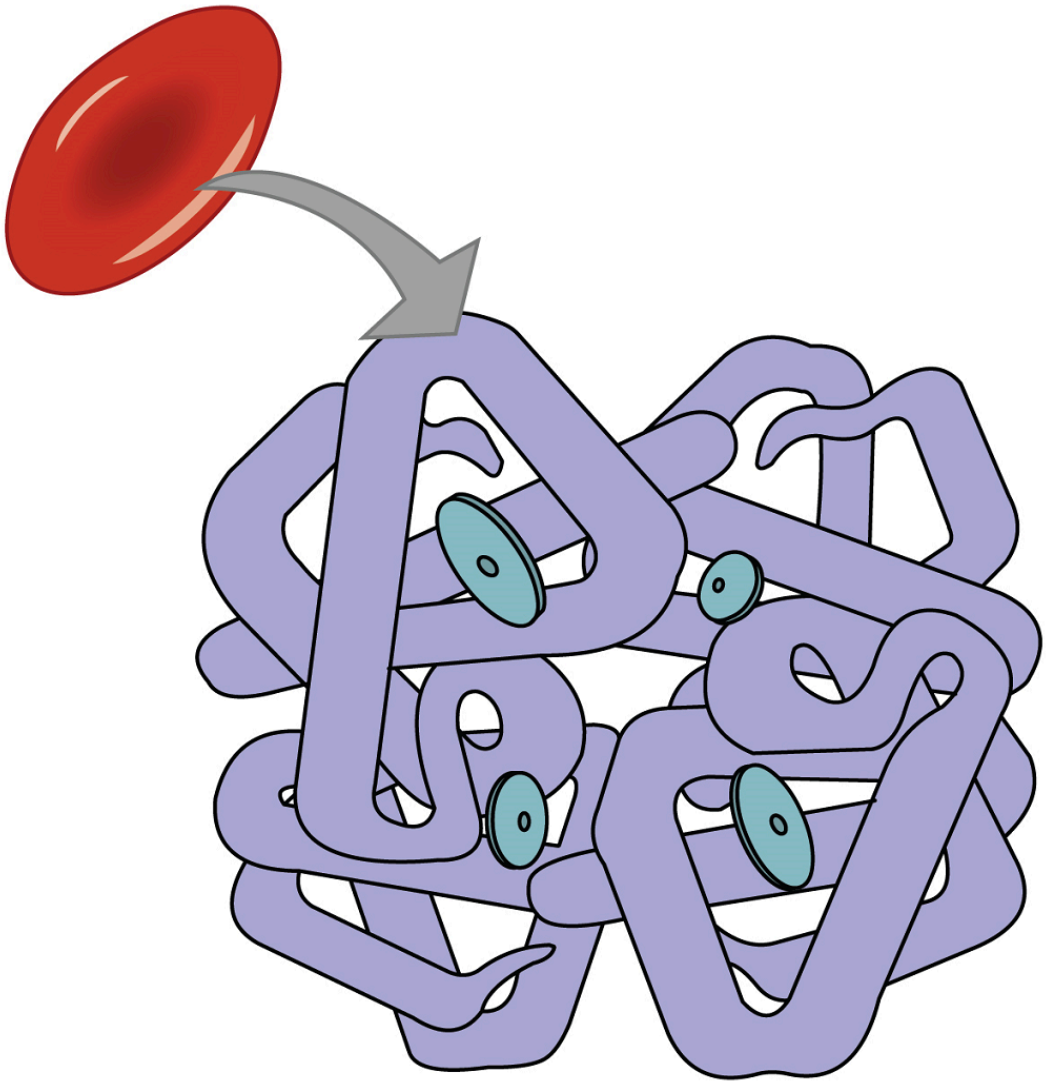


Figure 19. Erythrocyte and Hemoglobin. *Hemoglobin protein consists of four subunits, each of which contains one molecule of iron within the heme group (shown in blue).*

Binding of the first oxygen molecule causes a conformational change in hemoglobin that allows the second molecule of oxygen to bind more readily. As each molecule of oxygen is bound, it further facilitates the binding of the next molecule, until all four heme sites are occupied by oxygen. The opposite occurs as well: After the first oxygen molecule dissociates and is "dropped off" at the tissues, the next oxygen molecule dissociates more readily. When all four heme sites are occupied, the hemoglobin is said to be saturated. When one to three heme sites are occupied, the hemoglobin is said to be partially saturated. Therefore, when considering the blood as a whole, the percent of the available heme units that are bound to oxygen at a given time is called hemoglobin

saturation. In a healthy individual with normal hemoglobin levels, hemoglobin saturation generally ranges from 95 percent to 99 percent.

$$Hb + O_2 \leftrightarrow HbO_2$$

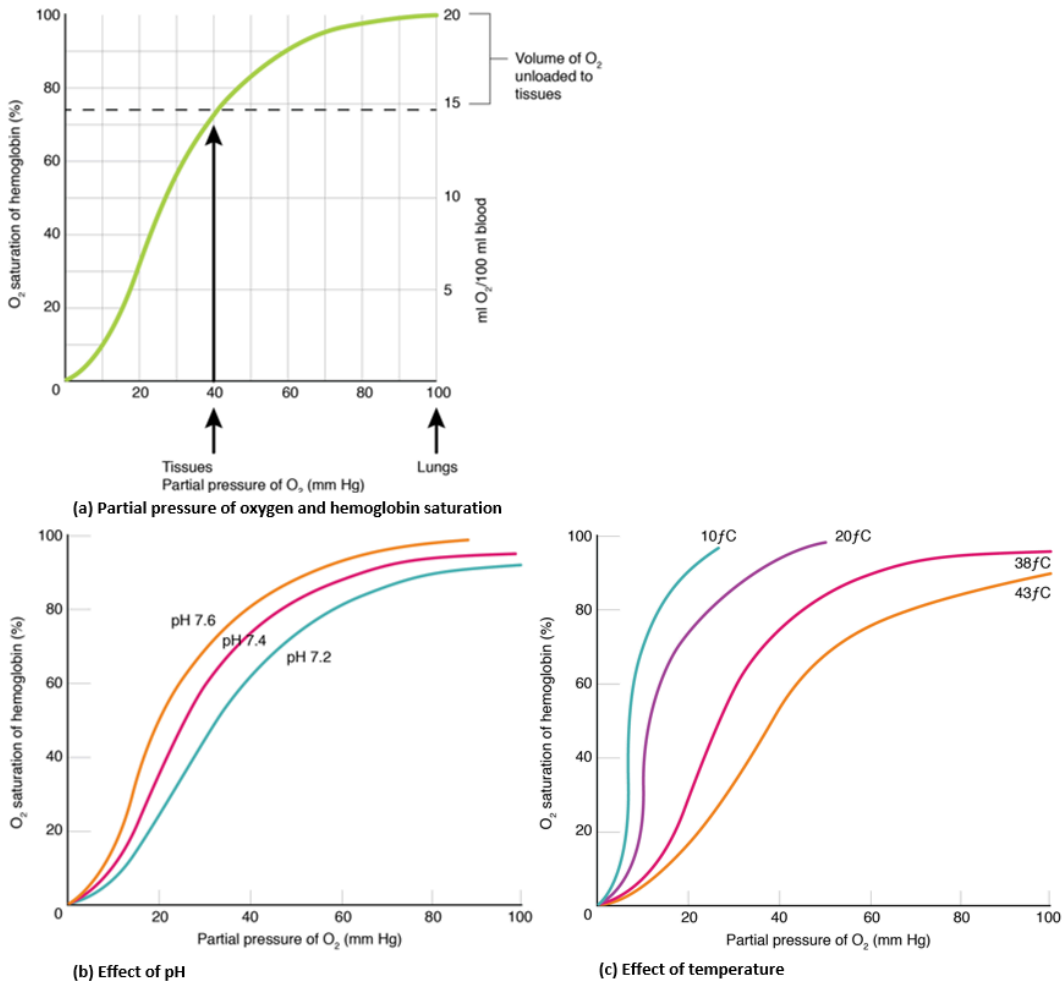


Figure 20. Oxygen-Hemoglobin Dissociation and Effects of pH and Temperature. *These three graphs show (a) the relationship between the partial pressure of oxygen and hemoglobin saturation, (b) the effect of pH on the oxygen–hemoglobin dissociation curve, and (c) the effect of temperature on the oxygen–hemoglobin dissociation curve.*

The mechanisms behind the oxygen–hemoglobin saturation/dissociation curve also serve as automatic control mechanisms that regulate how much oxygen is delivered to different tissues throughout the body. This is important because some tissues have a higher metabolic rate than others. Highly active tissues, such as muscle, rapidly use oxygen to produce ATP, lowering the partial pressure of oxygen in the tissue fluid. The difference in partial pressure of oxygen in the muscle tissue and the capillaries becomes quite high. As a result, a greater number of oxygen molecules dissociate from hemoglobin and enter the tissues. The reverse is true of tissues, such as adipose (body fat), which have lower metabolic rates. Because less oxygen is used by these cells, the partial pressure of oxygen within such tissues remains relatively high, resulting in fewer oxygen molecules dissociating from hemoglobin and entering the tissue fluid.

Although venous blood is said to be deoxygenated, some oxygen is still bound to hemoglobin in its red blood cells. This provides an oxygen reserve that can be used when tissues suddenly demand more oxygen.

Factors other than partial pressure of oxygen also affect the oxygen–hemoglobin saturation/dissociation curve. For example, a **higher temperature** promotes hemoglobin and oxygen to dissociate faster, whereas a lower temperature inhibits dissociation (see Figure 20c). However, the human body tightly regulates

temperature, so this factor may not affect gas exchange throughout the body. The exception to this is in highly active tissues, which may release a larger amount of energy than is given off as heat. As a result, oxygen readily dissociates from hemoglobin, which is a mechanism that helps to provide active tissues, such as muscles, with more oxygen.

The pH of the blood is another factor that influences the oxygen–hemoglobin saturation/dissociation curve (see Figure 20b). A lower, more acidic pH promotes oxygen dissociation from hemoglobin. In contrast, a higher, or more basic, pH inhibits oxygen dissociation from hemoglobin. The greater the amount of carbon dioxide in the blood, the more molecules that must be converted to carbonic acid, which in turn generates hydrogen ions and thus lowers blood pH. Furthermore, blood pH may become more acidic when certain byproducts of cell metabolism, such as lactic acid, carbonic acid, and carbon dioxide, are released into the bloodstream. As a result, in metabolically active tissues that create more metabolic acids, oxygen delivery is increased.

Carbon Dioxide Transport in the Blood: Carbon dioxide is transported by three major mechanisms. The first mechanism of carbon dioxide transport is by **blood plasma**, as some carbon dioxide molecules dissolve in the blood. The second mechanism is transport in the form of **bicarbonate (HCO_3^-)**, which also dissolves in plasma. The third mechanism of carbon dioxide transport is similar to the transport of oxygen by **hemoglobin** in **erythrocytes** (Figure 21).

1. Dissolved Carbon Dioxide: Although carbon dioxide is not considered to be highly soluble in blood, a small fraction—about 7 to 10 percent—of the carbon dioxide that diffuses into the blood from the tissues dissolves in plasma. The dissolved carbon dioxide then travels in the bloodstream and when the blood reaches the pulmonary capillaries, the dissolved carbon dioxide diffuses across the respiratory membrane into the alveoli, where it is then exhaled during pulmonary ventilation.

2. Bicarbonate Buffer: A large fraction—about 70 percent—of the carbon dioxide molecules that diffuse into the blood is transported to the lungs as bicarbonate (HCO_3^-). Most bicarbonate is produced in erythrocytes after carbon dioxide diffuses into the capillaries, and subsequently into red blood cells. Carbonic anhydrase (CA), an enzyme in red blood cells, causes carbon dioxide (CO_2) and water (H_2O) to form carbonic acid (H_2CO_3), which dissociates into two ions: a bicarbonate ion ($HCO3^-$) and a hydrogen ion (H^+). The following formula depicts this reaction:

$$CO_2 + H_2O \leftrightarrow H_2CO_3 \leftrightarrow H^+ + HCO_3^-$$

Bicarbonate tends to build up in the erythrocytes, so that there is a greater concentration of bicarbonate in the erythrocytes than in the surrounding blood plasma. As a result, some of the bicarbonate will leave the erythrocytes and move down its concentration gradient into the plasma.

At the pulmonary capillaries, the chemical reaction that produced bicarbonate (shown above) is reversed, and carbon dioxide and water are the products. Carbon dioxide diffuses out of the erythrocytes and into the plasma, where it can further diffuse across the respiratory membrane into the alveoli to be exhaled during pulmonary ventilation.

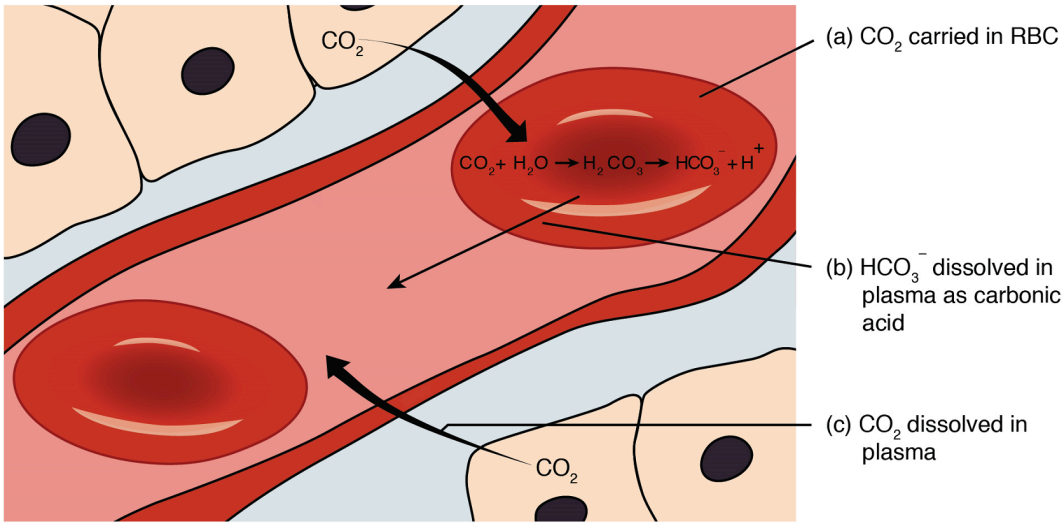


Figure 21. Carbon Dioxide Transport. *Carbon dioxide is transported by three different methods: (a) in erythrocytes on hemoglobin; (b) after forming carbonic acid (H2CO3), as HCO3- ion which is dissolved in plasma; (c) and in plasma.*

3. Carbaminohemoglobin: About 20 percent of carbon dioxide is bound by hemoglobin and is transported to the lungs. Carbon dioxide does not bind to iron as oxygen does; instead, carbon dioxide (CO_2) binds to the amino acid on the protein portions of hemoglobin (Hb) to form carbaminohemoglobin ($HbCO_2$). When hemoglobin is not transporting oxygen, it tends to have a bluish-purple tone to it, creating the darker burgundy color typical of deoxygenated blood. The following formula depicts this reversible reaction:

$$CO_2 + Hb \leftrightarrow HbCO_2$$

Similar to the transport of oxygen by heme, the binding and dissociation of carbon dioxide to and from hemoglobin is dependent on the partial pressure of carbon dioxide. Because carbon dioxide is released from the lungs, blood that leaves the lungs and reaches body tissues has a lower partial pressure of carbon dioxide than is found in the tissues.

As a result, carbon dioxide leaves the tissues because of its higher partial pressure, enters the blood, and then moves into red blood cells, binding to hemoglobin. In contrast, in the pulmonary capillaries, the partial pressure of carbon dioxide is high compared to within the alveoli. As a result, carbon dioxide dissociates readily from hemoglobin and diffuses across the respiratory membrane into the air.

Watch this CrashCourse video for an overview of how oxygen is exchanged! Direct link: https://youtu.be/Cqt4LjHnMEA

Part 5: Modifications in Respiratory Functions

At rest, the respiratory system performs its functions at a constant, rhythmic pace, as regulated by the respiratory centers of the brain. At this pace, ventilation provides sufficient oxygen to all the tissues of the body. However, there are times that the respiratory system must alter the pace of its functions in order to accommodate the oxygen demands of the body.

Hyperpnea: Hyperpnea is an increased depth and rate of ventilation to meet an increase in oxygen demand

as might be seen in exercise or disease, particularly diseases that target the respiratory or digestive tracts. This does not significantly alter blood oxygen or carbon dioxide levels, but merely increases the depth and rate of ventilation to meet the demand of the cells. In contrast, **hyperventilation** is an increased ventilation rate that is independent of the cellular oxygen needs and leads to abnormally low blood carbon dioxide levels and high (alkaline) blood pH.

Hyperventilation can occur for a multitude of reasons. It may be caused by abnormal functioning of the lungs, as a result of conditions such as asthma or early emphysema. It can occur at high altitudes where partial pressure of oxygen decreases (Table 2) and leads to lower hemoglobin saturation in the blood; hemoglobin saturation is about 67 percent at 19,000 feet above sea level, whereas it reaches about 98 percent at sea level. Finally, hyperventilation may be caused by increased metabolism as a result of such conditions as hyperthyroidism, infection, or fever.

Table 2: Partial Pressure of Oxygen at Different Altitudes

Example location	Altitude (km above sea level)	Atmospheric pressure (mmHg)	Partial pressure of oxygen (mmHg)
Vancouver, British Columbia	0-0.5 km	707-760	148-159
Lake Louise, Alberta	1.6 km	638	133
Aspen, Colorado	2.4 km	565	118
Mount Logan, Yukon	6.0 km	324	67
Mount Everest, Tibet	8.8 km	260	54

Although it has no effect on oxygen levels in the blood, hyperventilation significantly reduces the amount of carbon dioxide in the blood. This reduction in carbon dioxide levels in turn leads to reduced carbonic acid levels in the blood, which results in **alkalosis** (blood plasma pH higher than normal). Decreased blood carbon dioxide also **decreases blood pressure**, as the signals coming from peripheral carbon dioxide receptors (normally stimulated by CO_2) decrease in frequency and cause the vasomotor center in the medulla oblongata to reduce constriction of the smooth muscle in the walls of blood vessels and allow vasodilation. Finally, low carbon dioxide and the associated high pH interfere with the ability of hemoglobin to release oxygen molecules to body tissues, including the brain, which can cause **dizziness or unconsciousness**.

Hypoxia: Hypoxia is a reduction in the amount of oxygen reaching body tissues. It may be caused by a deficiency in atmospheric oxygen, whether due to high altitude (as described above) or being in an enclosed space with limited airflow (e.g. a crowded room with poor ventilation).

Hypoxia may also be caused by physiological problems with the respiratory or cardiovascular system. In the case of the respiratory system, any interference in the process of breathing (e.g. abnormal muscle contractions) or obstruction in the air passages (e.g. excessive mucus) will cause hypoxia by **ventilatory deficiency**. Alternatively, hypoxia may be caused by a **pulmonary diffusion defect** in which the diffusion of oxygen gas across the respiratory membrane is impaired. Fluid in the pulmonary alveoli, for example, increases the distance across which oxygen must diffuse through liquid, effectively increasing the thickness of the respiratory membrane and therefore slowing the rate at which oxygen can move into the blood. In the cardiovascular system, hypoxia may be caused by a **hemoglobin deficiency** or a **circulatory deficiency**. A hemoglobin deficiency may be the result of anemia, where there is a shortage of functional red blood cells. It may also be the result of conditions such as carbon monoxide poisoning, where carbon monoxide displaces oxygen bound to hemoglobin molecules, rendering the hemoglobin incapable of carrying oxygen and thus effectively nonfunctional. Circulatory deficiencies may be the result of obstruction of a blood vessel (e.g. as a result of atherosclerosis), of low blood pressure (hypotension), or of structural problems that make the cardiovascular system less efficient than it should normally be (e.g. if the ductus arteriosus or foramen ovale fail to close after birth).

Finally, hypoxia may result from **edema**, where excessive fluid accumulates around cells, for example as a

result of inflammation, renal failure, or congestive heart failure. This fluid buildup may occur in the lung tissue or alveoli (pulmonary edema), where it slows the diffusion of oxygen across respiratory membranes, or in other tissues where it can slow the diffusion of oxygen to body cells.

Hypoxia can result in cyanosis, where the skin and mucous membranes take on a bluish (or purplish) discoloration. It can also result in tachycardia, or increased heart rate, and dizziness as a result of insufficient oxygen reaching the brain.

Unit 5: The Digestive System

Unit outline

Part 1: Overview of the Digestive System

- Digestive System Organs
- Histology of the Alimentary Canal

Part 2: Digestive System Processes and Regulation

- Digestive Processes
- Regulatory Mechanisms

Part 3: The Mouth, Pharynx, and Esophagus

- The Mouth
- The Pharynx
- The Esophagus

Part 4: The Stomach

- Struture
- Digestive Functions of the Stomach

Part 5: The Small and Large Intestines

- The Small Intestine
- The Large Intestine
- Absorption, Feces Formation, and Defecation

Part 6: Accessory Organs in Digestion: The Liver, Pancreas, and Gallbladder

- The Liver
- The Pancreas
- The Gallbladder

Part 7: Chemical Digestion and Absorption: A Closer Look

- Carbohydrate Digestion
- Protein Digestion
- Lipid Digestion
- Nucleic Acid Digestion
- Absorption

Learning Objectives

At the end of this unit, you should be able to:

I. Describe the major functions of the digestive system.

II. Describe the relationship between the following processes in the gastrointestinal system: ingestion, digestion, absorption, defecation.

III. Distinguish between extracellular digestion and intracellular digestion.

IV. Describe the anatomy of the buccal cavity and explain its functions in digestion.

V. Describe the process of deglutition (swallowing), explaining why food, when swallowed, does not enter the respiratory tract or the nasal cavity.

VI. Describe the anatomy and functions of the esophagus.

VII. Describe the anatomy and functions of the stomach.

VIII. Describe the liver with reference to: anatomy, function, connection to the duodenum and gallbladder, blood supply.

IX. Describe the anatomy and functions of the pancreas.

X. Describe the anatomy and functions of the small intestine.

XI. Describe the anatomy and functions of the large intestine.

XII. Describe the process of defecation

XIII. Describe the chemical digestion of the following, specifying the source and the function of the principal enzymes involved: carbohydrates, proteins, lipids, nucleic acids.

XIV. Specify the end-products of the digestion of the following and explain how they are absorbed: carbohydrates, proteins, lipids, nucleic acids.

XV. Describe the control of the secretion of digestive juices in humans in terms of: nervous control, hormonal control.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Describe the major functions of the digestive system.

1. Describe the six major processes occurring during digestive system activity, and list all the organs of the gastrointestinal tract that perform each one.

II. Describe the relationship between the following processes in the gastrointestinal system: ingestion, digestion, absorption, defecation.

1. Clearly define each of the following terms as they relate to the gastrointestinal system:
    - Ingestion
    - Digestion
    - Absorption
    - Defecation

III. Distinguish between extracellular digestion and intracellular digestion.

1. Describe and clearly distinguish between extracellular digestion and intracellular digestion and state the specific location(s) in the human body where extracellular digestion occurs.

IV. Describe the anatomy of the buccal cavity and explain its functions in digestion.

1. Describe the anatomy of the buccal cavity, specifying the relative location and major tissue type(s) of each of the following structures:
    - Lips
    - Cheeks
    - Hard palate
    - Soft palate
    - Uvula
    - Teeth
    - Tongue
    - Salivary glands
2. Describe how each of the following structures contributes to the food-related functions served by the buccal cavity:
    - Lips
    - Cheeks
    - Hard palate
    - Soft palate
    - Uvula
    - Teeth
    - Tongue
    - Salivary glands

V. Describe the process of deglutition (swallowing), explaining why food, when swallowed, does not enter the respiratory tract or the nasal cavity.

1. Describe the process of deglutition in terms of its three major phases, describing the function and the neural control of each step.

VI. Describe the anatomy and functions of the esophagus.

1. Describe the anatomy of the esophagus by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its overall structure.

- The layers of tissue of which it is composed.

2. Describe how each tissue layer of the esophagus contributes to the primary function of the esophagus.

VII. Describe the anatomy and functions of the stomach.

1. Describe the anatomy of the stomach by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its overall structure.
    - The layers of tissue of which it is composed.
2. Describe how each tissue layer of the stomach performs (or contributes to):
    - Propulsion.
    - Mechanical digestion.
    - Chemical digestion.
3. Name the four secretory cell types that make up each gastric gland. For each cell type, state the product(s) it secretes and the function of its product(s).
4. Name one hormone secreted by the stomach, and state:
    - Its specific site (tissue and/or cell type) of production.
    - The stimulus for its production.
    - In which organ its target cells are located.
    - The effect(s) of its release.

VIII. Describe the liver with reference to: anatomy, function, connection to the duodenum and gallbladder, blood supply.

1. Describe the anatomy of the liver by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its connections to organs of the gastrointestinal tract and other accessory organs of the digestive system.
2. Describe the function served by the liver as part of the digestive system.
3. Describe the vasculature delivering blood to and from the liver, and explain how this vasculature relates to the functions served by the liver in the body.

IX. Describe the anatomy and functions of the pancreas.

1. Describe the anatomy of the pancreas by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its connections to organs of the gastrointestinal tract.
2. Describe and distinguish between the endocrine and exocrine functions of the pancreas.

X. Describe the anatomy and functions of the small intestine.

1. Describe the anatomy of the small intestine by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its three main anatomical subdivisions.
    - Its connections to other organs of the gastrointestinal tract, and to accessory organs of the digestive system.
2. Explain in detail how the small intestine performs (or contributes to):
    - Propulsion.
    - Mechanical digestion.
    - Chemical digestion.
    - Absorption.
3. Name two hormones secreted by the small intestine. For each hormone, state:
    - The stimulus for its production.
    - In which organ its target cells are located.
    - The effect(s) of its release.

XI. Describe the anatomy and functions of the large intestine.

1. Describe the anatomy of the large intestine by using correct anatomical terms to describe:
    - Its location in the human body.
    - Its main anatomical subdivisions.
    - The layers of tissue of which it is composed.
    - Its connections to other organs of the gastrointestinal tract.
2. Explain in detail how the large intestine performs (or contributes to):
    - Propulsion.
    - Mechanical digestion.
    - Chemical digestion.
    - Absorption.
    - Defecation.

XII. Describe the process of defecation

1. Describe the process of defecation, explaining the function and the neural control of each step.

XIII. Describe the chemical digestion of the following, specifying the source and the function of the principal enzymes involved: carbohydrates, proteins, lipids, nucleic acids.

XIV. Specify the end-products of the digestion of the following and explain how they are absorbed: carbohydrates, proteins, lipids, nucleic acids.

1. Describe the function of all the enzymes involved in carbohydrate digestion in the gastrointestinal tract. For each enzyme, state its name, source organ, site of action, substrate, and product.
2. Name the organ in the gastrointestinal tract within which the majority of chemical digestion of carbohydrates occurs.

3. Specify the end products of the carbohydrate digestion that occurs in the gastrointestinal tract.
4. Explain where and how each end product of carbohydrate digestion ultimately is absorbed from the lumen of the gastrointestinal tract into the blood.
5. Name the organ in the gastrointestinal tract within which the majority of chemical digestion of proteins occurs.
6. Specify the end products of the protein digestion that occurs in the gastrointestinal tract.
7. Explain where and how each end product of protein digestion ultimately is absorbed from the lumen of the gastrointestinal tract into the blood.
8. Describe the function of all the enzymes involved in lipid digestion in the gastrointestinal tract. For each enzyme, state its name, source organ, site of action, substrate, and product.
9. Name the organ in the gastrointestinal tract within which the majority of chemical digestion of lipids occurs.
10. Specify the end products of the lipid digestion that occurs in the gastrointestinal tract. Explain where and how each end product ultimately is absorbed from the lumen of the gastrointestinal tract into the blood.
11. Describe the type(s) of molecules that the end products of lipid digestion can be reassembled into, and what other functions they might serve.
12. Describe the function of all the enzymes involved in nucleic acid digestion in the gastrointestinal tract. For each enzyme, state its name, source organ, site of action, substrate, and product.
13. Name the organ in the gastrointestinal tract within which the majority of chemical digestion of nucleic acids occurs.
14. Specify the end products of the nucleic acid digestion that occurs in the gastrointestinal tract.
15. Explain where and how each end product of nucleic acid digestion ultimately is absorbed from the lumen of the gastrointestinal tract into the blood.
16. Describe the type(s) of molecules that the end products of nucleic acid digestion can be reassembled into, and what other functions they might serve.

XV. Describe the control of the secretion of digestive juices in humans in terms of: nervous control, hormonal control.

1. Describe the pathways by which the nervous system regulates:
    - Gastric secretory activity during the cephalic phase of gastric secretion.
    - Gastric secretory activity during the gastric phase of gastric secretion.
    - Gastric secretory activity during the intestinal phase of gastric secretion.
2. Describe the hormonal regulation of:
    - Gastric secretory activity during the gastric phase of gastric secretion.
    - Gastric secretory activity during the intestinal phase of gastric secretion.
    - Bile production and release.
    - Pancreatic juice production and release.

The Digestive System: The digestive system is continually at work, yet people seldom appreciate the complex tasks it performs in a choreographed biologic symphony. Consider what happens when you eat an apple. Of course, you enjoy the apple's taste as you chew it, but in the hours that follow, unless something goes amiss

and you get a stomachache, you don't notice that your digestive system is working. You may be taking a walk or studying or sleeping, having forgotten all about the apple, but your stomach and intestines are busy digesting it and absorbing its vitamins and other nutrients. By the time any waste material is excreted, the body has appropriated all it can use from the apple. In short, whether you pay attention or not, the organs of the digestive system perform their specific functions, allowing you to use the food you eat to keep you going. This chapter examines the structure and functions of these organs, and explores the mechanics and chemistry of the digestive processes.

Part 1: Overview of the Digestive System

The function of the digestive system is to break down the foods you eat, release their nutrients, and absorb those nutrients into the body. Although the small intestine is the workhorse of the system, where the majority of digestion occurs, and where most of the released nutrients are absorbed into the blood or lymph, each of the digestive system organs makes a vital contribution to this process (Figure 1).

As is the case with all body systems, the digestive system does not work in isolation; it functions cooperatively with the other systems of the body. Consider for example, the interrelationship between the digestive and cardiovascular systems. Arteries supply the digestive organs with oxygen and processed nutrients, and veins drain the digestive tract.

These intestinal veins, constituting the hepatic portal system, are unique; they do not return blood directly to the heart. Rather, this blood is diverted to the liver where its nutrients are off-loaded for processing before blood completes its circuit back to the heart. At the same time, the digestive system provides nutrients to the heart muscle and vascular tissue to support their functioning. The interrelationship of the digestive and endocrine systems is also critical. Hormones secreted by several endocrine glands, as well as endocrine cells of the pancreas, the stomach, and the small intestine, contribute to the control of digestion and nutrient metabolism. In turn, the digestive system provides the nutrients to fuel endocrine function. Table 1 gives a quick glimpse at how these other systems contribute to the functioning of the digestive system.

Digestive System Organs: The easiest way to understand the digestive system is to divide its organs into two main categories. The first group is the organs that make up the alimentary canal. Accessory digestive organs comprise the second group and are critical for orchestrating the breakdown of food and the assimilation of its nutrients into the body. Accessory digestive organs, despite their name, are critical to the function of the digestive system.

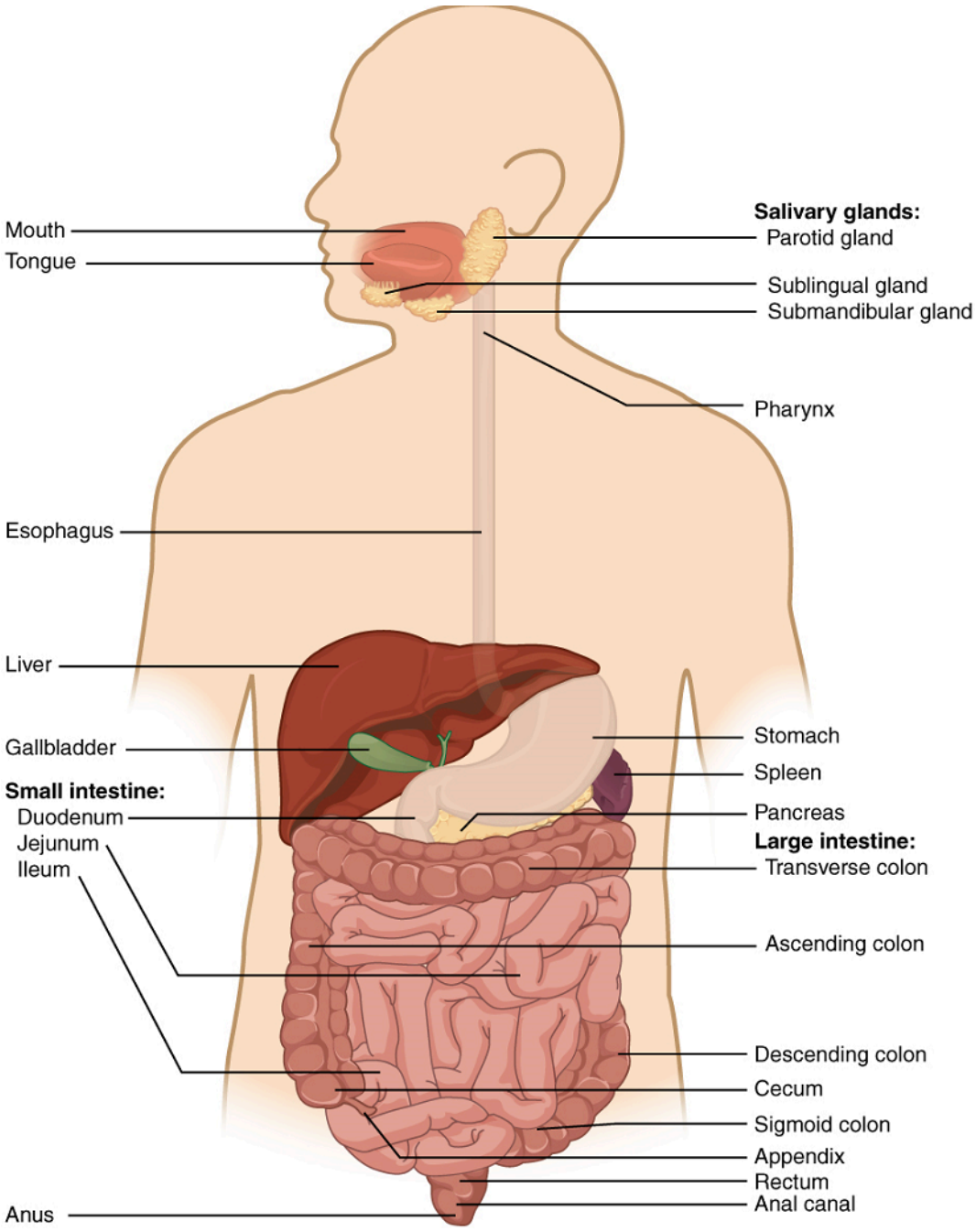


Figure 1. Components of the Digestive System. *All digestive organs play integral roles in the life-sustaining process of digestion.*

- **Alimentary Canal Organs:** Also called the gastrointestinal (GI) tract or gut, the alimentary canal (aliment- = "to nourish") is a one-way tube about 7.62 meters (25 feet) in length during life and closer to 10.67 meters (35 feet) in length when measured after death, once smooth muscle tone is lost. The main function of the organs of the alimentary canal is to nourish the body. This tube begins at the mouth and terminates at the anus. Between those two points, the canal is modified as the pharynx, esophagus, stomach, and small and large intestines to fit the functional needs of the body. Both the mouth and anus are open to the external environment; thus, food and wastes within the alimentary canal are technically considered to be outside the body. Only through the process of absorption do the nutrients in food enter into and nourish the body's "inner space."

- **Accessory Structures:** Each accessory digestive organ aids in the breakdown of food (Figure 2). Within the mouth, the teeth and tongue begin mechanical digestion, whereas the salivary glands begin chemical digestion. Once food products enter the small intestine, the gallbladder, liver, and pancreas release secretions—such as bile and enzymes—essential for digestion to continue. Together, these are called accessory organs because they sprout from the lining cells of the developing gut (mucosa) and augment its function; indeed, you could not live without their vital contributions, and many significant diseases result from their malfunction. Even after development is complete, they maintain a connection to the gut by way of ducts.

Table 1: Contribution of Other Body Systems to the Digestive System

Body system	Benefits received by the digestive system
Cardiovascular	Blood supplies digestive organs with oxygen and processed nutrients; capillaries receive absorbed nutrients
Endocrine	Hormones help regulate secretion in digestive glands and accessory organs
Integumentary	Skin helps protect digestive organs and synthesizes vitamin D to facilitate calcium absorption
Lymphatic	Mucosa-associated lymphoid tissue defend against entry of pathogens; lacteals absorb lipids; lymphatic vessels transport lipids to bloodstream
Muscular	Skeletal muscles support and protect abdominal organs
Nervous	Sensory and motor neurons help regulate secretions and muscle contractions in the digestive tract
Respiratory	Respiratory organs provide oxygen and remove carbon dioxide
Skeletal	Bones help protect and support digestive organs
Urinary	Kidneys convert vitamin D into its active form, allowing calcium absorption in the small intestine

Histology of the Alimentary Canal: Throughout its length, the alimentary tract is composed of the same four tissue layers; the details of their structural arrangements vary to fit their specific functions. Starting from the lumen and moving outwards, these layers are the mucosa, submucosa, muscularis, and serosa, which is continuous with the mesentery (Figure 2).

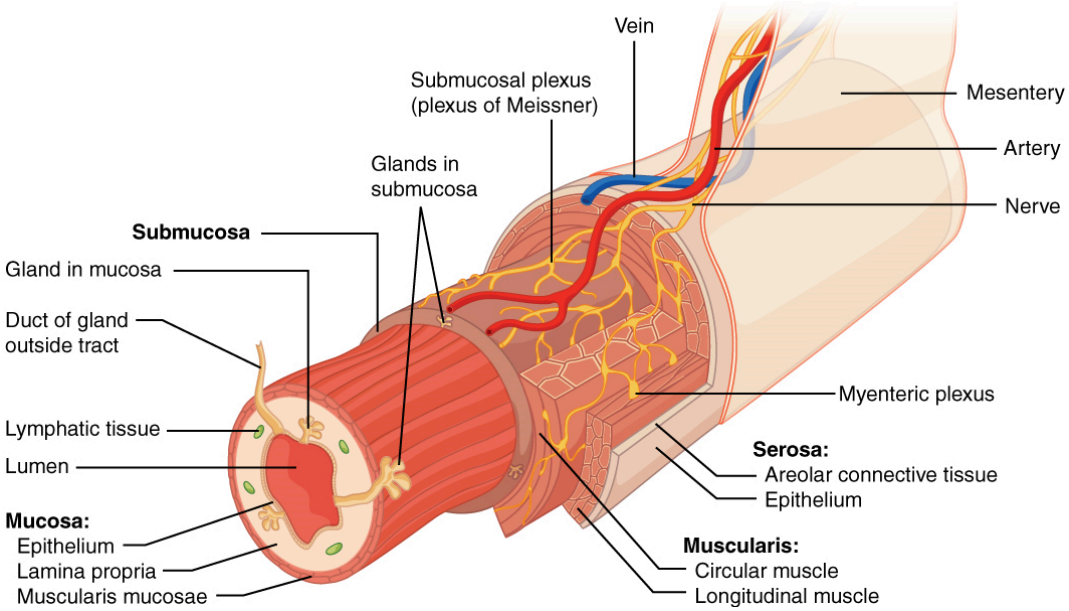


Figure 2. Layers of the Alimentary Canal. *The wall of the alimentary canal has four basic tissue layers: the mucosa, submucosa, muscularis, and serosa.*

The **mucosa** is referred to as a mucous membrane, because mucus production is a characteristic feature of gut epithelium. The membrane consists of epithelium, which is in direct contact with ingested food, and the lamina propria, a layer of connective tissue analogous to the dermis. In addition, the mucosa has a thin, smooth muscle layer, called the muscularis mucosa.

- *Epithelium*—In the mouth, pharynx, esophagus, and anal canal, the epithelium is primarily a non-keratinized, stratified squamous epithelium. In the stomach and intestines, it is a simple columnar epithelium. Notice that the epithelium is in direct contact with the lumen, the space inside the alimentary canal. Interspersed among its epithelial cells are goblet cells, which secrete mucus and fluid into the lumen, and enteroendocrine cells, which secrete hormones into the interstitial spaces between cells. Epithelial cells have a very brief lifespan, averaging from only a couple of days (in the mouth) to about a week (in the gut). This process of rapid renewal helps preserve the health of the alimentary canal, despite the wear and tear resulting from continued contact with foodstuffs.
- *Lamina propria*—In addition to loose connective tissue, the lamina propria contains numerous blood and lymphatic vessels that transport nutrients absorbed through the alimentary canal to other parts of the body.
- *Muscularis mucosa*—This thin layer of smooth muscle is in a constant state of tension, pulling the mucosa of the stomach and small intestine into undulating folds. These folds dramatically increase the surface area available for digestion and absorption.

As its name implies, the **submucosa** lies immediately beneath the mucosa. A broad layer of dense connective tissue, it connects the overlying mucosa to the underlying muscularis. It includes blood and lymphatic vessels (which transport absorbed nutrients), and a scattering of submucosal glands that release digestive secretions. Additionally, it serves as a conduit for a dense branching network of nerves, the submucosal plexus, which functions as described below.

The third layer of the alimentary canal is the **muscularis** (also called the muscularis externa). The muscularis in the small intestine is made up of a double layer of smooth muscle: an inner circular layer and an outer longitudinal layer. The contractions of these layers promote mechanical digestion, expose more of the food to digestive chemicals, and move the food along the canal. In the most proximal and distal regions of the alimentary canal, including the mouth, pharynx, proximal part of the esophagus, and external anal sphincter, the muscularis is made up of skeletal muscle, which gives you voluntary control over swallowing and defecation. The basic two-layer structure found in the small intestine is modified in the organs proximal and distal to it. The stomach is equipped for its churning function by the addition of a third layer, the oblique muscle. While the colon has two layers like the small intestine, its longitudinal layer is segregated into three narrow parallel bands, the tenia coli, which make it look like a series of pouches rather than a simple tube.

The **serosa** is the portion of the alimentary canal superficial to the muscularis. Present only in the region of the alimentary canal within the abdominal cavity, it consists of a layer of visceral peritoneum overlying a layer of loose connective tissue. Instead of serosa, the mouth, pharynx, and esophagus have a dense sheath of collagen fibers called the adventitia. These tissues serve to hold the alimentary canal in place near the ventral surface of the vertebral column.

Nerve Supply: As soon as food enters the mouth, it is detected by receptors that send impulses along the sensory neurons of cranial nerves. Without these nerves, not only would your food be without taste, but you would also be unable to feel either the food or the structures of your mouth, and you would be unable to avoid biting yourself as you chew, an action enabled by the motor branches of cranial nerves.

Intrinsic innervation of much of the alimentary canal is provided by the enteric nervous system, which runs from the esophagus to the anus, and contains approximately 100 million motor, sensory, and interneurons (unique to this system compared to all other parts of the peripheral nervous system). (see Figure 2).

Blood Supply: The blood vessels serving the digestive system have two functions. They transport the protein

and carbohydrate nutrients absorbed by mucosal cells after food is digested in the lumen. Lipids are absorbed via lacteals, tiny structures of the lymphatic system. The blood vessels' second function is to supply the organs of the alimentary canal with the nutrients and oxygen needed to drive their cellular processes.

The proximal parts of the alimentary canal are supplied with blood by arteries branching off the aortic arch and thoracic aorta. Below this point, the alimentary canal is supplied with blood by arteries branching from the abdominal aorta. The celiac trunk services the liver, stomach, and duodenum, whereas the superior and inferior mesenteric arteries supply blood to the remaining small and large intestines.

The veins that collect nutrient-rich blood from the small intestine (where most absorption occurs) empty into the hepatic portal system. This venous network takes the blood into the liver where the nutrients are either processed or stored for later use. Only then does the blood drained from the alimentary canal viscera circulate back to the heart. To appreciate just how demanding the digestive process is on the cardiovascular system, consider that while you are "resting and digesting," about one-fourth of the blood pumped with each heartbeat enters arteries serving the intestines.

The Peritoneum: The digestive organs within the abdominal cavity are held in place by the peritoneum, a broad serous membranous sac made up of squamous epithelial tissue surrounded by connective tissue. It is composed of two different regions: the parietal peritoneum, which lines the abdominal wall, and the visceral peritoneum, which envelopes the abdominal organs (Figure 3). The peritoneal cavity is the space bounded by the visceral and parietal peritoneal surfaces. A few milliliters of watery fluid act as a lubricant to minimize friction between the serosal surfaces of the peritoneum.

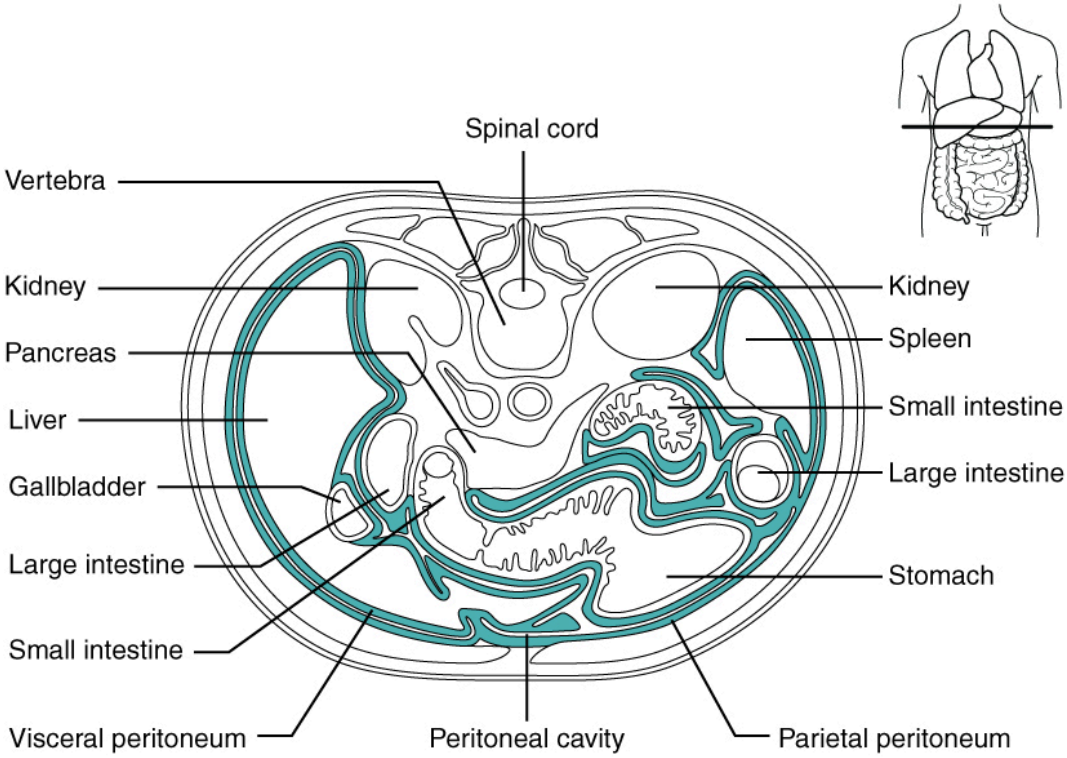


Figure 3. The Peritoneum. *A cross-section of the abdomen shows the relationship between abdominal organs and the peritoneum (darker lines).*

Watch this Crash Course video for an overview of the digestive system! Direct link: https://youtu.be/s06XzaKqELk

Part 2: Digestive System Processes and Regulation

The digestive system uses mechanical and chemical activities to break food down into absorbable substances during its journey through the digestive system. Table 2 provides an overview of the basic functions of the digestive organs.

Table 2: Functions of the Digestive Organs

Organ	Major functions	Other functions
Mouth	Ingests food Chews and mixes food Begins chemical breakdown of carbohydrates Moves food into pharynx Begins some breakdown of lipids via lingual lipase	Moistens and dissolves food, allowing taste Cleans and lubricates teeth and oral cavity Some antimicrobial activity
Pharynx	Propels food from oral cavity to esophagus	Lubricates food and passageways
Esophagus	Propels food to stomach	Lubricates food and passageways
Stomach	Mixes and churns food with gastric juices to form chyme Begins chemical breakdown of proteins Enhances activity of lingual lipase Releases food into duodenum as chyme Absorbs some fat-soluble substance (e.g., alcohol, aspirin) Secretes antimicrobial substances	Stimulates protein-digesting enzymes Secretes intrinsic factor required for vitamin B12 absorption in small intestine
Small intestine	Mixes chyme with digestive juices Propels food at a rate slow enough for digestion and absorption Absorbs breakdown products of carbohydrates, proteins, lipids, nucleic acids Absorbs vitamins, minerals, water Performs physical digestion via segmentation	Provides optimal medium for enzymatic activity
Accessory organs	Liver: produces bile salts which emulsify lipids, aiding their digestion and absorption Gallbladder: stores, concentrates, and releases bile Pancreas: produces digestive enzymes and bicarbonate	Bicarbonate-rich pancreatic juice helps neutralize acidic chyme and provide optimal environment for enzymatic activity
Large intestine	Further breaks down food residues Absorbs most residual water, electrolytes, vitamins produced by enteric bacteria Propels feces toward rectum Eliminates feces	Concentrates and temporarily stored food residue prior to defecation Mucus eases passage of feces through colon

Digestive Processes: The processes of digestion include six activities: ingestion, propulsion, mechanical or physical digestion, chemical digestion, absorption, and defecation.

The first of these processes, **ingestion**, refers to the entry of food into the alimentary canal through the

mouth. There, the food is chewed and mixed with saliva, which contains enzymes that begin breaking down the carbohydrates in the food plus some lipid digestion via lingual lipase. Chewing increases the surface area of the food and allows an appropriately sized bolus to be produced.

Food leaves the mouth when the tongue and pharyngeal muscles propel it into the esophagus. This act of swallowing, the last voluntary act until defecation, is an example of **propulsion**, which refers to the movement of food through the digestive tract. It includes both the voluntary process of swallowing and the involuntary process of peristalsis. **Peristalsis** consists of sequential, alternating waves of contraction and relaxation of the longitudinal and circular smooth muscle layers in the wall of the alimentary canal, which act to propel food along (Figure 4). These waves also play a role in mixing food with digestive juices.

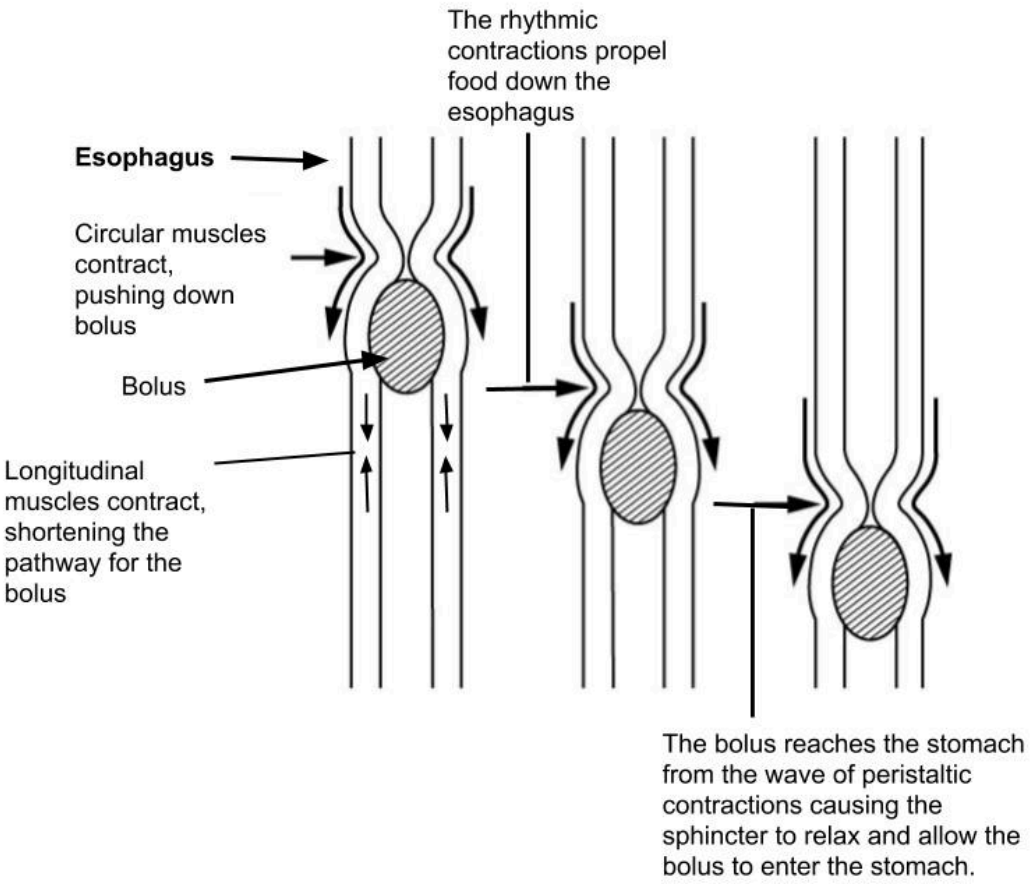


Figure 4. Peristalsis. *Peristalsis moves food through the digestive tract with alternating waves of muscle contraction and relaxation. (Image by Allison Calabrese CC-BY.)*

Digestion includes both mechanical and chemical processes. **Mechanical digestion** is a purely physical process that does not change the chemical nature of the food. Instead, it makes the food smaller to increase both surface area and mobility. It includes **mastication**, or chewing, as well as tongue movements that help break food into smaller bits and mix food with saliva. Although there may be a tendency to think that mechanical digestion is limited to the first steps of the digestive process, it occurs after the food leaves the mouth, as well. The mechanical churning of food in the stomach serves to further break it apart and expose more of its surface area to digestive juices, creating an acidic "soup" called **chyme**. **Segmentation**, which occurs mainly in the small intestine, consists of localized contractions of circular muscle of the muscularis layer of the alimentary canal. These contractions isolate small sections of the intestine, moving their contents back and forth while continuously subdividing, breaking up, and mixing the contents. By moving food back and forth in the intestinal lumen, segmentation mixes food with digestive juices and facilitates absorption.

In **chemical digestion**, starting in the mouth, digestive secretions break down complex food molecules into their chemical building blocks (for example, proteins into separate amino acids). These secretions vary in composition, but typically contain water, various enzymes, acids, and salts. The process is completed in the small intestine. Since this chemical digestion occurs in the lumen of the gastrointestinal tract as a result of secretions into the lumen, it is a form of **extracellular digestion**. (Contrast this with the intracellular digestion that occurs after phagocytosis, for example.)

Food that has been broken down is of no value to the body unless it enters the bloodstream and its nutrients are put to work. This occurs through the process of **absorption**, which takes place primarily within the small intestine. There, most nutrients are absorbed from the lumen of the alimentary canal into the bloodstream through the epithelial cells that make up the mucosa. Lipids are absorbed into lacteals and are transported via the lymphatic vessels to the bloodstream (the subclavian veins near the heart). The details of these processes will be discussed later.

In **defecation**, the final step in digestion, undigested materials are removed from the body as feces.

In some cases, a single organ is in charge of a digestive process. For example, ingestion occurs only in the mouth and defecation only in the anus. However, most digestive processes involve the interaction of several organs and occur gradually as food moves through the alimentary canal (Figure 5).

Some chemical digestion occurs in the mouth. Some absorption can occur in the mouth and stomach, for example, alcohol and aspirin.

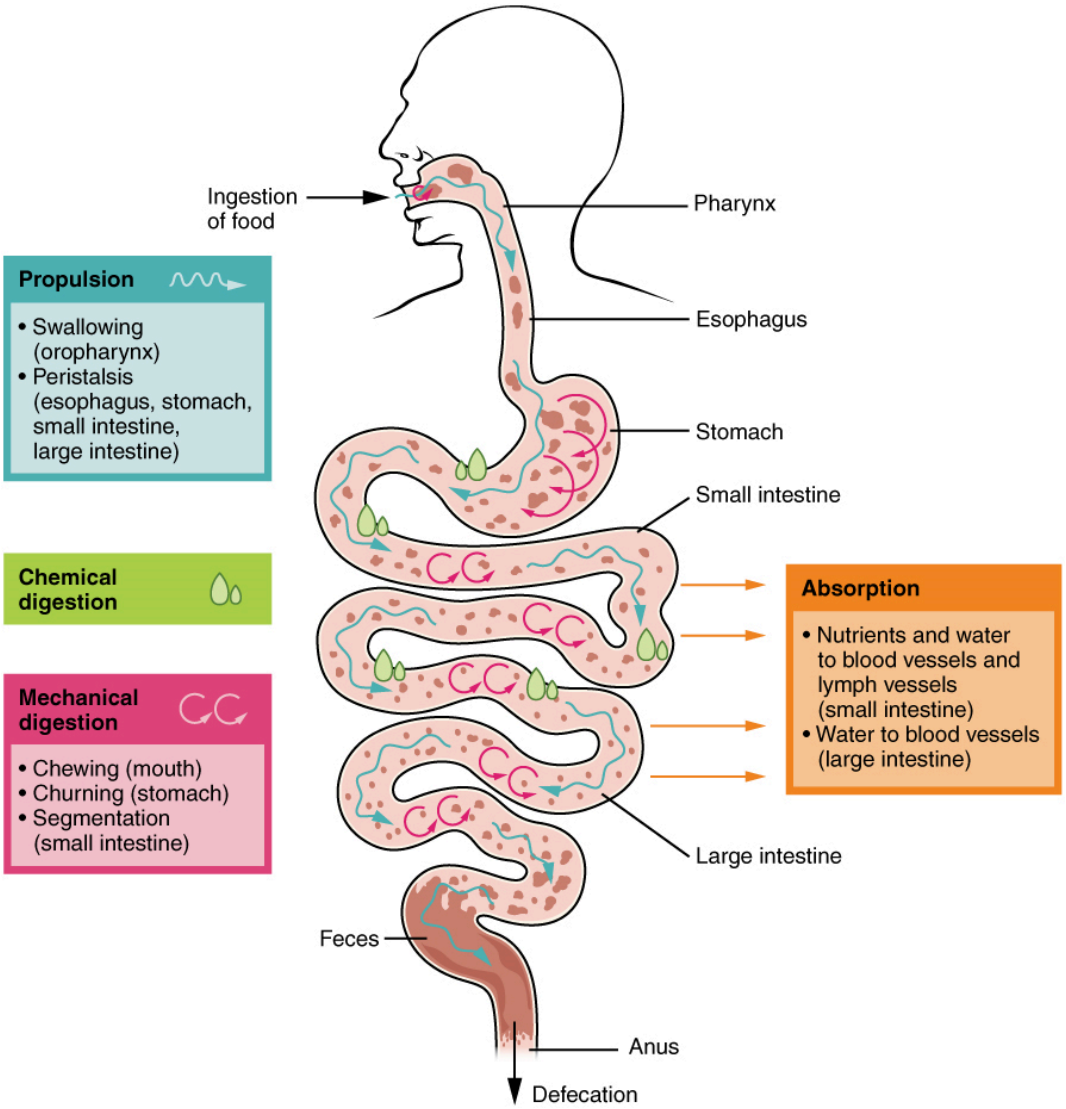


Figure 5. Digestive Processes. *The digestive processes are ingestion, propulsion, mechanical digestion, chemical digestion, absorption, and defecation.*

Regulatory Mechanisms: Neural and endocrine regulatory mechanisms work to maintain the optimal conditions in the lumen needed for digestion and absorption. These regulatory mechanisms, which stimulate digestive activity through mechanical and chemical activity, are controlled both extrinsically and intrinsically.

Neural Controls: The walls of the alimentary canal contain a variety of sensors that help regulate digestive functions. These include mechanoreceptors, chemoreceptors, and osmoreceptors, which are capable of detecting mechanical, chemical, and osmotic stimuli, respectively. For example, these receptors can sense when the presence of food has caused the stomach to expand, whether food particles have been sufficiently broken down, how much liquid is present, and the type of nutrients in the food (lipids, carbohydrates, and/or proteins). Stimulation of these receptors provokes an appropriate reflex that furthers the process of digestion. This may entail sending a message that activates the glands that secrete digestive juices into the lumen, or it may mean the stimulation of muscles within the alimentary canal, thereby activating peristalsis and segmentation that move food along the intestinal tract.

Hormonal Controls: A variety of hormones are involved in the digestive process. The main digestive hormone of the stomach is gastrin, which is secreted in response to the presence of food. Gastrin stimulates the secretion of gastric acid by the parietal cells of the stomach mucosa. Other GI hormones are produced and act upon the gut and its accessory organs. Hormones produced by the duodenum include secretin, which stimulates a watery secretion of bicarbonate by the pancreas; cholecystokinin (CCK), which stimulates the secretion of pancreatic enzymes and bile from the liver and release of bile from the gallbladder; and gastric inhibitory peptide, which inhibits gastric secretion and slows gastric emptying and motility.

Watch this Crash Course video to learn more about digestion! Direct link: https://youtu.be/yloTRGfcMqM

Part 3: The Mouth, Pharynx, and Esophagus

In this section, you will examine the anatomy and functions of the three main organs of the upper alimentary canal—the mouth, pharynx, and esophagus—as well as three associated accessory organs—the tongue, salivary glands, and teeth.

The Mouth: The cheeks, tongue, and palate frame the mouth, which is also called the **oral cavity** (or buccal cavity). The structures of the mouth are illustrated in Figure 6, and the digestive functions of the mouth are summarized in Table 3.

At the entrance to the mouth are the lips, or **labia** (singular = labium). Their outer covering is skin, which transitions to a mucous membrane in the mouth proper. Lips are very vascular with a thin layer of keratin; hence, the reason they are "red." They have a huge representation on the cerebral cortex, which probably explains the human fascination with kissing! The lips cover the orbicularis oris muscle, which regulates what comes in and goes out of the mouth. The **labial frenulum** is a midline fold of mucous membrane that attaches the inner surface of each lip to the gum. The cheeks make up the oral cavity's sidewalls. While their outer covering is skin, their inner covering is mucous membrane. This membrane is made up of non-keratinized, stratified squamous epithelium. Beneath the skin and mucous membranes are connective tissue and buccinator muscles. The next time you eat some food, notice how the buccinator muscles in your cheeks and the orbicularis oris muscle in your lips contract, helping you keep the food from falling out of your mouth. Additionally, notice how these muscles work when you are speaking.

The pocket-like part of the mouth that is framed on the inside by the gums and teeth, and on the outside by

the cheeks and lips is called the **oral vestibule**. Moving farther into the mouth, the opening between the oral cavity and throat (oropharynx) is called the **fauces** (like the kitchen "faucet"). The main open area of the mouth, or oral cavity proper, runs from the gums and teeth to the fauces.

When you are chewing, you do not find it difficult to breathe simultaneously. The next time you have food in your mouth, notice how the arched shape of the roof of your mouth allows you to handle both digestion and respiration at the same time. This arch is called the palate. The anterior region of the palate serves as a wall (or septum) between the oral and nasal cavities as well as a rigid shelf against which the tongue can push food. It is created by the maxillary and palatine bones of the skull and, given its bony structure, is known as the hard palate.

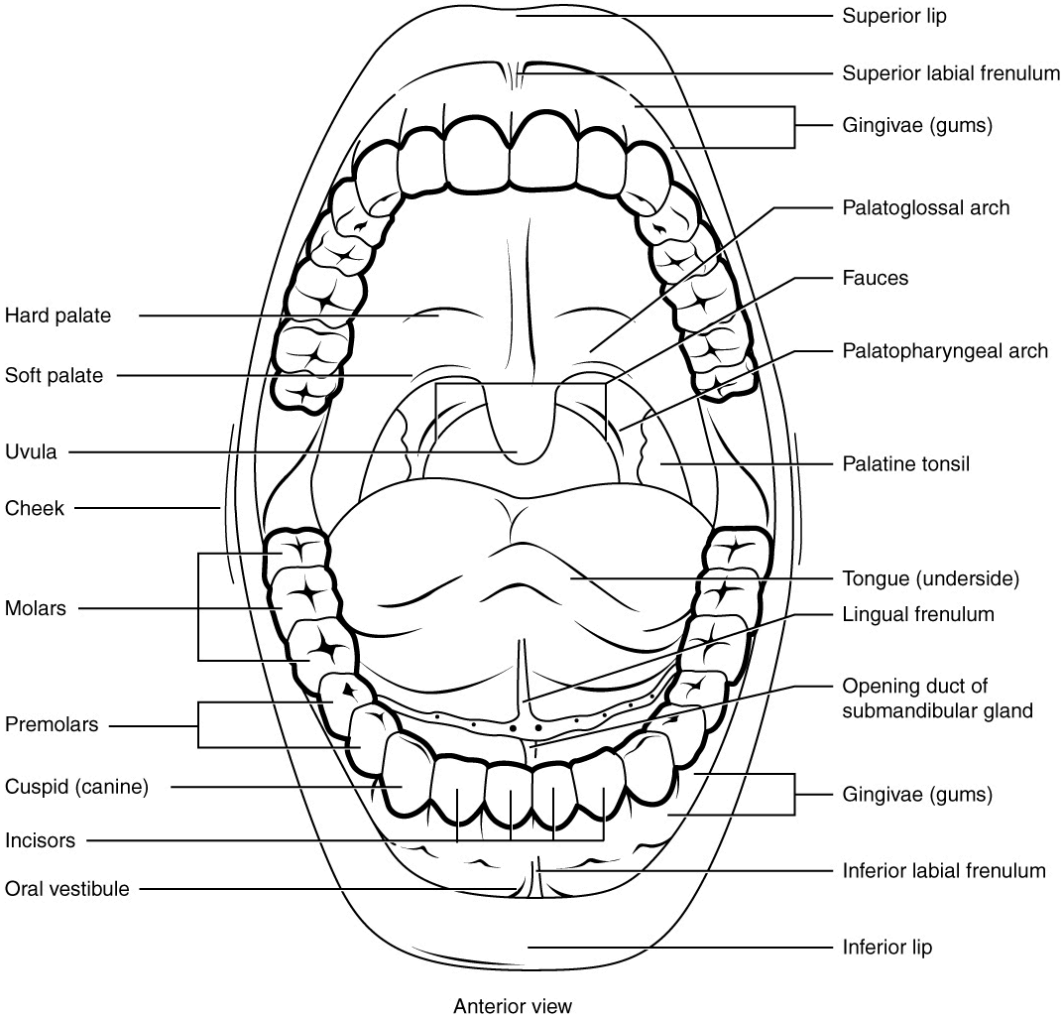


Figure 6. Mouth. *The mouth includes the lips, tongue, palate, gums, and teeth.*

If you run your tongue along the roof of your mouth, you'll notice that the hard palate ends in the posterior oral cavity, and the tissue becomes fleshier. This part of the palate, known as the **soft palate**, is composed mainly of skeletal muscle. You can therefore manipulate, subconsciously, the soft palate—for instance, to yawn, swallow, or sing (see Figure 6).

A fleshy bead of tissue called the **uvula** drops down from the center of the posterior edge of the soft palate. Although some have suggested that the uvula is a vestigial organ, it serves an important purpose. When you swallow, the soft palate and uvula move upward, helping to keep foods and liquid from entering the nasal cavity. Unfortunately, it can also contribute to the sound produced by snoring. Two muscular folds extend downward from the soft palate, on either side of the uvula

The Tongue: Perhaps you have heard it said that the **tongue** is the strongest muscle in the body. Those who stake this claim cite its strength proportionate to its size. Although it is difficult to quantify the relative strength of different muscles, it remains indisputable that the tongue is a workhorse, facilitating ingestion, mechanical digestion, chemical digestion (lingual lipase), sensation (of taste, texture, and temperature of food), swallowing, and vocalization.

The tongue is attached to the mandible, the styloid processes of the temporal bones, and the hyoid bone.

Working in concert, the muscles of the tongue perform three important digestive functions in the mouth: (1) position food for optimal chewing, (2) gather food into a **bolus** (rounded mass), and (3) position food so it can be swallowed.

The top and sides of the tongue are studded with papillae, extensions of lamina propria of the mucosa, which are covered in stratified squamous epithelium (Figure 7). Lingual glands in the lamina propria of the tongue secrete mucus and a watery serous fluid that contains the enzyme **lingual lipase**, which plays a minor role in breaking down triglycerides.

The Salivary Glands: Many small salivary glands are housed within the mucous membranes of the mouth and tongue. These minor exocrine glands are constantly secreting saliva, either directly into the oral cavity or indirectly through ducts, even while you sleep. In fact, an average of 1 to 1.5 liters of saliva is secreted each day. Usually just enough saliva is present to moisten the mouth and teeth. Secretion increases when you eat, because saliva is essential to moisten food and initiate the chemical breakdown of carbohydrates. Small amounts of saliva are also secreted by the labial glands in the lips. In addition, the buccal glands in the cheeks, palatal glands in the palate, and lingual glands in the tongue help ensure that all areas of the mouth are supplied with adequate saliva.

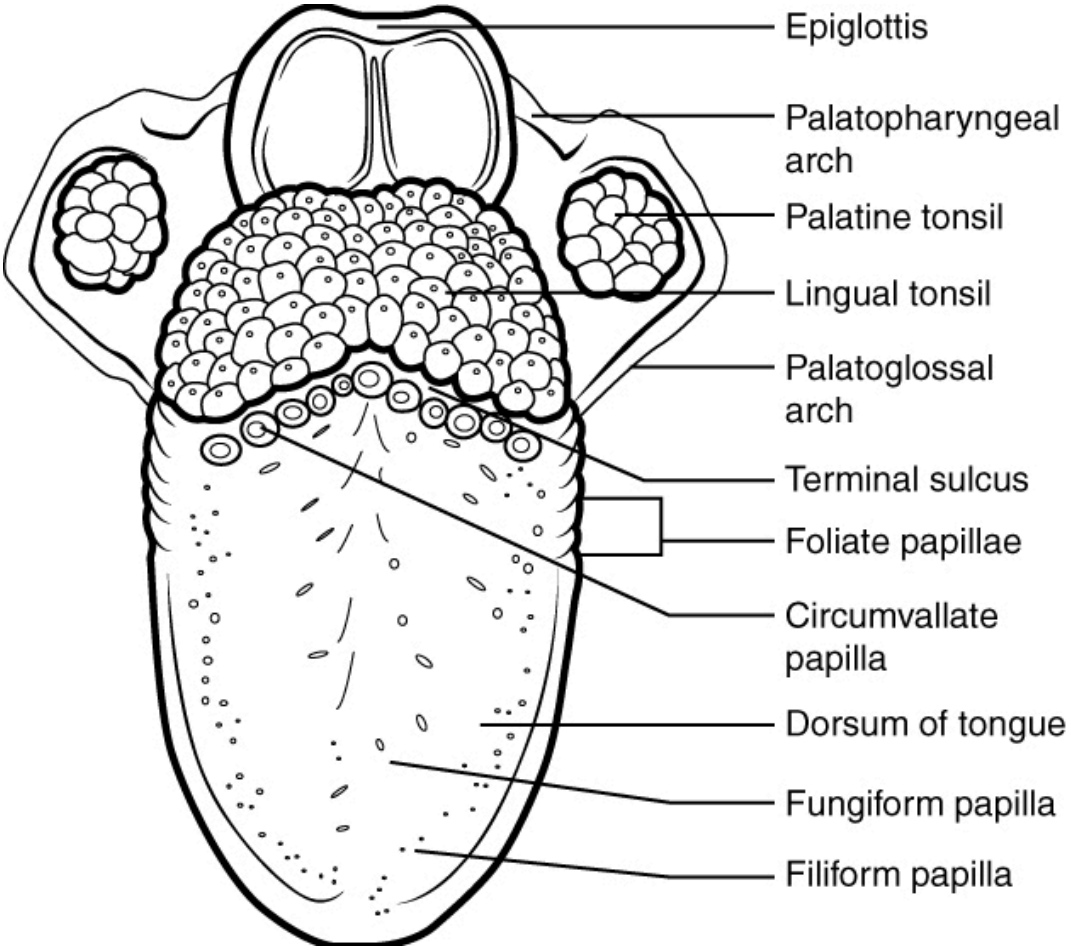


Figure 7. Tongue. *This superior view of the tongue shows the locations and types of lingual papillae.*

The Major Salivary Glands: Outside the oral mucosa are three pairs of major salivary glands, which secrete the majority of saliva into ducts that open into the mouth:

- The submandibular glands, which are in the floor of the mouth, secrete saliva into the mouth through the submandibular ducts.
- The sublingual glands, which lie below the tongue, use the lesser sublingual ducts to secrete saliva into the oral cavity.
- The parotid glands lie between the skin and the masseter muscle, near the ears. They secrete saliva into the mouth through the parotid duct, which is located near the second upper molar tooth (Figure 8).

Saliva: Saliva is essentially (>95%) water. The remainder is a complex mixture of ions, glycoproteins, enzymes, growth factors, and waste products. Perhaps the most important ingredient in salvia from the perspective of digestion is the enzyme **salivary amylase**, which initiates the breakdown of starch. Food does not spend enough time in the mouth to allow all the carbohydrates to break down, but salivary amylase continues acting until it is inactivated by stomach acids. Bicarbonate and phosphate ions function as chemical buffers, maintaining saliva at a pH between 6.35 and 6.85.

Salivary mucus helps lubricate food, facilitating movement in the mouth, bolus formation, and swallowing. Saliva contains immunoglobulin A, which prevents microbes from penetrating the epithelium, and lysozyme, which makes saliva antimicrobial.

Each of the major salivary glands secretes a unique formulation of saliva according to its cellular makeup. For example, the parotid glands secrete a watery solution that contains salivary amylase. The submandibular glands have cells similar to those of the parotid glands, as well as mucus-secreting cells. Therefore, saliva secreted by the submandibular glands also contains amylase but in a liquid thickened with mucus. The sublingual glands contain mostly mucous cells, and they secrete the thickest saliva with the least amount of salivary amylase.

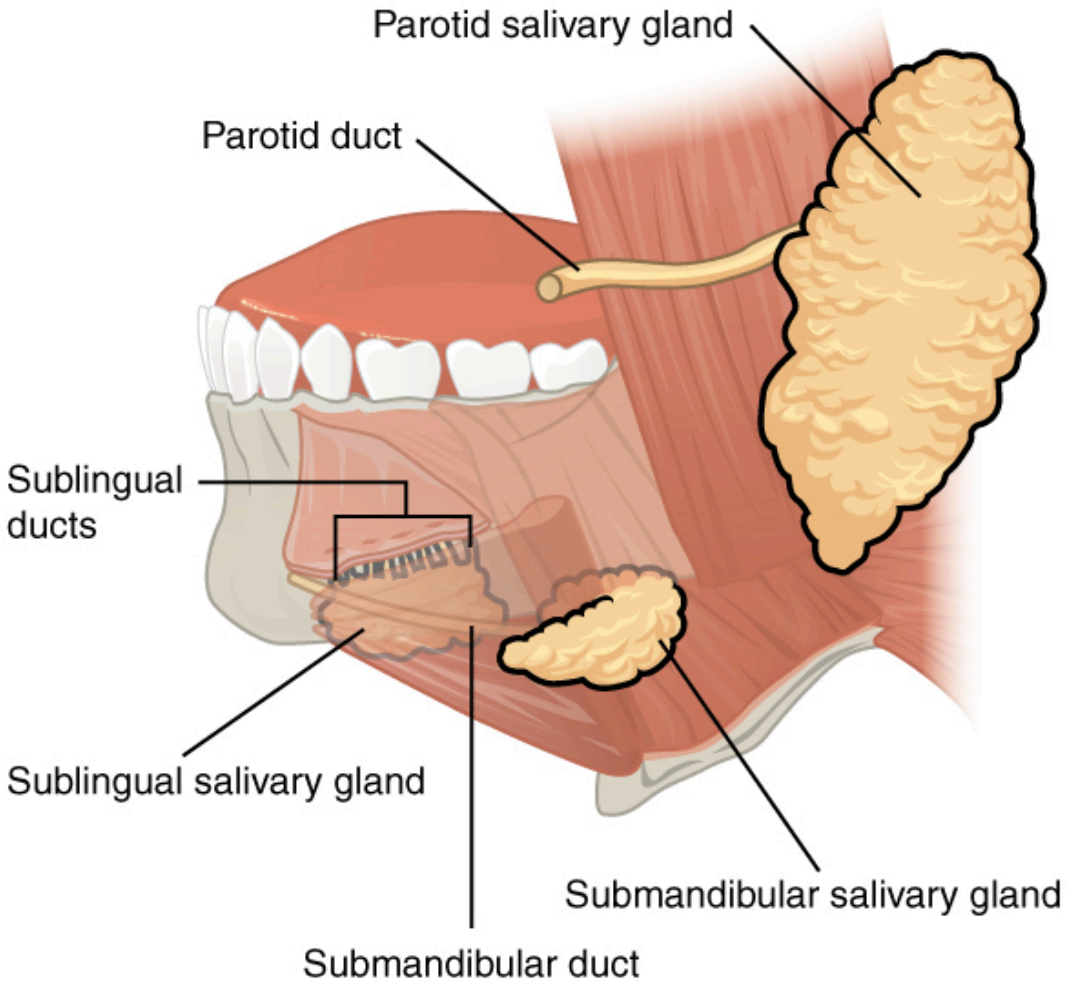


Figure 8. Salivary Glands. *The major salivary glands are located outside the oral mucosa and deliver saliva into the mouth through ducts.*

The Teeth: The teeth, or dentes (singular = dens), are organs similar to bones that you use to tear, grind, and otherwise mechanically break down food.

Types of Teeth: During the course of your lifetime, you have two sets of teeth (one set of teeth is a dentition). Your 20 deciduous teeth, or baby teeth, first begin to appear at about 6 months of age. Between approximately age 6 and 12, these teeth are replaced by 32 permanent teeth. Moving from the center of the mouth toward the side, these are as follows:

- The eight incisors, four top and four bottom, are the sharp front teeth you use for biting into food.
- The four cuspids (or canines) flank the incisors and have a pointed edge (cusp) to tear up food. These fang-like teeth are superb for piercing tough or fleshy foods.
- Posterior to the cuspids are the eight premolars (or bicuspids), which have an overall flatter shape with two rounded cusps useful for mashing foods.

The Pharynx: The pharynx (throat) is involved in both digestion and respiration. It receives food and air from the mouth, and air from the nasal cavities. When food enters the pharynx, involuntary muscle contractions close off the air passageways.

A short tube of skeletal muscle lined with a mucous membrane, the pharynx runs from the posterior oral and nasal cavities to the opening of the esophagus and larynx. It has three subdivisions. The most superior, the nasopharynx, is involved only in breathing and speech. The other two subdivisions, the oropharynx and the

laryngopharynx, are used for both breathing and digestion. The oropharynx begins inferior to the nasopharynx and is continuous below with the laryngopharynx (Figure 9). The inferior border of the laryngopharynx connects to the esophagus, whereas the anterior portion connects to the larynx, allowing air to flow into the bronchial tree. During swallowing, the elevator skeletal muscles of the pharynx contract, raising and expanding the pharynx to receive the bolus of food. Once received, these muscles relax and the constrictor muscles of the pharynx contract, forcing the bolus into the esophagus and initiating peristalsis.

Usually during swallowing, the soft palate and uvula rise reflexively to close off the entrance to the nasopharynx. At the same time, the larynx is pulled superiorly and the cartilaginous epiglottis, its most superior structure, folds inferiorly, covering the glottis (the opening to the larynx); this process effectively blocks access to the trachea and bronchi. When the food "goes down the wrong way," it goes into the trachea. When food enters the trachea, the reaction is to cough, which usually forces the food up and out of the trachea, and back into the pharynx.

The Esophagus: The esophagus is a muscular tube that connects the pharynx to the stomach. It is approximately 25.4 cm (10 in) in length, located posterior to the trachea, and remains in a collapsed form when not engaged in swallowing. The esophagus runs a mainly straight route through the mediastinum of the thorax (Figure 10). To enter the abdomen, the esophagus penetrates the diaphragm through an opening called the esophageal hiatus.

Table 3: Digestive Functions of the Mouth

Structure	Action	Outcome
Lips and cheeks	Confine food between teeth	Food is chewed evenly during mastication
		Moisten and lubricate lining of the mouth and pharynx
Salivary glands	Secrete saliva	Moisten, soften, dissolve food Clean mouth and teeth Salivary amylase breaks down starch
Tongue's extrinsic muscles	Move tongue sideways, and in and out	Manipulate food for chewing Shape food into a bolus Manipulate food for swallowing
Tongue's intrinsic muscles	Change tongue shape	Manipulate food for swallowing
Taste buds	Sense food in mouth, sense taste	Nerve impulses from taste buds are conducted to salivary nuclei in the brain stem and then to salivary glands, stimulating saliva secretion
Lingual glands	Secrete lingual lipase	Functions optimally in the stomach, breaks down triglycerides into fatty acids and diglycerides
Teeth	Shred and crush food	Break down solid food into smaller particles for deglutition

Passage of Food through the Esophagus: The upper esophageal sphincter, which is continuous with the inferior pharyngeal constrictor, controls the movement of food from the pharynx into the esophagus. The upper two-thirds of the esophagus consists of both smooth and skeletal muscle fibers, with the latter fading out in the bottom third of the esophagus. Rhythmic waves of peristalsis, which begin in the upper esophagus, propel the bolus of food toward the stomach. Meanwhile, secretions from the esophageal mucosa lubricate the esophagus and food. Food passes from the esophagus into the stomach at the lower esophageal sphincter (also called the gastroesophageal or cardiac sphincter). Recall that sphincters are muscles that surround tubes and serve as valves, closing the tube when the sphincters contract and opening it when they relax. The lower esophageal sphincter relaxes to let food pass into the stomach, and then contracts to prevent stomach acids from backing up into the esophagus. Surrounding this sphincter is the muscular diaphragm, which helps close off the sphincter when no food is being swallowed.

Histology of the Esophagus: The mucosa of the esophagus is made up of an epithelial lining that contains

non-keratinized, stratified squamous epithelium, with a layer of basal and parabasal cells. This epithelium protects against erosion from food particles. The mucosa's lamina propria contains mucus-secreting glands. The muscularis layer changes according to location: In the upper third of the esophagus, the muscularis is skeletal muscle. In the middle third, it is both skeletal and smooth muscle. In the lower third, it is smooth muscle. As mentioned previously, the most superficial layer of the esophagus is called the adventitia, not the serosa. In contrast to the stomach and intestines, the loose connective tissue of the adventitia is not covered by a fold of visceral peritoneum. The digestive functions of the esophagus are identified in Table 4.

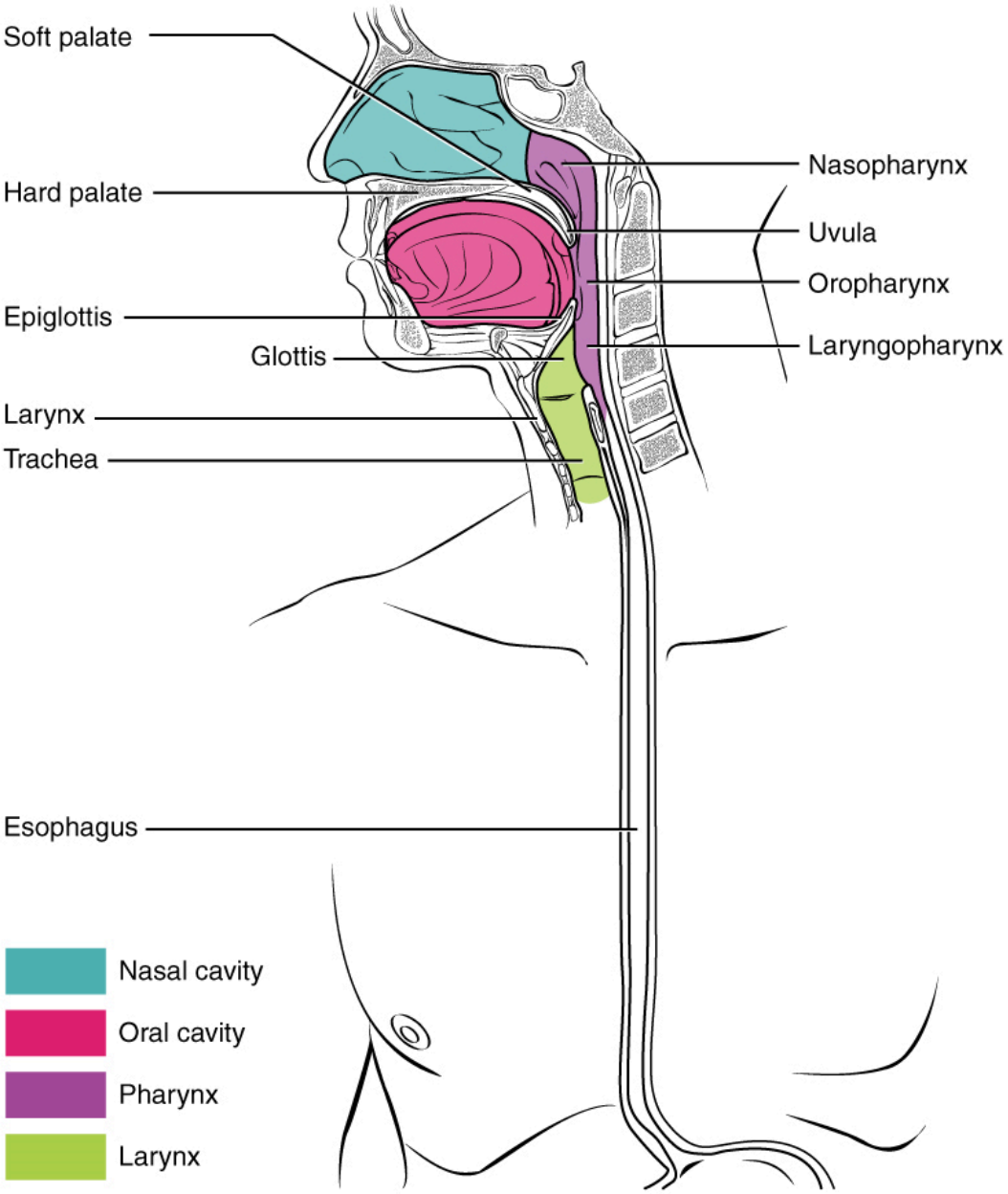


Figure 9. Pharynx. *The pharynx runs from the nostrils to the esophagus and the larynx.*

Table 4: Digestive Functions of the Esophagus

Action	Outcome
Upper esophageal sphincter relaxation	Allows bolus to move from laryngopharynx to esophagus
Peristalsis	Propels bolus through esophagus
Lower esophageal sphincter relaxation	Allows bolus to move from esophagus into stomach; prevents chyme from entering esophagus
Mucus secretion	Lubricates esophagus, allowing easy passage of bolus

Deglutition: Deglutition is another word for swallowing—the movement of food from the mouth to the stomach. The entire process takes about 4 to 8 seconds for solid or semisolid food, and about 1 second for very soft food and liquids. Although this sounds quick and effortless, deglutition is, in fact, a complex process that involves both the skeletal muscle of the tongue and the muscles of the pharynx and esophagus. It is aided by the presence of mucus and saliva. There are three stages in deglutition: the voluntary phase, the pharyngeal phase, and the esophageal phase (Figure 11). The autonomic nervous system controls the latter two phases.

The Voluntary Phase: The voluntary phase of deglutition (also known as the oral or buccal phase) is so called because you can control when you swallow food. In this phase, chewing has been completed and swallowing is set in motion. The tongue moves upward and backward against the palate, pushing the bolus to the back of the oral cavity and into the oropharynx. Other muscles keep the mouth closed and prevent food from falling out. At this point, the two involuntary phases of swallowing begin.

The Pharyngeal Phase: In the pharyngeal phase, stimulation of receptors in the oropharynx sends impulses to the deglutition center (a collection of neurons that controls swallowing) in the medulla oblongata. Impulses are then sent back to the uvula and soft palate, causing them to move upward and close off the nasopharynx. The laryngeal muscles also constrict to prevent aspiration of food into the trachea. At this point, deglutition apnea takes place, which means that breathing ceases for a very brief time. Contractions of the pharyngeal constrictor muscles move the bolus through the oropharynx and laryngopharynx. Relaxation of the upper esophageal sphincter then allows food to enter the esophagus.

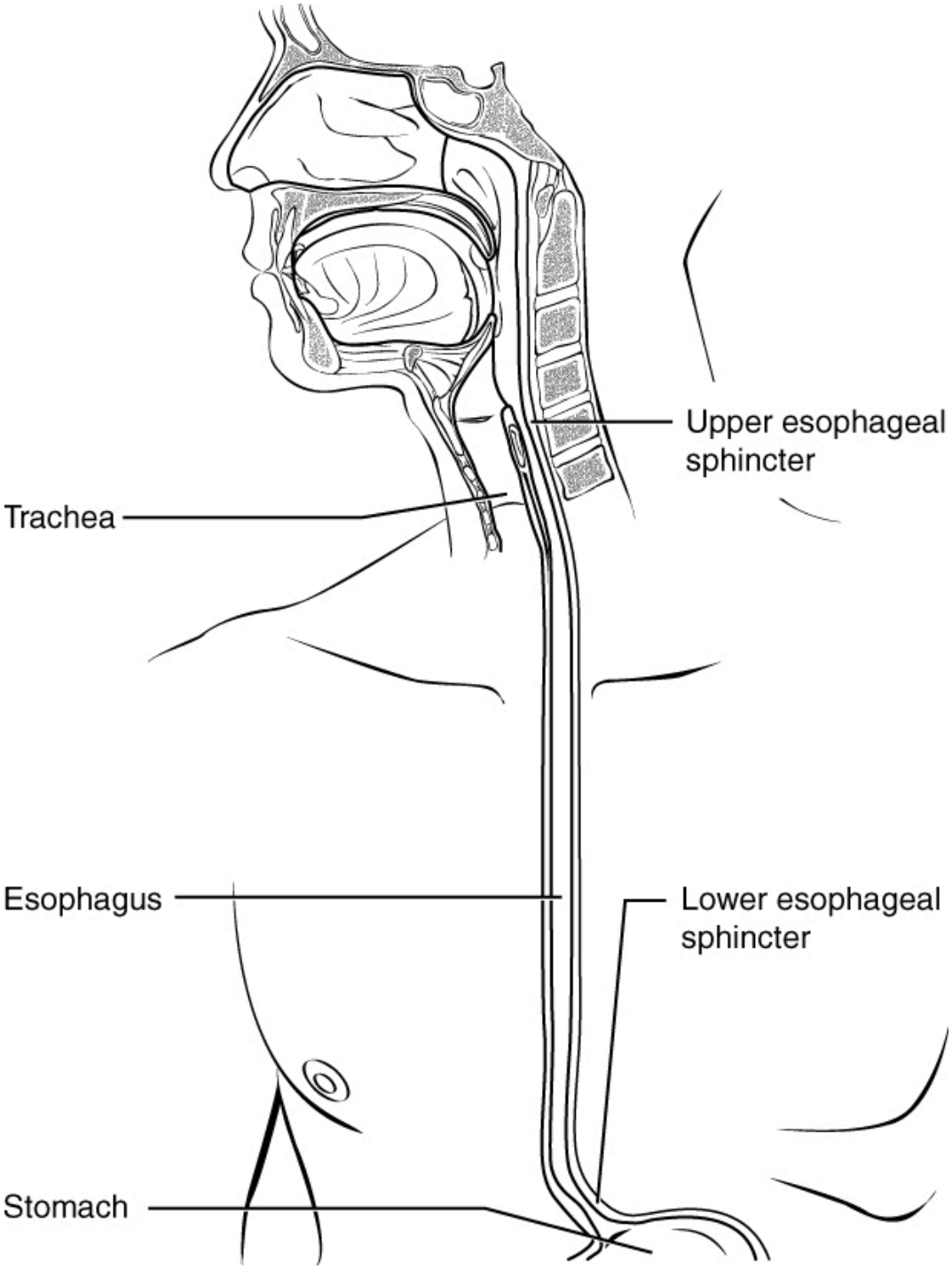


Figure 10. Deglutition. *Deglutition includes the voluntary phase and two involuntary phases: the pharyngeal phase and the esophageal phase.*

The Esophageal Phase: The entry of food into the esophagus marks the beginning of the esophageal phase of deglutition and the initiation of peristalsis. As in the previous phase, the complex neuromuscular actions are controlled by the medulla oblongata. Peristalsis propels the bolus through the esophagus and toward the stomach. The circular muscle layer of the muscularis contracts, pinching the esophageal wall and forcing the bolus forward. At the same time, the longitudinal muscle layer of the muscularis also contracts, shortening this area and pushing out its walls to receive the bolus. In this way, a series of contractions keeps moving food toward the stomach. When the bolus nears the stomach, distention of the esophagus initiates a short

reflex relaxation of the lower esophageal sphincter that allows the bolus to pass into the stomach. During the esophageal phase, esophageal glands secrete mucus that lubricates the bolus and minimizes friction.

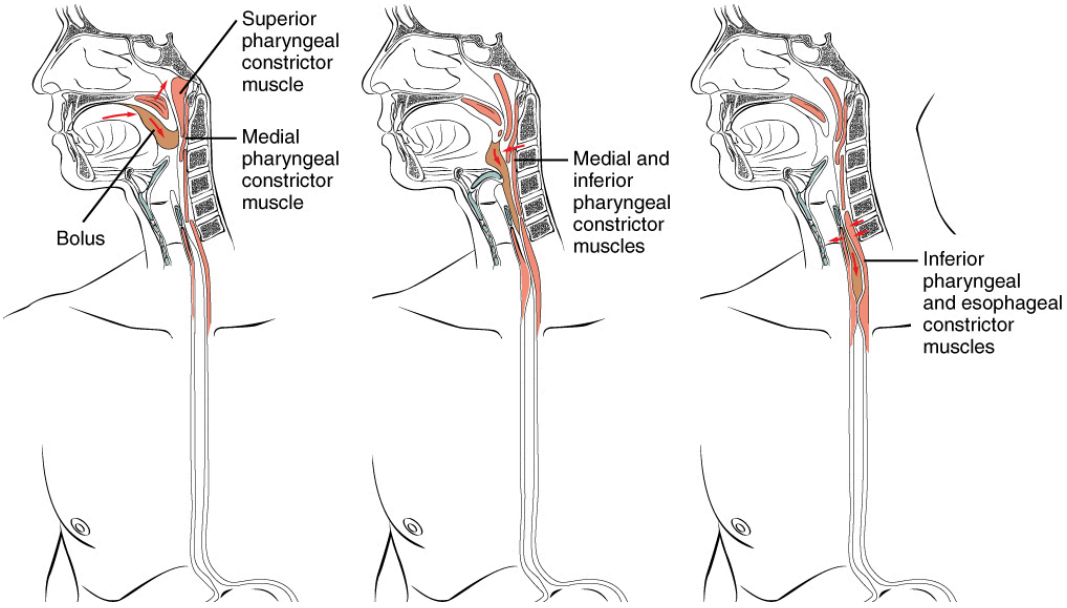


Figure 11. Deglutition. *Deglutition includes the voluntary phase and two involuntary phases: the pharyngeal phase and the esophageal phase.*

Watch this animation to see how swallowing is a complex process that involves the nervous system to coordinate the actions of upper respiratory and digestive activities. Direct link: http://openstaxcollege.org/l/swallowing

Part 4: The Stomach

Although a minimal amount of carbohydrate digestion occurs in the mouth, chemical digestion really gets underway in the stomach. An expansion of the alimentary canal that lies immediately inferior to the esophagus, the stomach links the esophagus to the first part of the small intestine (the duodenum) and is relatively fixed in place at its esophageal and duodenal ends. In between, however, it can be a highly active structure, contracting and continually changing position and size. These contractions provide mechanical assistance to digestion. The empty stomach is only about the size of your fist, but can stretch to hold as much as 4 liters of food and fluid, or more than 75 times its empty volume, and then return to its resting size when empty. Although you might think that the size of a person's stomach is related to how much food that individual consumes, body weight does not correlate with stomach size. Rather, when you eat greater quantities of food—such as at holiday dinner—you stretch the stomach more than when you eat less.

An important function of the stomach is to serve as a temporary holding chamber. You can ingest a meal far more quickly than it can be digested and absorbed by the small intestine. Thus, the stomach holds food and parses only small amounts into the small intestine at a time. Foods are not processed in the order they are eaten; rather, they are mixed together with digestive juices in the stomach until they are converted into chyme, which is released into the small intestine.

As you will see in the sections that follow, the stomach plays several important roles in chemical digestion, including the continued digestion of carbohydrates and the initial digestion of proteins and triglycerides. Little if any nutrient absorption occurs in the stomach, with the exception of the negligible amount of nutrients in alcohol.

Structure: There are four main regions in the **stomach**: the cardia, fundus, body, and pylorus (Figure 12). The **cardia** (or cardiac region) is the point where the esophagus connects to the stomach and through which food passes into the stomach. Located inferior to the diaphragm, above and to the left of the cardia, is the dome-shaped **fundus**. Below the fundus is the **body**, the main part of the stomach. The funnel-shaped **pylorus** connects the stomach to the duodenum. The wider end of the funnel, the **pyloric antrum**, connects to the body of the stomach. The narrower end is called the **pyloric canal**, which connects to the duodenum. The smooth muscle **pyloric sphincter** is located at this latter point of connection and controls stomach emptying. In the absence of food, the stomach deflates inward, and its mucosa and submucosa fall into a large fold called a **ruga**.

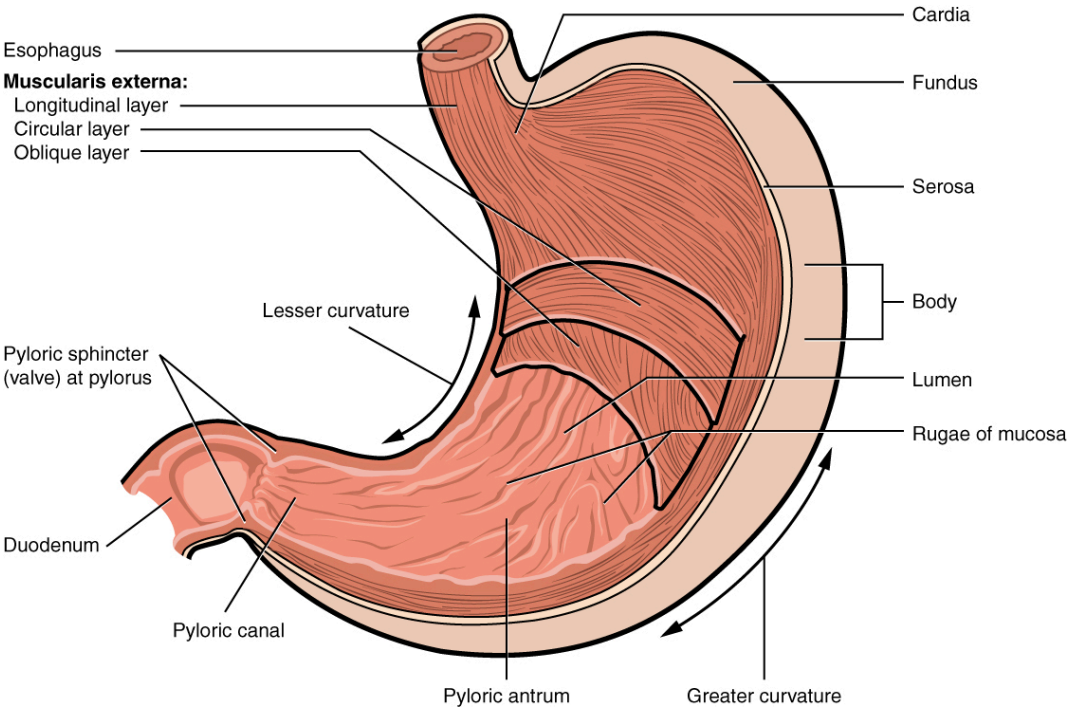


Figure 12. Stomach. *The stomach has four major regions: the cardia, fundus, body, and pylorus. The addition of an inner oblique smooth muscle layer gives the muscularis the ability to vigorously churn and mix food.*

Histology: The wall of the stomach is made of the same four layers as most of the rest of the alimentary canal, but with adaptations to the mucosa and muscularis for the unique functions of this organ. In addition to the typical circular and longitudinal smooth muscle layers, the muscularis has an inner oblique smooth muscle layer (Figure 13). As a result, in addition to moving food through the canal, the stomach can vigorously churn food, mechanically breaking it down into smaller particles.

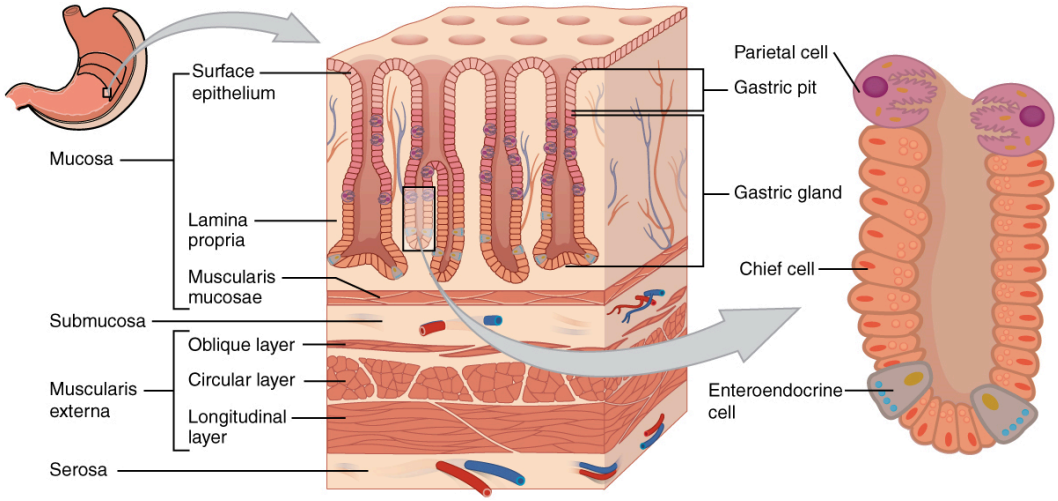


Figure 13. Histology of the Stomach. *The stomach wall is adapted for the functions of the stomach. In the epithelium, gastric pits lead to gastric glands that secrete gastric juice. The gastric glands (one gland is shown enlarged on the right) contain different types of cells that secrete a variety of enzymes, including hydrochloride acid, which activates the protein-digesting enzyme pepsin.*

The stomach mucosa's epithelial lining consists only of surface mucus cells, which secrete a protective coat of alkaline mucus. A vast number of gastric pits dot the surface of the epithelium, giving it the appearance of a well-used pincushion, and mark the entry to each gastric gland, which secretes a complex digestive fluid referred to as gastric juice.

Although the walls of the gastric pits are made up primarily of mucus cells, the gastric glands are made up of different types of cells. The glands of the cardia and pylorus are composed primarily of mucus-secreting cells. Cells that make up the pyloric antrum secrete mucus and a number of hormones, including the majority of the stimulatory hormone, **gastrin**. The much larger glands of the fundus and body of the stomach, the site of most chemical digestion, produce most of the gastric secretions. These glands are made up of a variety of secretory cells. These include parietal cells, chief cells, mucous neck cells, and enteroendocrine cells.

- *Parietal cells*—Located primarily in the middle region of the gastric glands are **parietal cells,** which are among the most highly differentiated of the body's epithelial cells. These relatively large cells produce both **hydrochloric acid** (HCl) and **intrinsic factor**. HCl is responsible for the high acidity (pH 1.5 to 3.5) of the stomach contents and is needed to activate the protein-digesting enzyme, pepsin. The acidity also kills much of the bacteria you ingest with food and helps to denature proteins, making them more available for enzymatic digestion. Intrinsic factor is a glycoprotein necessary for the absorption of vitamin B12 in the small intestine.
- *Chief cells*—Located primarily in the basal regions of gastric glands are **chief cells**, which secrete **pepsinogen**, the inactive proenzyme form of pepsin. HCl is necessary for the conversion of pepsinogen to pepsin.
- *Mucous neck cells*—Gastric glands in the upper part of the stomach contain **mucous neck cells** that secrete thin, acidic mucus that is much different from the mucus secreted by the goblet cells of the surface epithelium. The role of this mucus is not currently known.
- *Enteroendocrine cells*—Finally, **enteroendocrine cells** found in the gastric glands secrete various hormones into the interstitial fluid of the lamina propria. These include gastrin, which is released mainly by enteroendocrine **G cells**.

Watch this animation that depicts the structure of the stomach and how this structure functions in the initiation of protein digestion. Direct link: http://openstaxcollege.org/l/stomach1

Gastric Secretion: The secretion of gastric juice is controlled by both nerves and hormones. Stimuli in the brain, stomach, and small intestine activate or inhibit gastric juice production. This is why the three phases of gastric secretion are called the cephalic, gastric, and intestinal phases (Figure 14). However, once gastric secretion begins, all three phases can occur simultaneously.

The **cephalic phase** (reflex phase) of gastric secretion, which is relatively brief, takes place before food enters the stomach. The smell, taste, sight, or thought of food triggers this phase. For example, when you bring a piece of sushi to your lips, impulses from receptors in your taste buds or the nose are relayed to your brain, which returns signals that increase gastric secretion to prepare your stomach for digestion. This enhanced secretion is a conditioned reflex, meaning it occurs only if you like or want a particular food. Depression and loss of appetite can suppress the cephalic reflex.

The **gastric phase** of secretion lasts 3 to 4 hours, and is set in motion by local neural and hormonal mechanisms triggered by the entry of food into the stomach. For example, when your sushi reaches the stomach, it creates distention that activates the stretch receptors. This stimulates parasympathetic neurons to release acetylcholine, which then provokes increased secretion of gastric juice. Partially digested proteins, caffeine, and rising pH stimulate the release of gastrin from enteroendocrine G cells, which in turn induces parietal cells to increase their production of HCl, which is needed to create an acidic environment for the conversion of pepsinogen to pepsin, and protein digestion. Additionally, the release of gastrin activates vigorous smooth muscle contractions. However, it should be noted that the stomach does have a natural means of avoiding excessive acid secretion and potential heartburn. Whenever pH levels drop too low, cells in the stomach react by suspending HCl secretion and increasing mucous secretions.

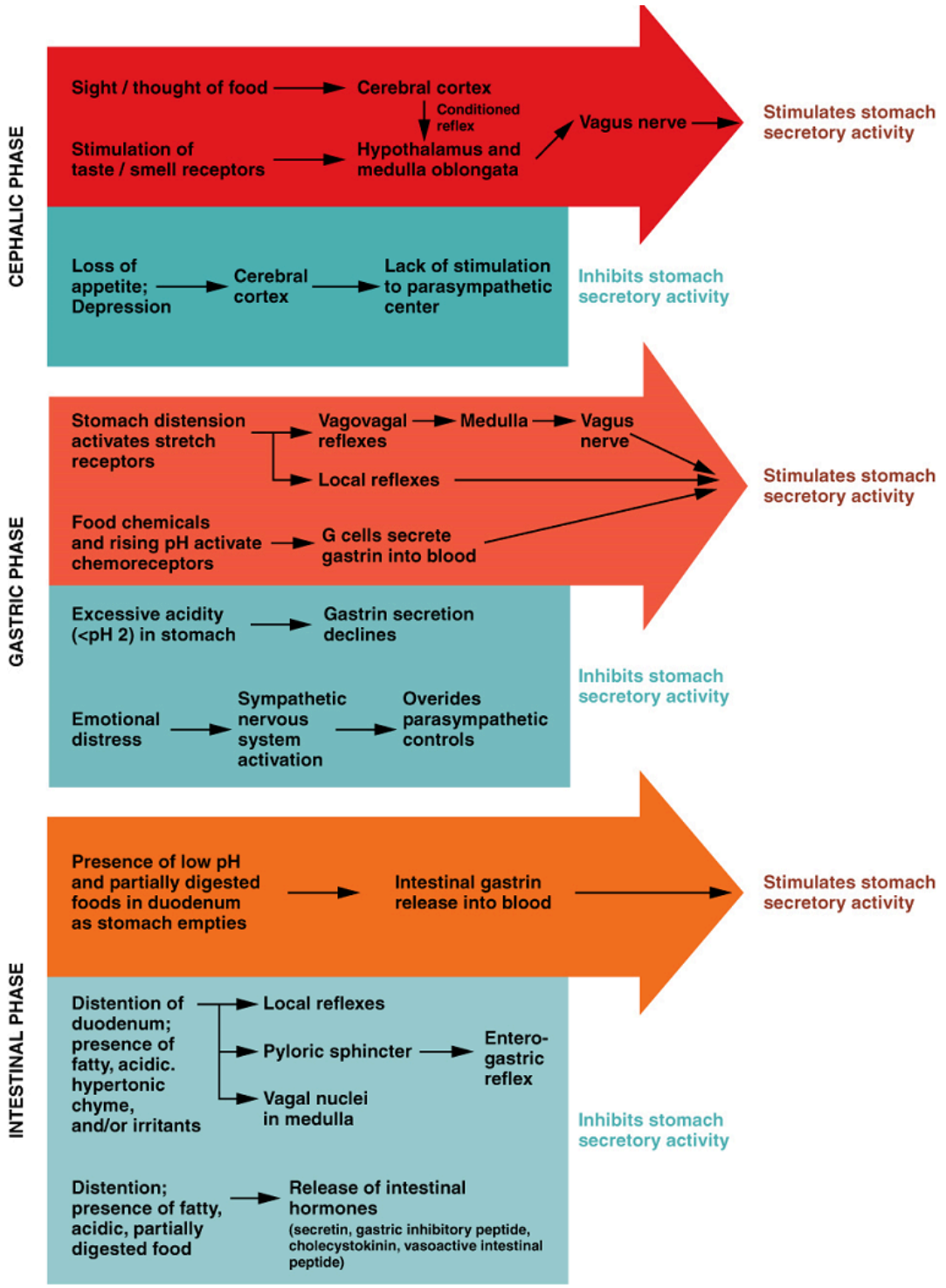


Figure 14. The Three Phases of Gastric Secretion. *Gastric secretion occurs in three phases: cephalic, gastric, and intestinal. During each phase, the secretion of gastric juice can be stimulated or inhibited.*

The **intestinal phase** of gastric secretion has both excitatory and inhibitory elements. The duodenum has a major role in regulating the stomach and its emptying. When partially digested food fills the duodenum, intestinal mucosal cells release a hormone called intestinal (enteric) gastrin, which further excites gastric juice secretion. This stimulatory activity is brief, however, because when the intestine distends with chyme, the enterogastric reflex inhibits secretion. One of the effects of this reflex is to close the pyloric sphincter, which blocks additional chyme from entering the duodenum.

The Mucosal Barrier: The mucosa of the stomach is exposed to the highly corrosive acidity of gastric juice. Gastric enzymes that can digest protein can also digest the stomach itself. The stomach is protected from self-digestion by the mucosal barrier. This barrier has several components. First, the stomach wall is covered by a thick coating of bicarbonate-rich mucus. This mucus forms a physical barrier, and its bicarbonate ions neutralize acid. Second, the epithelial cells of the stomach's mucosa meet at tight junctions, which block gastric juice from penetrating the underlying tissue layers. Finally, stem cells located where gastric glands join the gastric pits quickly replace damaged epithelial mucosal cells, when the epithelial cells are shed. In fact, the surface epithelium of the stomach is completely replaced every 3 to 6 days.

Digestive Functions of the Stomach: The stomach participates in virtually all the digestive activities with the exception of ingestion and defecation. Although almost all absorption takes place in the small intestine, the stomach does absorb some nonpolar substances, such as alcohol and aspirin.

Mechanical Digestion: Within a few moments after food after enters your stomach, mixing waves begin to occur at intervals of approximately 20 seconds. A **mixing wave** is a unique type of peristalsis that mixes and softens the food with gastric juices to create chyme. The initial mixing waves are relatively gentle, but these are followed by more intense waves, starting at the body of the stomach and increasing in force as they reach the pylorus. It is fair to say that long before your sushi exits through the pyloric sphincter, it bears little resemblance to the sushi you ate.

The pylorus, which holds around 30 mL (1 fluid ounce) of chyme, acts as a filter, permitting only liquids and small food particles to pass through the mostly, but not fully, closed pyloric sphincter. In a process called **gastric emptying**, rhythmic mixing waves force about 3 mL of chyme at a time through the pyloric sphincter and into the duodenum. Release of a greater amount of chyme at one time would overwhelm the capacity of the small intestine to handle it. The rest of the chyme is pushed back into the body of the stomach, where it continues mixing. This process is repeated when the next mixing waves force more chyme into the duodenum.

Gastric emptying is regulated by both the stomach and the duodenum. The presence of chyme in the duodenum activates receptors that inhibit gastric secretion. This prevents additional chyme from being released by the stomach before the duodenum is ready to process it.

Chemical Digestion: The fundus plays an important role, because it stores both undigested food and gases that are released during the process of chemical digestion. Food may sit in the fundus of the stomach for a while before being mixed with the chyme. While the food is in the fundus, the digestive activities of salivary amylase continue until the food begins mixing with the acidic chyme. Ultimately, mixing waves incorporate this food with the chyme, the acidity of which inactivates salivary amylase. The acidity of the chyme also allows lingual lipase to break down triglycerides into free fatty acids and diglycerides more efficiently than it could in the less acidic environment of the mouth.

The breakdown of protein begins in the stomach through the actions of HCl and the enzyme pepsin. During infancy, gastric glands also produce rennin, an enzyme that helps digest milk protein.

Its numerous digestive functions notwithstanding, there is only one stomach function necessary to life: the production of intrinsic factor. The intestinal absorption of vitamin B_{12}, which is necessary for both the production of mature red blood cells and normal neurological functioning, cannot occur without intrinsic factor. People who undergo total gastrectomy (stomach removal)—for life-threatening stomach cancer, for example—can survive with minimal digestive dysfunction if they receive vitamin B_{12} injections.

The contents of the stomach are completely emptied into the duodenum within 2 to 4 hours after you eat a meal. Different types of food take different amounts of time to process. Foods heavy in carbohydrates empty fastest, followed by high-protein foods. Meals with a high triglyceride content remain in the stomach the longest. Since enzymes in the small intestine digest fats slowly, food can stay in the stomach for 6 hours or longer when the duodenum is processing fatty chyme. However, note that this is still a fraction of the 24 to 72 hours that full digestion typically takes from start to finish.

Part 5: The Small and Large Intestines

The word intestine is derived from a Latin root meaning "internal," and indeed, the two organs together nearly

fill the interior of the abdominal cavity. In addition, called the small and large bowel, or colloquially the "guts," they constitute the greatest mass and length of the alimentary canal and, with the exception of ingestion, perform all digestive system functions.

The Small Intestine: Chyme released from the stomach enters the **small intestine**, which is the primary digestive organ in the body. Not only is this where most digestion occurs, it is also where practically all absorption occurs. The longest part of the alimentary canal, the small intestine is about 3.05 meters (10 feet) long in a living person (but about twice as long in a cadaver due to the loss of muscle tone). Since this makes it about five times longer than the large intestine, you might wonder why it is called "small." In fact, its name derives from its relatively smaller diameter of only about 2.54 cm (1 in), compared with 7.62 cm (3 in) for the large intestine. As we'll see shortly, in addition to its length, the folds and projections of the lining of the small intestine work to give it an enormous surface area, which is approximately 200 m^2, more than 100 times the surface area of your skin. This large surface area is necessary for complex processes of digestion and absorption that occur within it.

Structure: The coiled tube of the small intestine is subdivided into three regions. From proximal (at the stomach) to distal, these are the duodenum, jejunum, and ileum (Figure 15).

The shortest region is the 25.4-cm (10-in) **duodenum**, which begins at the pyloric sphincter. Just past the pyloric sphincter, it bends posteriorly behind the peritoneum, becoming retroperitoneal, and then makes a C-shaped curve around the head of the pancreas before ascending anteriorly again to return to the peritoneal cavity and join the jejunum. The duodenum can therefore be subdivided into four segments: the superior, descending, horizontal, and ascending duodenum.

Of particular interest is the **hepatopancreatic ampulla** (ampulla of Vater). Located in the duodenal wall, the ampulla marks the transition from the anterior portion of the alimentary canal to the mid-region, and is where the bile duct (through which bile passes from the liver) and the **main pancreatic duct** (through which pancreatic juice passes from the pancreas) join. This ampulla opens into the duodenum at a tiny volcano-shaped structure called the **major duodenal papilla**. The **hepatopancreatic sphincter** (sphincter of Oddi) regulates the flow of both bile and pancreatic juice from the ampulla into the duodenum.

The **jejunum** is about 0.9 meters (3 feet) long (in life) and runs from the duodenum to the ileum. Jejunum means "empty" in Latin and supposedly was so named by the ancient Greeks who noticed it was always empty at death.

No clear demarcation exists between the jejunum and the final segment of the small intestine, the ileum.

The **ileum** is the longest part of the small intestine, measuring about 1.8 meters (6 feet) in length. It is thicker, more vascular, and has more developed mucosal folds than the jejunum. The ileum joins the cecum, the first portion of the large intestine, at the **ileocecal sphincter** (or valve). The jejunum and ileum are tethered to the posterior abdominal wall by the mesentery. The large intestine frames these three parts of the small intestine.

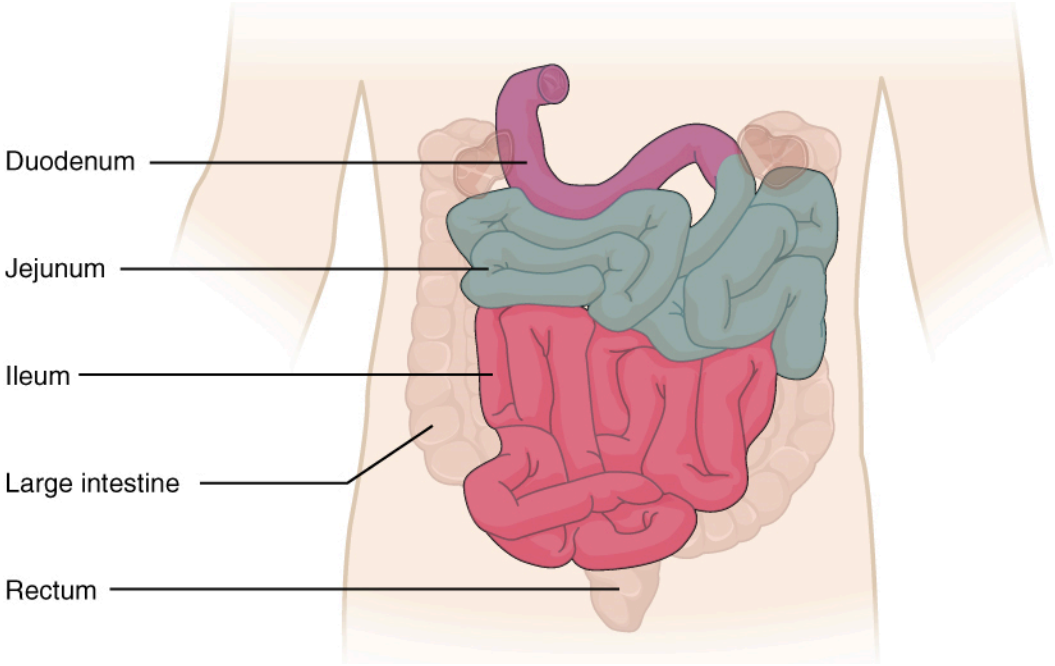


Figure 15. Small Intestine. *The three regions of the small intestine are the duodenum, jejunum, and ileum.*

Parasympathetic nerve fibers from the vagus nerve and sympathetic nerve fibers from the thoracic splanchnic nerve provide extrinsic innervation to the small intestine. The superior mesenteric artery is its main arterial supply. Veins run parallel to the arteries and drain into the superior mesenteric vein. Nutrient-rich blood from the small intestine is then carried to the liver via the hepatic portal vein.

Histology: The wall of the small intestine is composed of the same four layers typically present in the alimentary system. However, three features of the mucosa and submucosa are unique. These features, which increase the absorptive surface area of the small intestine more than 600-fold, include circular folds, villi, and microvilli (Figure 16). These adaptations are most abundant in the proximal two-thirds of the small intestine, where the majority of absorption occurs.

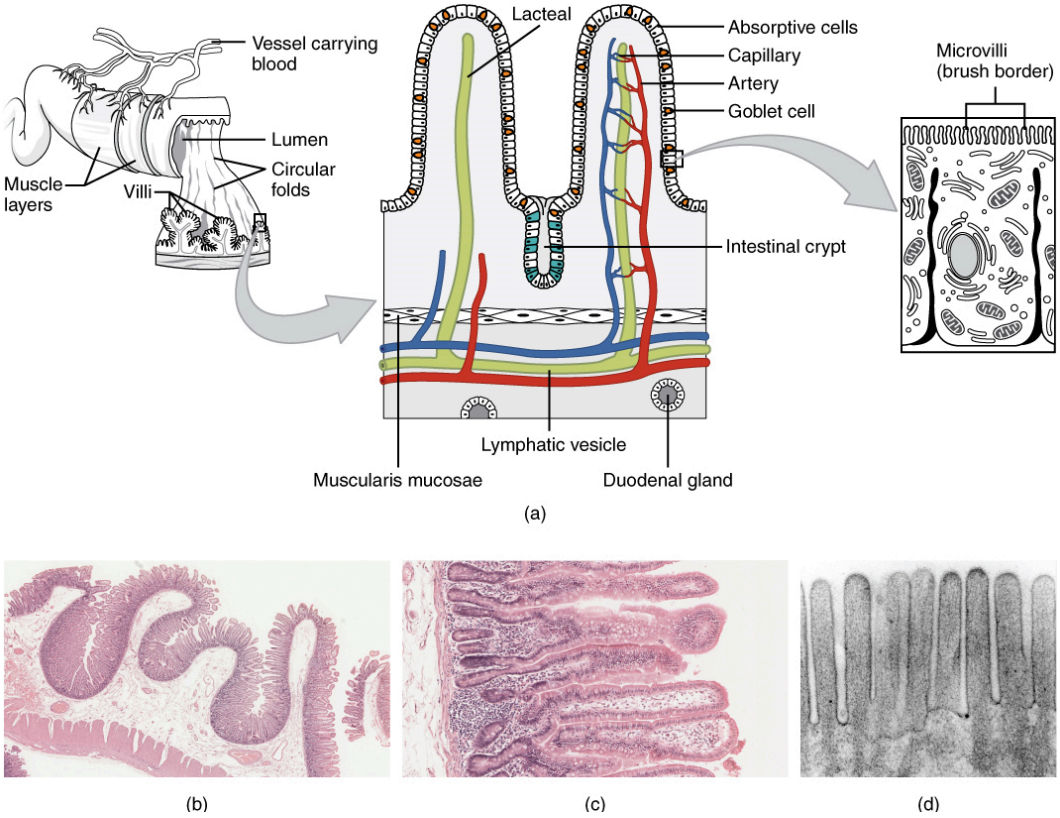


Figure 16. Histology of the Small Intestine. *(a) The absorptive surface of the small intestine is vastly enlarged by the presence of circular folds, villi, and microvilli. (b) Micrograph of the circular folds. (c) Micrograph of the villi. (d) Electron micrograph of the microvilli. From left to right, LM x 56, LM x 508, EM x 196,000. (credit b-d: Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

Circular folds: Also called a plica circulare, a **circular fold** is a deep ridge in the mucosa and submucosa. Beginning near the proximal part of the duodenum and ending near the middle of the ileum, these folds facilitate absorption. Their shape causes the chyme to spiral, rather than move in a straight line, through the small intestine. Spiraling slows the movement of chyme and provides the time needed for nutrients to be fully absorbed.

Villi: Within the circular folds are small (0.5–1 mm long) hairlike vascularized projections called **villi** (singular = villus) that give the mucosa a furry texture. There are about 20 to 40 villi per square millimeter, increasing the surface area of the epithelium tremendously. The mucosal epithelium, primarily composed of absorptive cells, covers the villi. In addition to muscle and connective tissue to support its structure, each villus contains a capillary bed composed of one arteriole and one venule, as well as a lymphatic capillary called a **lacteal**. The breakdown products of carbohydrates and proteins (sugars and amino acids) can enter the bloodstream directly, but lipid breakdown products are absorbed by the lacteals and transported to the bloodstream via the lymphatic system.

Microvilli: As their name suggests, microvilli (singular = microvillus) are much smaller (1 μm) than villi. They are cylindrical apical surface extensions of the plasma membrane of the mucosa's epithelial cells, and are supported by microfilaments within those cells. Although their small size makes it difficult to see each microvillus, their combined microscopic appearance suggests a mass of bristles, which is termed the brush border. Fixed to the surface of the microvilli membranes are enzymes that finish digesting carbohydrates and proteins. There are an estimated 200 million microvilli per square millimeter of small intestine, greatly expanding the surface area of the plasma membrane and thus greatly enhancing absorption.

Intestinal Glands: In addition to the three specialized absorptive features just discussed, the mucosa between the villi is dotted with deep crevices that each lead into a tubular **intestinal gland** (crypt of Lieberkühn), which is formed by cells that line the crevices (see Figure 16). These produce **intestinal juice**, a slightly alkaline (pH 7.4

to 7.8) mixture of water and mucus. Each day, about 0.95 to 1.9 liters (1 to 2 quarts) are secreted in response to the distention of the small intestine or the irritating effects of chyme on the intestinal mucosa.

The submucosa of the duodenum is the only site of the complex mucus-secreting **duodenal glands** (Brunner's glands), which produce a bicarbonate-rich alkaline mucus that buffers the acidic chyme as it enters from the stomach.

Mechanical Digestion in the Small Intestine: The movement of intestinal smooth muscles includes both segmentation and a form of peristalsis called migrating motility complexes. The kind of peristaltic mixing waves seen in the stomach are not observed here.

If you could see into the small intestine when it was going through segmentation, it would look as if the contents were being shoved incrementally back and forth, as the rings of smooth muscle repeatedly contract and then relax. Segmentation in the small intestine does not force chyme through the tract. Instead, it combines the chyme with digestive juices and pushes food particles against the mucosa to be absorbed. The duodenum is where the most rapid segmentation occurs, at a rate of about 12 times per minute. In the ileum, segmentations are only about eight times per minute (Figure 17).

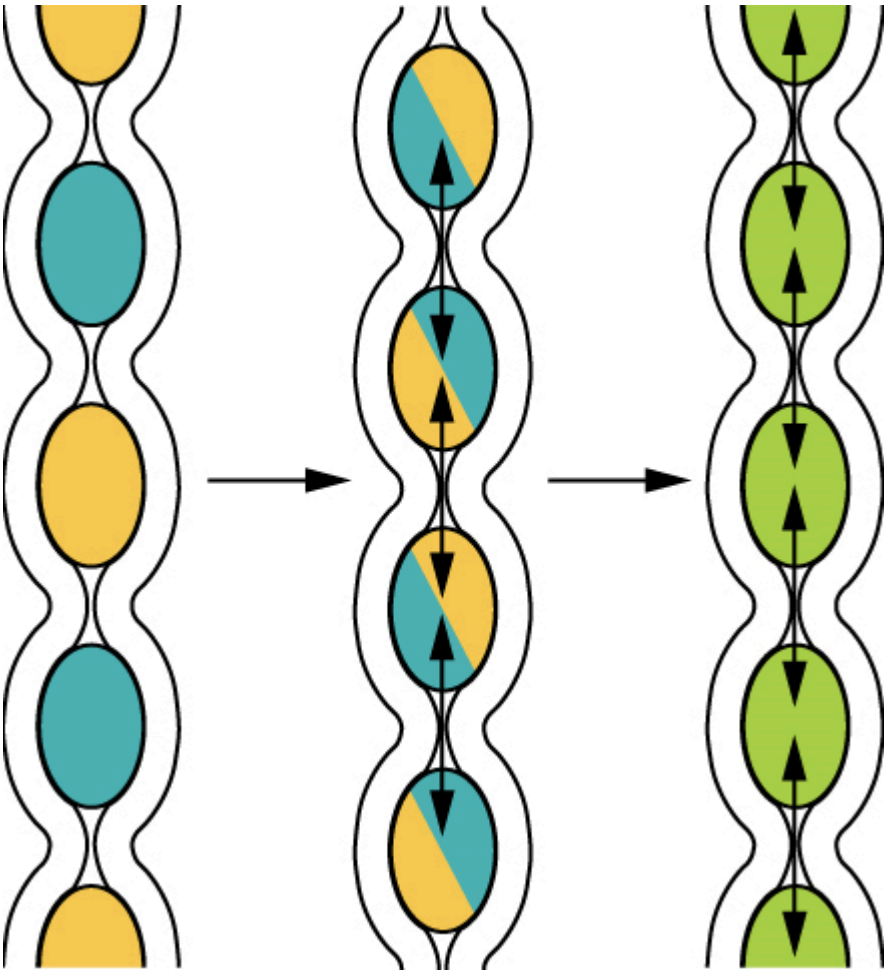


Figure 17. Segmentation. *Segmentation separates chyme and then pushes it back together, mixing it and providing time for digestion and absorption.*

When most of the chyme has been absorbed, the small intestinal wall becomes less distended. At this point, the localized segmentation process is replaced by transport movements. The duodenal mucosa secretes the hormone **motilin**, which initiates peristalsis. These complexes, which begin in the duodenum, force chyme through a short section of the small intestine and then stop. The next contraction begins a little bit farther down

than the first, forces chyme a bit farther through the small intestine, then stops. These complexes move slowly down the small intestine, forcing chyme on the way, taking around 90 to 120 minutes to finally reach the end of the ileum. At this point, the process is repeated, starting in the duodenum.

The ileocecal valve, a sphincter, is usually in a constricted state, but when motility in the ileum increases, this sphincter relaxes, allowing food residue to enter the first portion of the large intestine, the cecum. Relaxation of the ileocecal sphincter is controlled by both nerves and hormones. First, digestive activity in the stomach provokes the **gastroileal reflex**, which increases the force of ileal segmentation. Second, the stomach releases the hormone gastrin, which enhances ileal motility, thus relaxing the ileocecal sphincter. After chyme passes through, backward pressure helps close the sphincter, preventing backflow into the ileum. Because of this reflex, your lunch is completely emptied from your stomach and small intestine by the time you eat your dinner. It takes about 3 to 5 hours for all chyme to leave the small intestine.

Chemical Digestion in the Small Intestine: The digestion of proteins and carbohydrates, which partially occurs in the stomach, is completed in the small intestine with the aid of intestinal and pancreatic juices. Lipids arrive in the intestine largely undigested, so much of the focus here is on lipid digestion, which is facilitated by bile and the enzyme pancreatic lipase.

Moreover, intestinal juice combines with pancreatic juice to provide a liquid medium that facilitates absorption. The intestine is also where most water is absorbed, via osmosis. The small intestine's absorptive cells also synthesize digestive enzymes and then place them in the plasma membranes of the microvilli. This distinguishes the small intestine from the stomach; that is, enzymatic digestion occurs not only in the lumen, but also on the luminal surfaces of the mucosal cells.

For optimal chemical digestion, chyme must be delivered from the stomach slowly and in small amounts. This is because chyme from the stomach is typically hypertonic, and if large quantities were forced all at once into the small intestine, the resulting osmotic water loss from the blood into the intestinal lumen would result in potentially life-threatening low blood volume. In addition, continued digestion requires an upward adjustment of the low pH of stomach chyme, along with rigorous mixing of the chyme with bile and pancreatic juices. Both processes take time, so the pumping action of the pylorus must be carefully controlled to prevent the duodenum from being overwhelmed with chyme.

The Large Intestine: The **large intestine is** the terminal part of the alimentary canal. The primary function of this organ is to finish absorption of nutrients and water, synthesize certain vitamins, form feces, and eliminate feces from the body.

Structure: The large intestine runs from the appendix to the anus. It frames the small intestine on three sides. Despite its being about one-half as long as the small intestine, it is called large because it is more than twice the diameter of the small intestine, about 3 inches.

Subdivisions: The large intestine is subdivided into four main regions: the cecum, the colon, the rectum, and the anus. The ileocecal valve, located at the opening between the ileum and the large intestine, controls the flow of chyme from the small intestine to the large intestine.

1. Cecum: The first part of the large intestine is the cecum, a sac-like structure that is suspended inferior to the ileocecal valve. It is about 6 cm (2.4 in) long, receives the contents of the ileum, and continues the absorption of water and salts. The appendix (or vermiform appendix) is a winding tube that attaches to the cecum. Although the 7.6-cm (3-in) long appendix contains lymphoid tissue, suggesting an immunologic function, this organ is generally considered vestigial. However, at least one recent report postulates a survival advantage conferred by the appendix: In diarrheal illness, the appendix may serve as a bacterial reservoir to repopulate the enteric bacteria for those surviving the initial phases of the illness. Moreover, its twisted anatomy provides a haven for the accumulation and multiplication of enteric bacteria. The mesoappendix, the mesentery of the appendix, tethers it to the mesentery of the ileum.

2. Colon: The cecum blends seamlessly with the colon. Upon entering the colon, the food residue first travels up the ascending colon on the right side of the abdomen. At the inferior surface of the liver, the colon bends to the **right colic flexure** (hepatic flexure) and becomes the **transverse colon**. The region defined as hindgut

begins with the last third of the transverse colon and continues on. Food residue passing through the transverse colon travels across to the left side of the abdomen, where the colon angles sharply immediately inferior to the spleen, at the **left colic flexure** (splenic flexure). From there, food residue passes through the **descending colon**, which runs down the left side of the posterior abdominal wall. After entering the pelvis inferiorly, it becomes the s-shaped **sigmoid colon**, which extends medially to the midline (Figure 18).

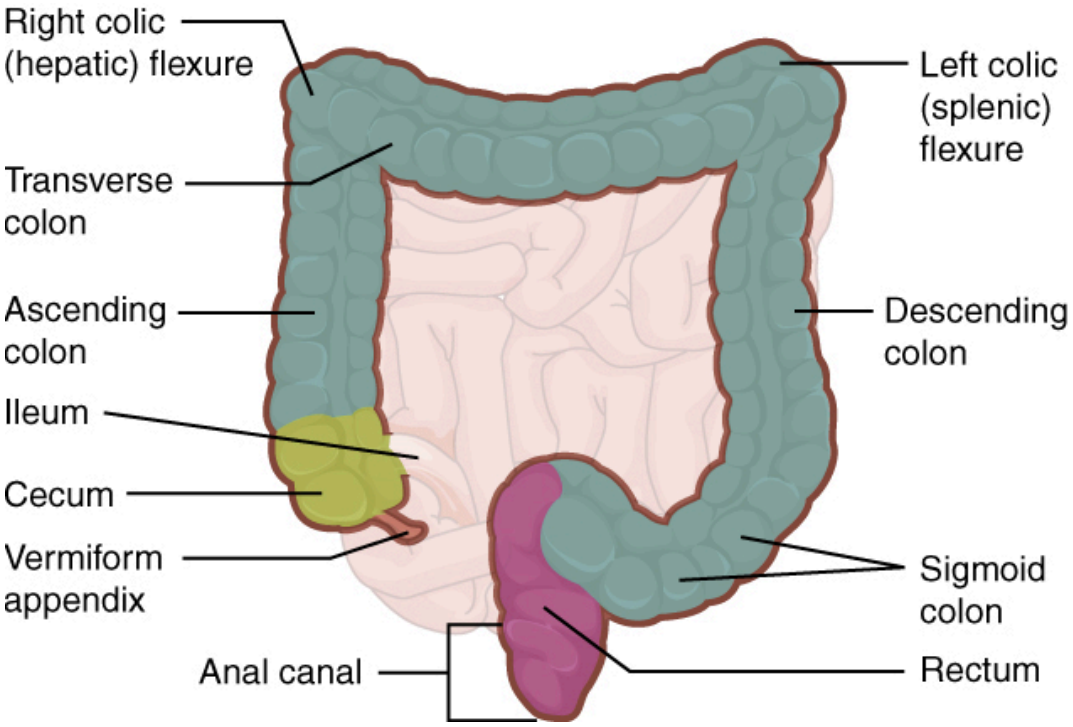


Figure 18. Large Intestine. *The large intestine includes the cecum, colon, and rectum.*

3. Rectum: Food residue leaving the sigmoid colon enters the **rectum** in the pelvis, near the third sacral vertebra. The final 20.3 cm (8 in) of the alimentary canal, the rectum extends anterior to the sacrum and coccyx. Even though rectum is Latin for "straight," this structure follows the curved contour of the sacrum and has three lateral bends that create a trio of internal transverse folds called the **rectal valves**. These valves help separate the feces from gas to prevent the simultaneous passage of feces and gas.

4. Anal Canal: Finally, food residue reaches the last part of the large intestine, the anal canal, which is located in the perineum, completely outside of the abdominopelvic cavity. This 3.8–5 cm (1.5–2 in) long structure opens to the exterior of the body at the anus. The anal canal includes two sphincters. The internal anal sphincter is made of smooth muscle, and its contractions are involuntary. The external anal sphincter is made of skeletal muscle, which is under voluntary control. Except when defecating, both usually remain closed.

Histology: There are several notable differences between the walls of the large and small intestines (Figure 19). For example, few enzyme-secreting cells are found in the wall of the large intestine, and there are no circular folds or villi. Other than in the anal canal, the mucosa of the colon is simple columnar epithelium made mostly of enterocytes (absorptive cells) and goblet cells. In addition, the wall of the large intestine has far more intestinal glands, which contain a vast population of enterocytes and goblet cells. These goblet cells secrete mucus that eases the movement of feces and protects the intestine from the effects of the acids and gases produced by enteric bacteria. The enterocytes absorb water and salts as well as vitamins produced by your intestinal bacteria.

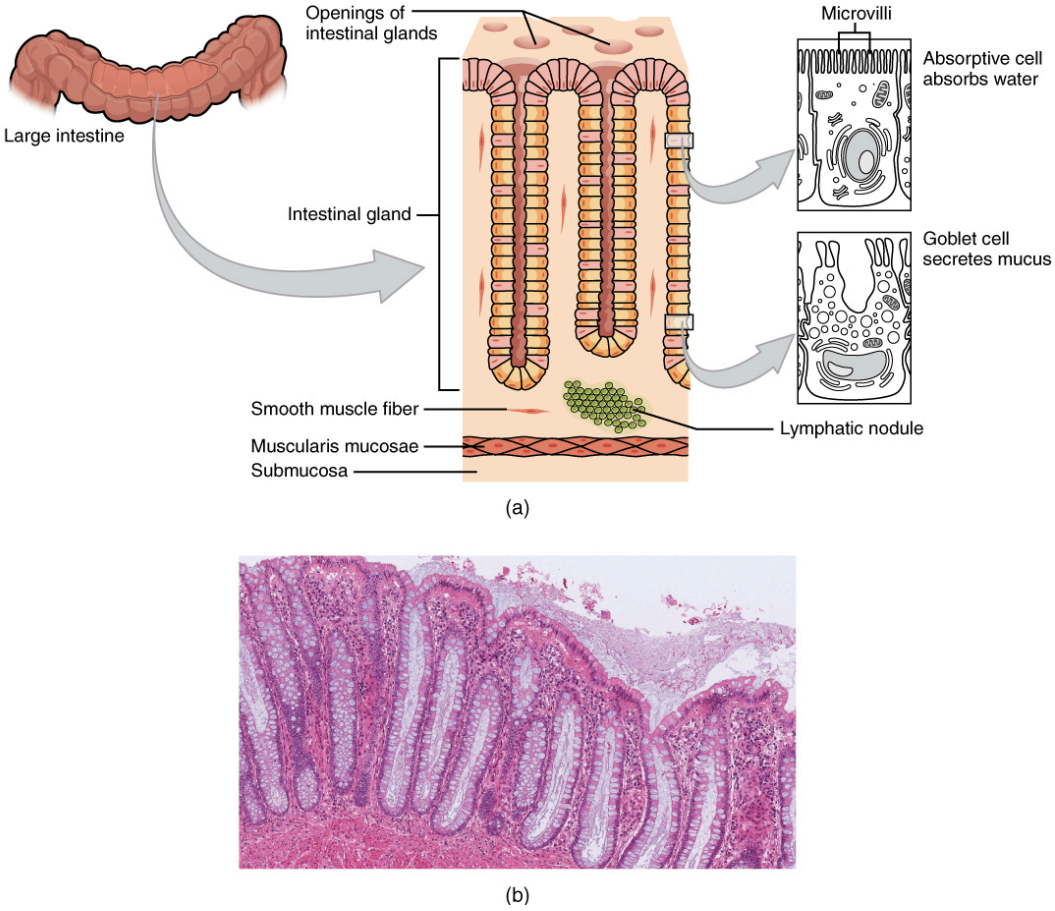


Figure 19. Histology of the Large Intestine. *(a) The histologies of the large intestine and small intestine (not shown) are adapted for the digestive functions of each organ. (b) This micrograph shows the colon's simple columnar epithelium and goblet cells. LM x 465. (credit b: Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

Digestive Functions of the Large Intestine: The residue of chyme that enters the large intestine contains few nutrients except water, which is reabsorbed as the residue lingers in the large intestine, typically for 12 to 24 hours. Thus, it may not surprise you that the large intestine can be completely removed without significantly affecting digestive functioning. For example, in severe cases of inflammatory bowel disease, the large intestine can be removed by a procedure known as a colectomy. Often, a new fecal pouch can be crafted from the small intestine and sutured to the anus, but if not, an ileostomy can be created by bringing the distal ileum through the abdominal wall, allowing the watery chyme to be collected in a bag-like adhesive appliance.

Mechanical Digestion: In the large intestine, mechanical digestion begins when chyme moves from the ileum into the cecum, an activity regulated by the ileocecal sphincter. Right after you eat, peristalsis in the ileum forces chyme into the cecum. When the cecum is distended with chyme, contractions of the ileocecal sphincter strengthen. Once chyme enters the cecum, colon movements begin.

Mechanical digestion in the large intestine includes a combination of three types of movements. The presence of food residues in the colon stimulates a slow-moving **haustral contraction**. This type of movement involves sluggish segmentation, primarily in the transverse and descending colons. When a haustrum is distended with chyme, its muscle contracts, pushing the residue into the next haustrum. These contractions occur about every 30 minutes, and each last about 1 minute. These movements also mix the food residue, which helps the large intestine absorb water. The second type of movement is **peristalsis**, which, in the large intestine, is slower than in the more proximal portions of the alimentary canal. The third type is a **mass movement**. These strong waves start midway through the transverse colon and quickly force the contents toward the rectum. Mass movements usually occur three or four times per day, either while you eat or immediately afterward. Distension in the stomach and the breakdown products of digestion in the small intestine provoke

the **gastrocolic reflex**, which increases motility, including mass movements, in the colon. Fiber in the diet both softens the stool and increases the power of colonic contractions, optimizing the activities of the colon.

Chemical Digestion: Although the glands of the large intestine secrete mucus, they do not secrete digestive enzymes. Therefore, chemical digestion in the large intestine occurs exclusively because of bacteria in the lumen of the colon. Through the process of **saccharolytic fermentation**, bacteria break down some of the remaining carbohydrates. This results in the discharge of hydrogen, carbon dioxide, and methane gases that create **flatus** (gas) in the colon; flatulence is excessive flatus. Each day, up to 1500 mL of flatus is produced in the colon. More is produced when you eat foods such as beans, which are rich in otherwise indigestible sugars and complex carbohydrates like soluble dietary fiber.

Absorption, Feces Formation, and Defecation: The small intestine absorbs about 90 percent of the water you ingest (either as liquid or within solid food). The large intestine absorbs most of the remaining water, a process that converts the liquid chyme residue into semisolid **feces** ("stool"). Feces is composed of undigested food residues, unabsorbed digested substances, millions of bacteria, old epithelial cells from the GI mucosa, inorganic salts, and enough water to let it pass smoothly out of the body. Of every 500 mL (17 ounces) of food residue that enters the cecum each day, about 150 mL (5 ounces) become feces.

Feces are eliminated through contractions of the rectal muscles. You help this process by a voluntary procedure called **Valsalva's maneuver**, in which you increase intra-abdominal pressure by contracting your diaphragm and abdominal wall muscles, and closing your glottis.

The process of defecation begins when mass movements force feces from the colon into the rectum, stretching the rectal wall and provoking the defecation reflex, which eliminates feces from the rectum. This parasympathetic reflex is mediated by the spinal cord. It contracts the sigmoid colon and rectum, relaxes the internal anal sphincter, and initially contracts the external anal sphincter. The presence of feces in the anal canal sends a signal to the brain, which gives you the choice of voluntarily opening the external anal sphincter (defecating) or keeping it temporarily closed. If you decide to delay defecation, it takes a few seconds for the reflex contractions to stop and the rectal walls to relax. The next mass movement will trigger additional defecation reflexes until you defecate.

If defecation is delayed for an extended time, additional water is absorbed, making the feces firmer and potentially leading to constipation. On the other hand, if the waste matter moves too quickly through the intestines, not enough water is absorbed, and diarrhea can result. This can be caused by the ingestion of foodborne pathogens. In general, diet, health, and stress determine the frequency of bowel movements. The number of bowel movements varies greatly between individuals, ranging from two or three per day to three or four per week.

Watch this Crash Course video to learn more about the role of the intestines in digestion! Direct link: https://youtu.be/jGme7BRkpuQ

Part 6: Accessory Organs in Digestion: The Liver, Pancreas, and Gallbladder

Chemical digestion in the small intestine relies on the activities of three accessory digestive organs: the liver, pancreas, and gallbladder (Figure 20). The digestive role of the liver is to produce bile and export it to the duodenum. The gallbladder primarily stores, concentrates, and releases bile. The pancreas produces pancreatic juice, which contains digestive enzymes and bicarbonate ions, and delivers it to the duodenum.

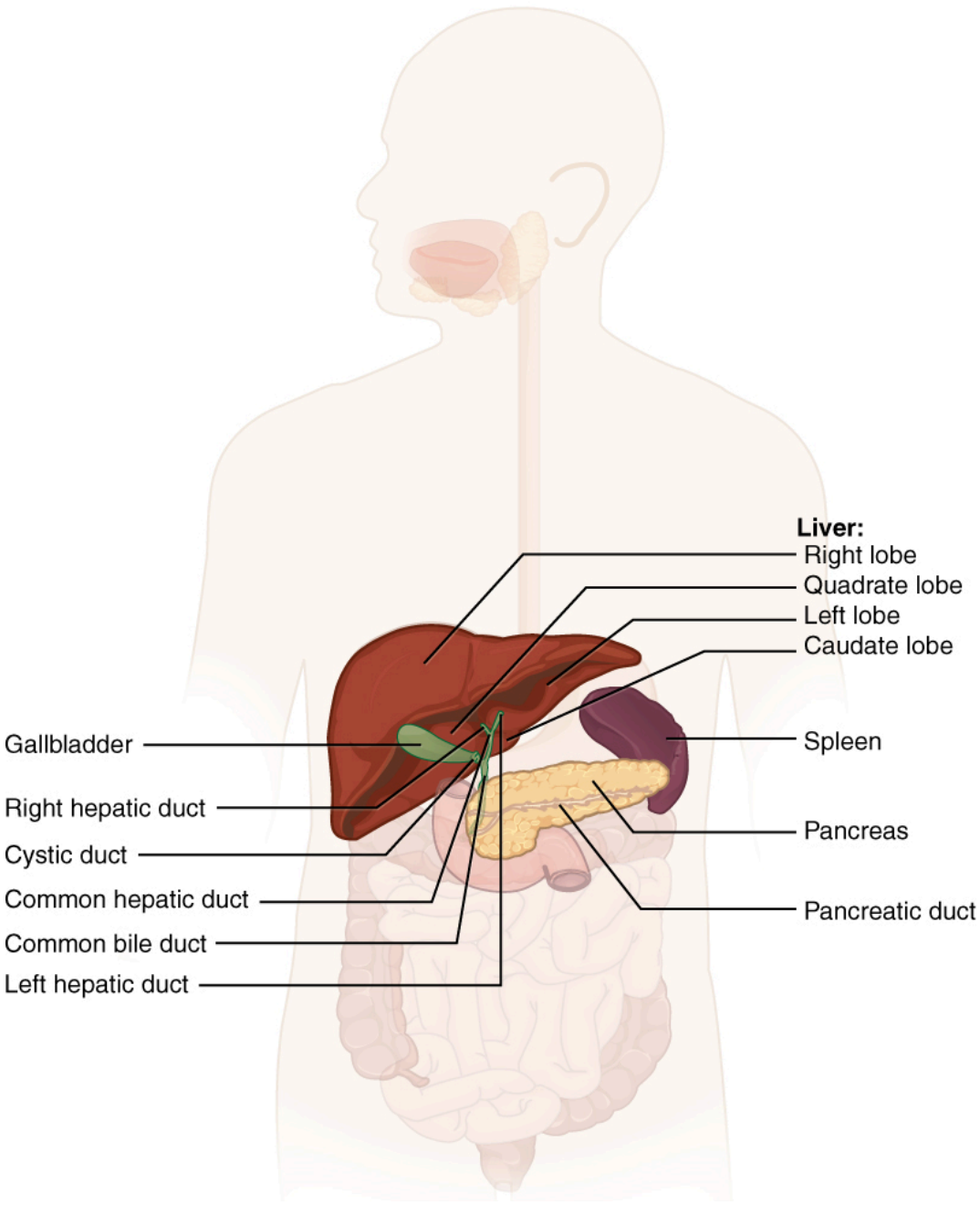


Figure 20. Accessory Organs. *The liver, pancreas, and gallbladder are considered accessory digestive organs, but their roles in the digestive system are vital.*

The Liver: The liver is the largest gland in the body, weighing about three pounds in an adult. It is also one of the most important organs. In addition to being an accessory digestive organ, it plays a number of roles in metabolism and regulation. The liver lies inferior to the diaphragm in the right upper quadrant of the abdominal cavity and receives protection from the surrounding ribs.

The liver is divided into two primary lobes: a large right lobe and a much smaller left lobe. In the right lobe, some anatomists also identify an inferior quadrate lobe and a posterior caudate lobe, which are defined by internal features. The liver is connected to the abdominal wall and diaphragm by five peritoneal folds referred to as ligaments. These are the falciform ligament, the coronary ligament, two lateral ligaments, and the ligamentum teres hepatis. The falciform ligament and ligamentum teres hepatis are actually remnants of the

umbilical vein, and separate the right and left lobes anteriorly. The lesser omentum tethers the liver to the lesser curvature of the stomach.

The **porta hepatis** ("gate to the liver") is where the **hepatic artery** and **hepatic portal vein** enter the liver. These two vessels, along with the common hepatic duct, run behind the lateral border of the lesser omentum on the way to their destinations. The hepatic artery delivers oxygenated blood from the heart to the liver (Figure 21). The hepatic portal vein delivers partially deoxygenated blood containing nutrients absorbed from the small intestine and actually supplies more oxygen to the liver than do the much smaller hepatic arteries. In addition to nutrients, drugs and toxins are also absorbed. After processing the bloodborne nutrients and toxins, the liver releases nutrients needed by other cells back into the blood, which drains into the central vein and then through the hepatic vein to the inferior vena cava. With this hepatic portal circulation, all blood from the alimentary canal passes through the liver.

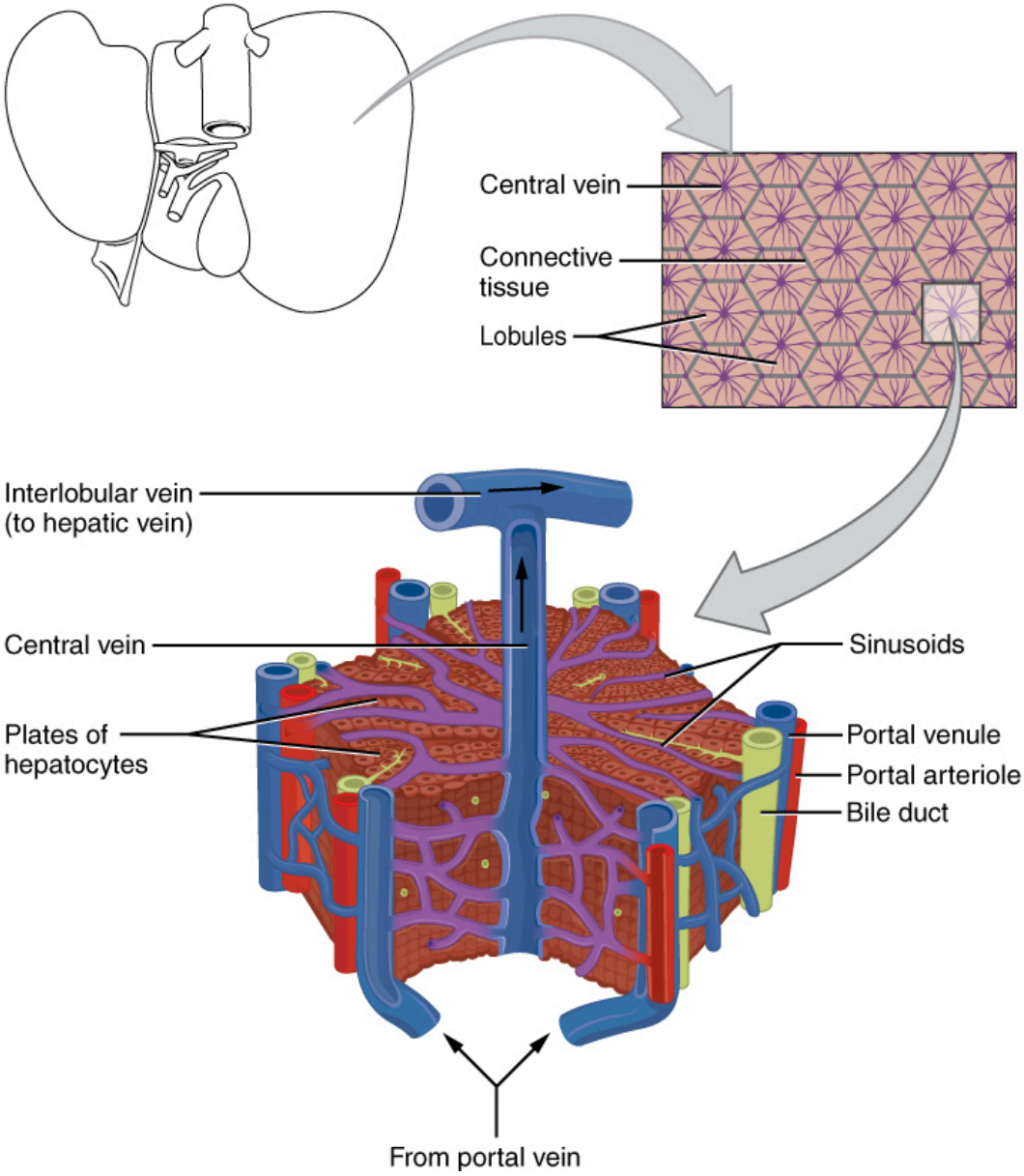


Figure 21. Microscopic Anatomy of the Liver. *The liver receives oxygenated blood from the hepatic artery and nutrient-rich deoxygenated blood from the hepatic portal vein.*

Bile: Recall that lipids are hydrophobic, that is, they do not dissolve in water. Thus, before they can be digested

in the watery environment of the small intestine, large lipid globules must be broken down into smaller lipid globules, a process called emulsification. **Bile** is a mixture secreted by the liver to accomplish the emulsification of lipids in the small intestine.

Hepatocytes secrete about one liter of bile each day. A yellow-brown or yellow-green alkaline solution (pH 7.6 to 8.6), bile is a mixture of water, bile salts, bile pigments, phospholipids (such as lecithin), electrolytes, cholesterol, and triglycerides. The components most critical to emulsification are bile salts and phospholipids, which have a nonpolar (hydrophobic) region as well as a polar (hydrophilic) region. The hydrophobic region interacts with the large lipid molecules, whereas the hydrophilic region interacts with the watery chyme in the intestine. This results in the large lipid globules being pulled apart into many tiny lipid fragments of about 1 μm in diameter. This change dramatically increases the surface area available for lipid-digesting enzyme activity. This is the same way dish soap works on fats mixed with water.

Bile salts act as emulsifying agents, so they are also important for the absorption of digested lipids. While most constituents of bile are eliminated in feces, bile salts are reclaimed by the **enterohepatic circulation**. Once bile salts reach the ileum, they are absorbed and returned to the liver in the hepatic portal blood. The hepatocytes then excrete the bile salts into newly formed bile. Thus, this precious resource is recycled.

Bilirubin, the main bile pigment, is a waste product produced when the spleen removes old or damaged red blood cells from the circulation. These breakdown products, including proteins, iron, and toxic bilirubin, are transported to the liver via the splenic vein of the hepatic portal system. In the liver, proteins and iron are recycled, whereas bilirubin is excreted in the bile. It accounts for the green color of bile. Bilirubin is eventually transformed by intestinal bacteria into stercobilin, a brown pigment that gives your stool its characteristic color! In some disease states, bile does not enter the intestine, resulting in white ('acholic') stool with a high fat content, since virtually no fats are broken down or absorbed.

Hepatocytes work non-stop, but bile production increases when fatty chyme enters the duodenum and stimulates the secretion of the gut hormone secretin. Between meals, bile is produced but conserved. The valve-like hepatopancreatic ampulla closes, allowing bile to divert to the gallbladder, where it is concentrated and stored until the next meal.

The Pancreas: The soft, oblong, glandular **pancreas** lies transversely in the retroperitoneum behind the stomach. Its head is nestled into the "c-shaped" curvature of the duodenum with the body extending to the left about 15.2 cm (6 in) and ending as a tapering tail in the hilum of the spleen. It is a curious mix of exocrine (secreting digestive enzymes) and endocrine (releasing hormones into the blood) functions (Figure 22).

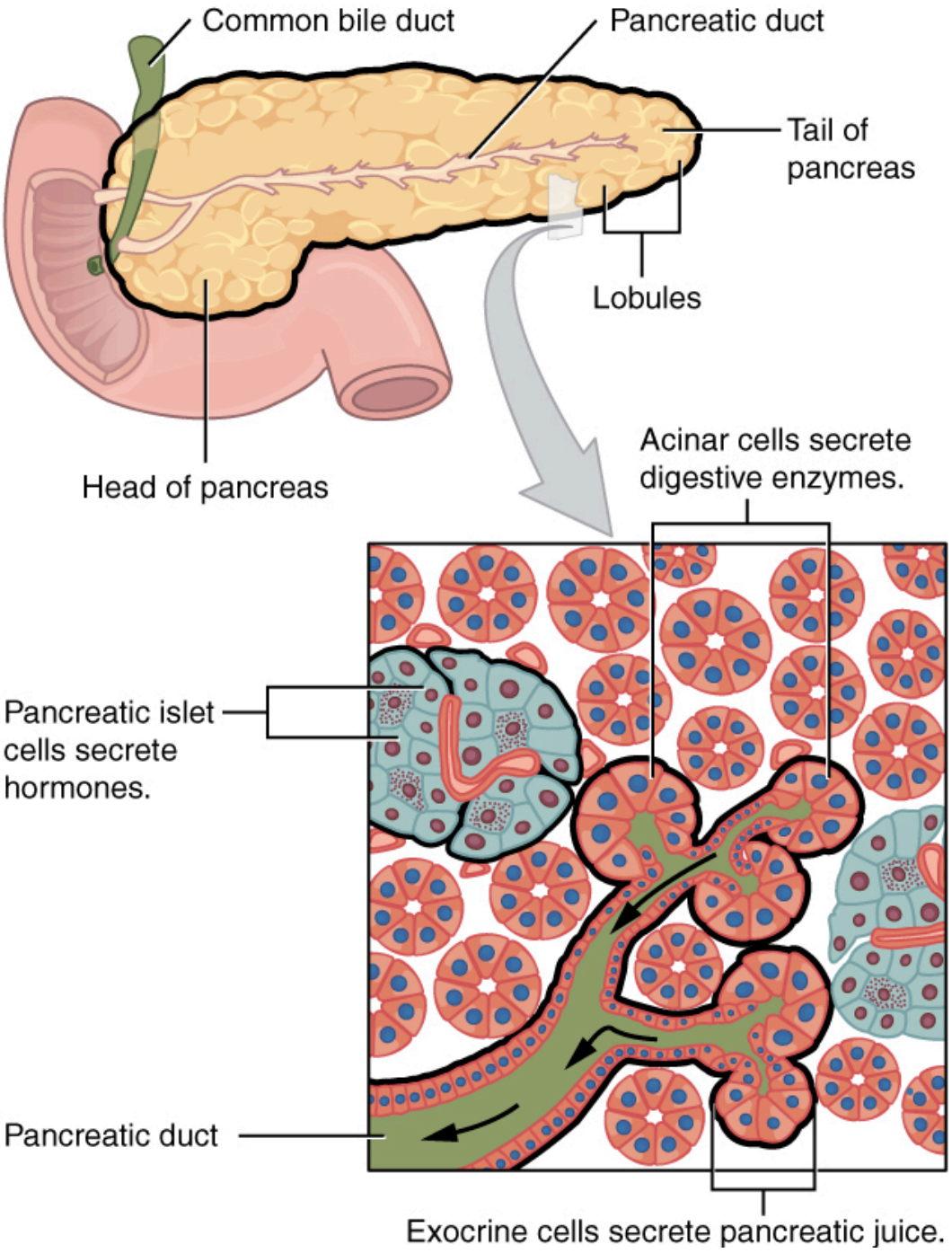


Figure 22. Exocrine and Endocrine Pancreas. *The pancreas has a head, a body, and a tail. It delivers pancreatic juice to the duodenum through the pancreatic duct.*

The exocrine part of the pancreas arises as little grape-like cell clusters, each called an **acinus** (plural = acini), located at the terminal ends of pancreatic ducts. These acinar cells secrete enzyme-rich **pancreatic juice** into tiny merging ducts that form two dominant ducts. The larger duct fuses with the common bile duct (carrying bile from the liver and gallbladder) just before entering the duodenum via a common opening (the hepatopancreatic ampulla). The smooth muscle sphincter of the hepatopancreatic ampulla controls the release of pancreatic juice and bile into the small intestine. The second and smaller pancreatic duct, the **accessory duct** (duct of Santorini), runs from the pancreas directly into the duodenum, approximately 1 inch above the hepatopancreatic ampulla. When present, it is a persistent remnant of pancreatic development.

Scattered through the sea of exocrine acini are small islands of endocrine cells, the islets of Langerhans. These vital cells produce the hormones pancreatic polypeptide, insulin, glucagon, and somatostatin.

Pancreatic Juice: The pancreas produces over a liter of pancreatic juice each day. Unlike bile, it is clear and composed mostly of water along with some salts, sodium bicarbonate, and several digestive enzymes. Sodium bicarbonate is responsible for the slight alkalinity of pancreatic juice (pH 7.1 to 8.2), which serves to buffer the acidic gastric juice in chyme, inactivate pepsin from the stomach, and create an optimal environment for the activity of pH-sensitive digestive enzymes in the small intestine. Pancreatic enzymes are active in the digestion of sugars, proteins, and fats.

The pancreas produces protein-digesting enzymes in their inactive forms. These enzymes are activated in the duodenum. If produced in an active form, they would digest the pancreas (which is exactly what occurs in the disease, pancreatitis). The intestinal brush border enzyme **enteropeptidase** stimulates the activation of trypsin from trypsinogen of the pancreas, which in turn changes the pancreatic enzymes procarboxypeptidase and chymotrypsinogen into their active forms, carboxypeptidase and chymotrypsin.

The enzymes that digest starch (amylase), fat (lipase), and nucleic acids (nuclease) are secreted in their active forms, since they do not attack the pancreas as do the protein-digesting enzymes.

Pancreatic Secretion: Regulation of pancreatic secretion is the job of hormones and the parasympathetic nervous system. The entry of acidic chyme into the duodenum stimulates the release of secretin, which in turn causes the duct cells to release bicarbonate-rich pancreatic juice. The presence of proteins and fats in the duodenum stimulates the secretion of cholecystokinin, which then stimulates the acini to secrete enzyme-rich pancreatic juice and enhances the activity of secretin. Parasympathetic regulation occurs mainly during the cephalic and gastric phases of gastric secretion, when vagal stimulation prompts the secretion of pancreatic juice.

Usually, the pancreas secretes just enough bicarbonate to counterbalance the amount of HCl produced in the stomach. Hydrogen ions enter the blood when bicarbonate is secreted by the pancreas. Thus, the acidic blood draining from the pancreas neutralizes the alkaline blood draining from the stomach, maintaining the pH of the venous blood that flows to the liver.

The Gallbladder: The **gallbladder** is 8–10 cm (~3–4 in) long and is nested in a shallow area on the posterior aspect of the right lobe of the liver. This muscular sac stores, concentrates, and, when stimulated, propels the bile into the duodenum via the common bile duct. It is divided into three regions. The fundus is the widest portion and tapers medially into the body, which in turn narrows to become the neck. The neck angles slightly superiorly as it approaches the hepatic duct. The cystic duct is 1–2 cm long and turns inferiorly as it bridges the neck and hepatic duct.

The simple columnar epithelium of the gallbladder mucosa is organized in rugae, similar to those of the stomach. There is no submucosa in the gallbladder wall. The wall's middle, muscular coat is made of smooth muscle fibers. When these fibers contract, the gallbladder's contents are ejected through the **cystic duct** and into the **bile duct** (Figure 23). Visceral peritoneum reflected from the liver capsule holds the gallbladder against the liver and forms the outer coat of the gallbladder. The gallbladder's mucosa absorbs water and ions from bile, concentrating it by up to 10-fold.

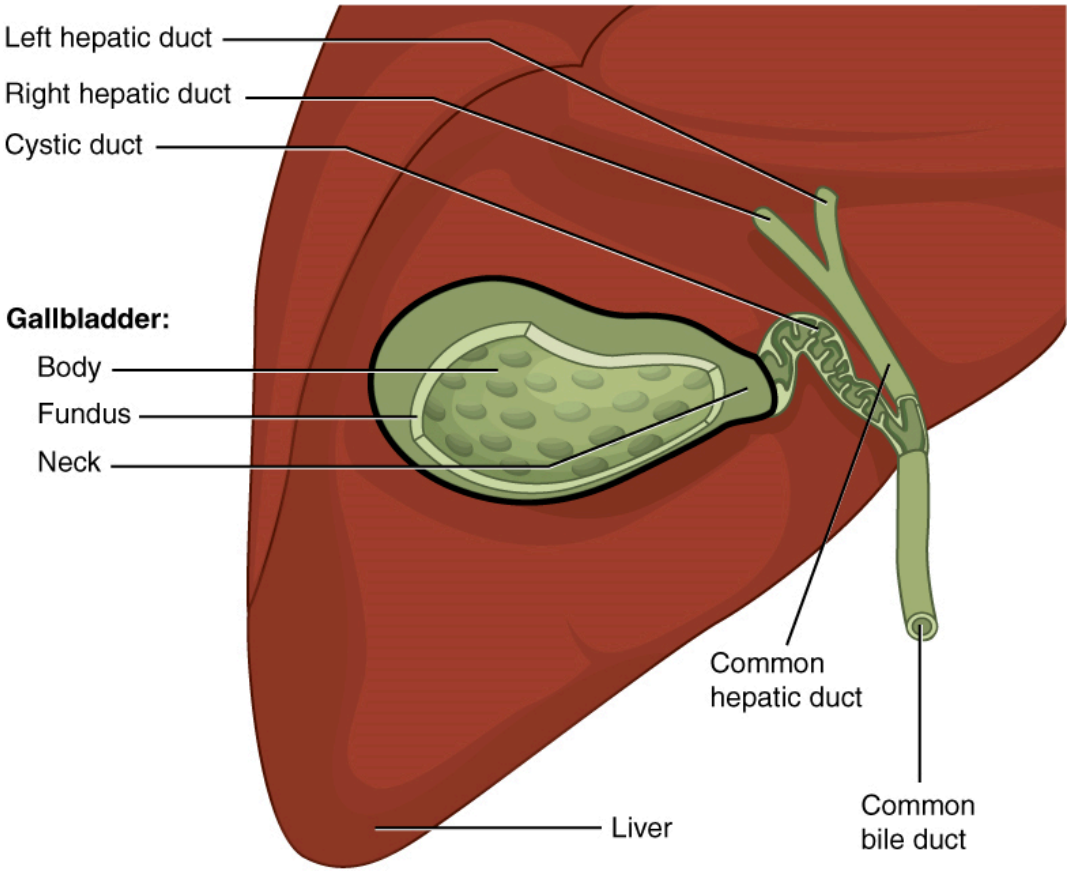


Figure 23. Gallbladder. *The gallbladder stores and concentrates bile, and releases it into the two-way cystic duct when it is needed by the small intestine.*

Part 7: Chemical Digestion and Absorption: A Closer Look

As you have learned, the process of mechanical digestion is relatively simple. It involves the physical breakdown of food but does not alter its chemical makeup. Chemical digestion, on the other hand, is a complex process that reduces food into its chemical building blocks, which are then absorbed to nourish the cells of the body (Figure 24). In this section, you will look more closely at the processes of chemical digestion and absorption.

Chemical Digestion: Large food molecules (for example, proteins, lipids, nucleic acids, and starches) must be broken down into subunits that are small enough to be absorbed by the lining of the alimentary canal. This is accomplished by enzymes through hydrolysis. The many enzymes involved in chemical digestion are summarized in Table 5.

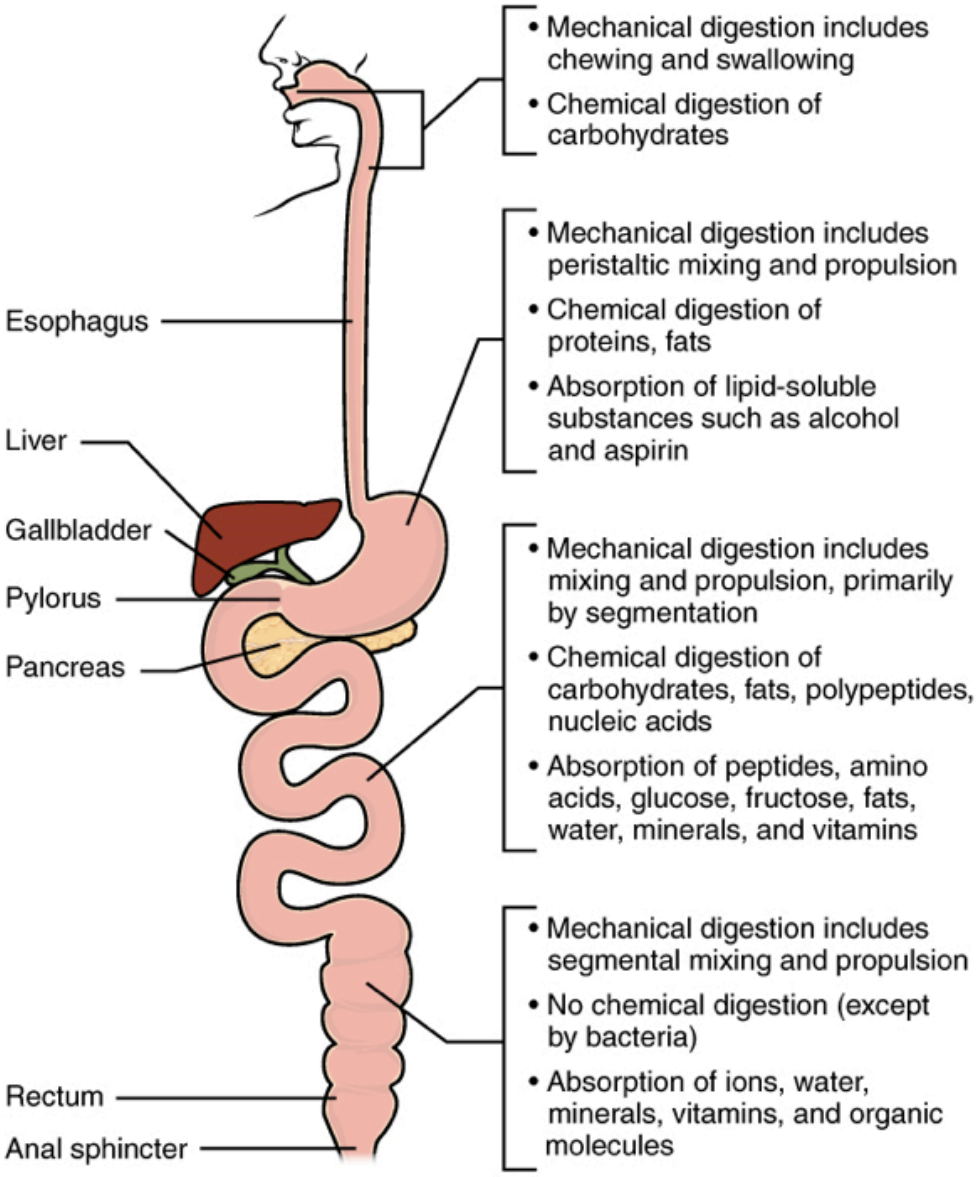


Figure 24. Digestion and Absorption. *Digestion begins in the mouth and continues as food travels through the small intestine. Most absorption occurs in the small intestine.*

Table 5: The Digestive Enzymes

Enzyme category	Enzyme name	Source	Substrate	Product
Salivary enzymes	Lingual lipase	Lingual glands	Triglycerides	Free fatty acids + diglycerides
Salivary enzymes	Salivary amylase	Salivary glands	Polysaccharides (starch, glycogen)	Maltose (and dextrins)
Gastric enzymes	Gastric lipase	Chief cells	Triglycerides	Fatty acids + monoglycerides
Gastric enzymes	Pepsin	Chief cells	Proteins	Peptides
Brush border enzymes	Lactase	Small intestine	Lactose	Glucose + galactose
Brush border enzymes	Maltase	Small intestine	Maltose	Glucose
Brush border enzymes	Sucrase	Small intestine	Sucrose	Glucose + fructose
Brush border enzymes	Nucleotidases & phosphatases	Small intestine	Nucleotides	Phosphate ions + nitrogenous bases + pentoses
Brush border enzymes	Peptidases	Small intestine	Aminopeptidase: amino acids at amino end of peptides Carboxypeptidase: amino acids at carboxyl end of peptides Dipeptidase: dipeptides Enteropeptidase: trypsinogen	Aminopeptidase & carboxypeptidase: amino acids + peptides Dipeptidase: amino acids Enteropeptidase: trypsin
Pancreatic enzymes	Carboxypeptidase	Acinar cells	Amino acids at carboxyl end of proteins/ polypeptides	Amino acids + peptides
Pancreatic enzymes	Chymotrypsin (released as chymotrypsinogen)	Acinar cells	Proteins/ polypeptides	Peptides
Pancreatic enzymes	Trypsin (released as trypsinogen)	Acinar cells	Proteins/ polypeptides (including chymotrypsinogen)	Peptides (including chymotrypsin)
Pancreatic enzymes	Nucleases	Acinar cells	Ribonuclease: ribonucleic acids Deoxyribonuclease: deoxyribonucleic acids	Nucleotides

Pancreatic enzymes	Pancreatic amylase	Acinar cells	Polysaccharides (starch, glycogen)	Maltose (and dextrins)
Pancreatic enzymes	Pancreatic lipase	Acinar cells	Triglycerides	Free fatty acids + diglycerides

Carbohydrate Digestion: The average Canadian diet is about 50 percent carbohydrates, which may be classified according to the number of monomers they contain of simple sugars (monosaccharides and disaccharides)and/or complex sugars (polysaccharides). Glucose, galactose, and fructose are the three monosaccharides that are commonly consumed and are readily absorbed.

Your digestive system is also able to break down the disaccharides sucrose (regular table sugar: glucose + fructose), lactose (milk sugar: glucose + galactose), and maltose (grain sugar: glucose + glucose), and the polysaccharides glycogen and starch (chains of monosaccharides). Your bodies do not produce enzymes that can break down most fibrous polysaccharides, such as cellulose. While indigestible polysaccharides do not provide any nutritional value, they do provide dietary fiber, which helps propel food through the alimentary canal.

The chemical digestion of starches begins in the mouth, where **salivary amylase** acts on starch (Table 3). There is little further chemical digestion of carbohydrates until they reach the small intestine.

In the small intestine, **pancreatic amylase** does the 'heavy lifting' for starch and carbohydrate digestion (Figure 25). After amylases break down starch into smaller fragments, the brush border enzymes continue the process. Three brush border enzymes hydrolyze sucrose, lactose, and maltose into monosaccharides. **Sucrase** splits sucrose into one molecule of fructose and one molecule of glucose; **maltase** breaks down maltose and maltotriose into two and three glucose molecules, respectively; and **lactase** breaks down lactose into one molecule of glucose and one molecule of galactose. Insufficient lactase can lead to lactose intolerance.

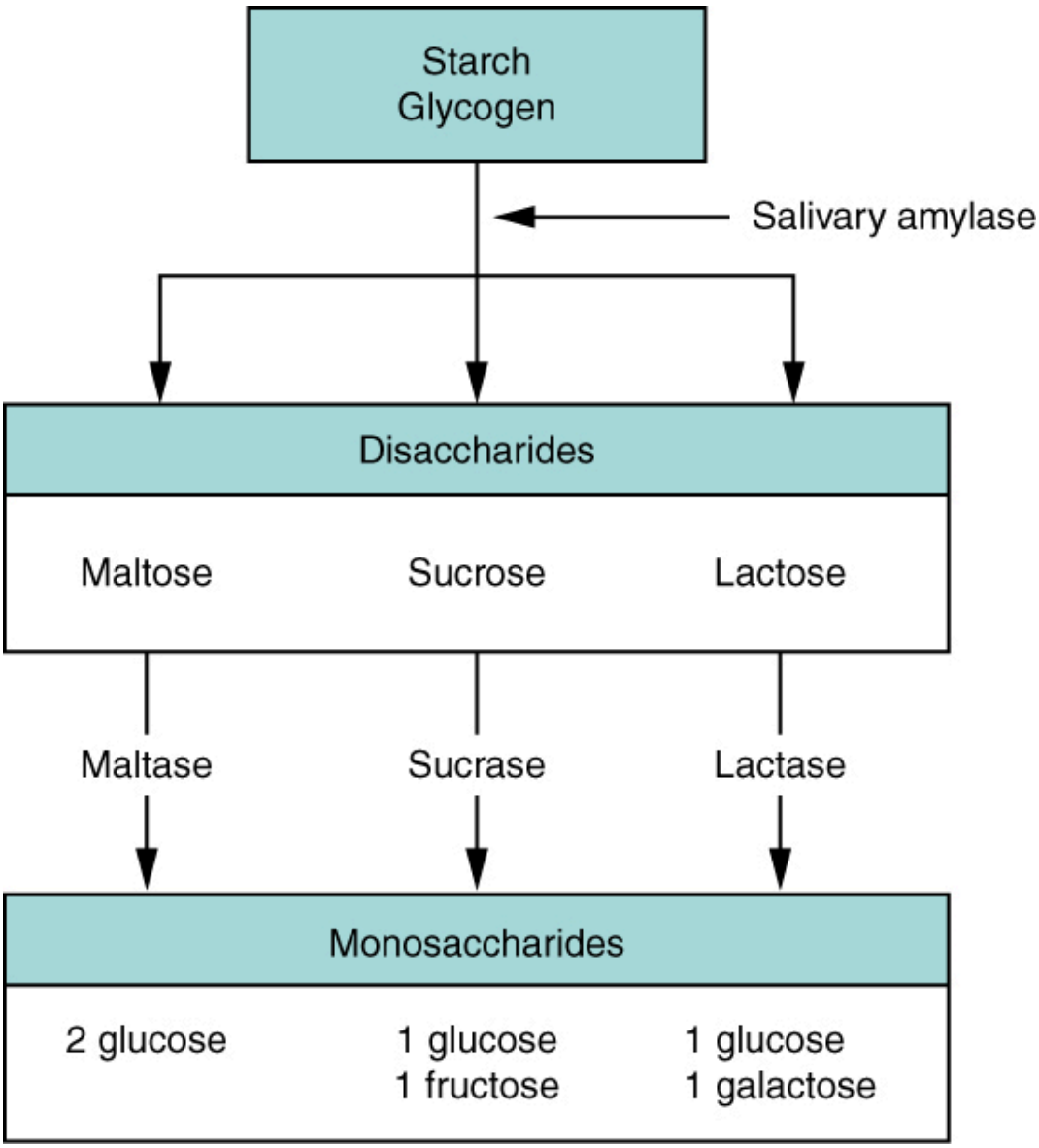


Figure 25. Carbohydrate Digestion. *Carbohydrates are broken down into their monomers in a series of steps.*

Protein Digestion: Proteins are polymers composed of amino acids linked by peptide bonds to form long chains. Digestion reduces them to their constituent amino acids. You usually consume about 15 to 20 percent of your total calorie intake as protein.

The digestion of protein starts in the stomach, where **pepsin** breaks proteins into smaller polypeptides, which then travel to the small intestine (Figure 26). Chemical digestion in the small intestine is continued by pancreatic enzymes, including **chymotrypsin** and **trypsin**, each of which act on specific bonds in amino acid sequences. At the same time, the cells of the brush border secrete enzymes such as **aminopeptidase**, **carboxypeptidase** and **dipeptidase**, which further break down peptide chains. This results in molecules small enough to enter the bloodstream (Figure 27).

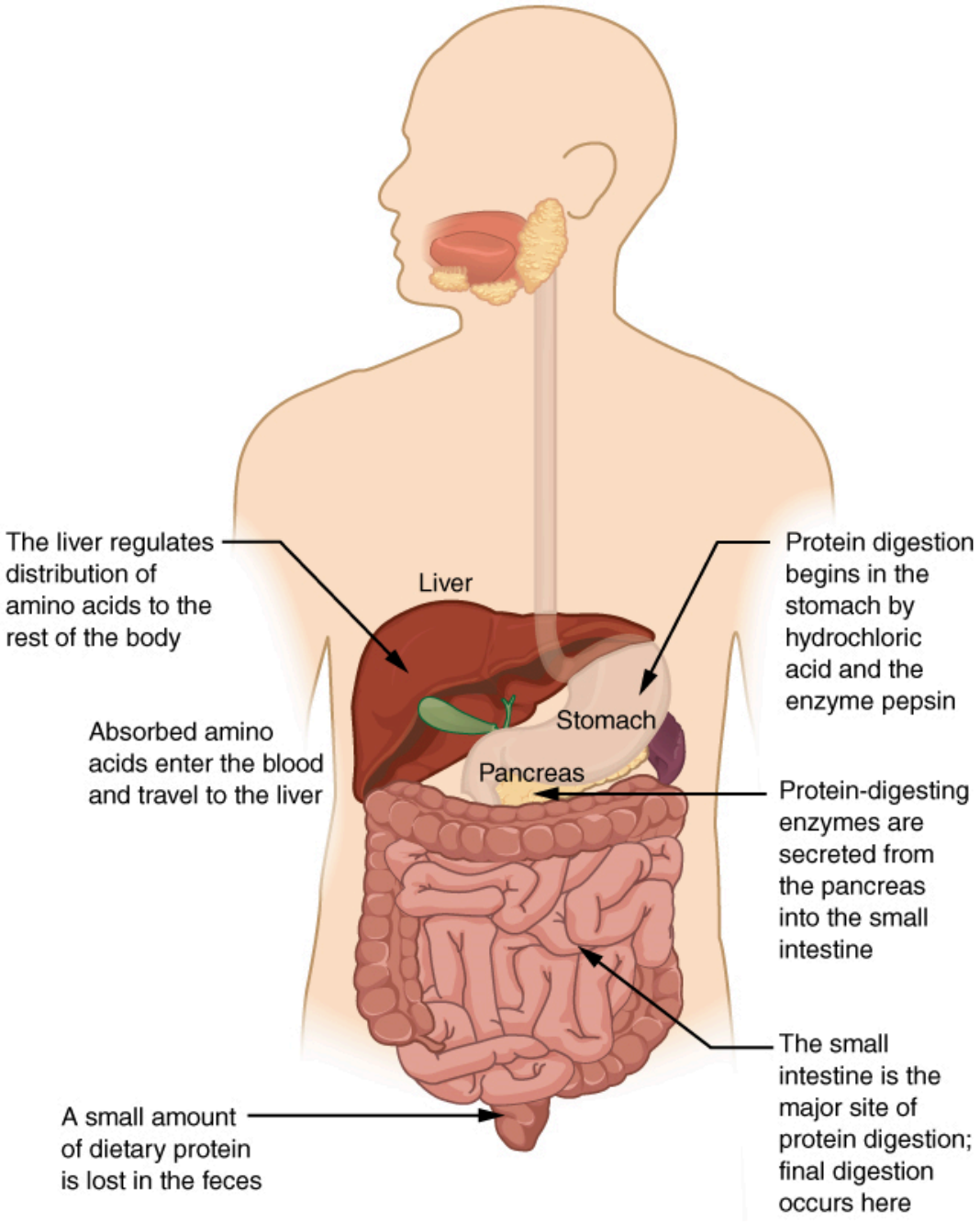


Figure 26. Sites of Protein Digestion. *The digestion of protein begins in the stomach and is completed in the small intestine.*

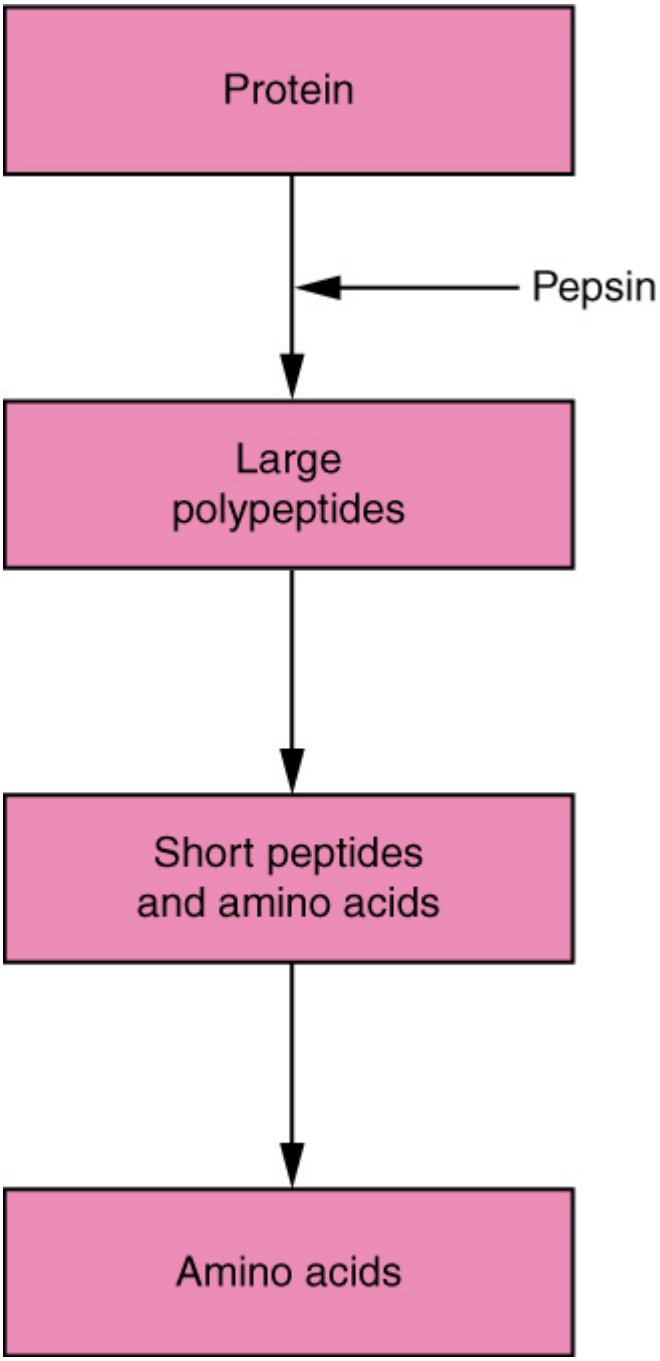


Figure 27. Protein Digestion. *Proteins are successively broken down into their amino acid components.*

Lipid Digestion: A healthy diet limits lipid intake to 35 percent of total calorie intake. The most common dietary lipids are triglycerides, which are made up of a glycerol molecule bound to three fatty acid chains. Small amounts of dietary cholesterol and phospholipids are also consumed.

The three lipases responsible for lipid digestion are **lingual lipase**, **gastric lipase**, and **pancreatic lipase**. However, because the pancreas is the only consequential source of lipase, virtually all lipid digestion occurs in the small intestine. Pancreatic lipase breaks down each triglyceride into two free fatty acids and a monoglyceride.

Nucleic Acid Digestion: The nucleic acids DNA and RNA are found in most of the foods you eat. Two types of **pancreatic nuclease** are responsible for their digestion: **deoxyribonuclease**, which digests DNA, and

ribonuclease, which digests RNA. The nucleotides produced by this digestion are further broken down by two intestinal brush border enzymes (**nucleosidase** and **phosphatase**) into pentoses, phosphates, and nitrogenous bases, which can be absorbed through the alimentary canal wall.

The large food molecules that must be broken down into subunits are summarized Table 6.

Table 6: Absorbable Food Substances

Source	Substance
Carbohydrates	Monosaccharides: glucose, galactose, fructose
Proteins	Amino acids, dipeptides, tripeptides
Triglycerides	Monoglycerides, glycerol, free fatty acids
Nucleic acids	Pentose sugars, phosphates, nitrogenous bases

Absorption: The mechanical and digestive processes have one goal: to convert food into molecules small enough to be absorbed by the epithelial cells of the intestinal villi. The absorptive capacity of the alimentary canal is almost endless. Each day, the alimentary canal processes up to 10 liters of food, liquids, and GI secretions, yet less than one liter enters the large intestine. Almost all ingested food, 80 percent of electrolytes, and 90 percent of water are absorbed in the small intestine. Although the entire small intestine is involved in the absorption of water and lipids, most absorption of carbohydrates and proteins occurs in the jejunum. Notably, bile salts and vitamin B12 are absorbed in the terminal ileum. By the time chyme passes from the ileum into the large intestine, it is essentially indigestible food residue (mainly plant fibers like cellulose), some water, and millions of bacteria (Figure 28).

Absorption can occur through five mechanisms: (1) active transport, (2) passive diffusion, (3) facilitated diffusion,(4) co-transport (or secondary active transport), and (5) endocytosis. As you will recall, active transport refers to the movement of a substance across a cell membrane going from an area of lower concentration to an area of higher concentration (up the concentration gradient). In this type of transport, proteins within the cell membrane act as “pumps,” using cellular energy (ATP) to move the substance. Passive diffusion refers to the movement of substances from an area of higher concentration to an area of lower concentration, while facilitated diffusion refers to the movement of substances from an area of higher to an area of lower concentration using a carrier protein in the cell membrane. Co-transport uses the movement of one molecule through the membrane from higher to lower concentration to power the movement of another from lower to higher. Finally, endocytosis is a transportation process in which the cell membrane engulfs material. It requires energy, generally in the form of ATP.

Because the cell’s plasma membrane is made up of hydrophobic phospholipids, water-soluble nutrients must use transport molecules embedded in the membrane to enter cells. Moreover, many substances cannot pass between the epithelial cells of the intestinal mucosa because these cells are bound together by tight junctions. Thus, nutrients generally enter blood capillaries by passing through the apical surface of epithelial cells and then out the basal surface into the interstitial fluid. Water-soluble nutrients then enter the capillary blood in the villi and travel to the liver via the hepatic portal vein.

In contrast to the water-soluble nutrients, lipid-soluble nutrients can diffuse through the plasma membrane of an intestinal epithelial cell. Once inside the cell, they are packaged for transport, exit via the base of the cell, and then enter the lacteals of the villi to be transported by lymphatic vessels to the systemic circulation via

the thoracic duct. The absorption of most nutrients through the mucosa of the intestinal villi requires active transport fueled by ATP. The routes of absorption for each food category are summarized in Table 7.

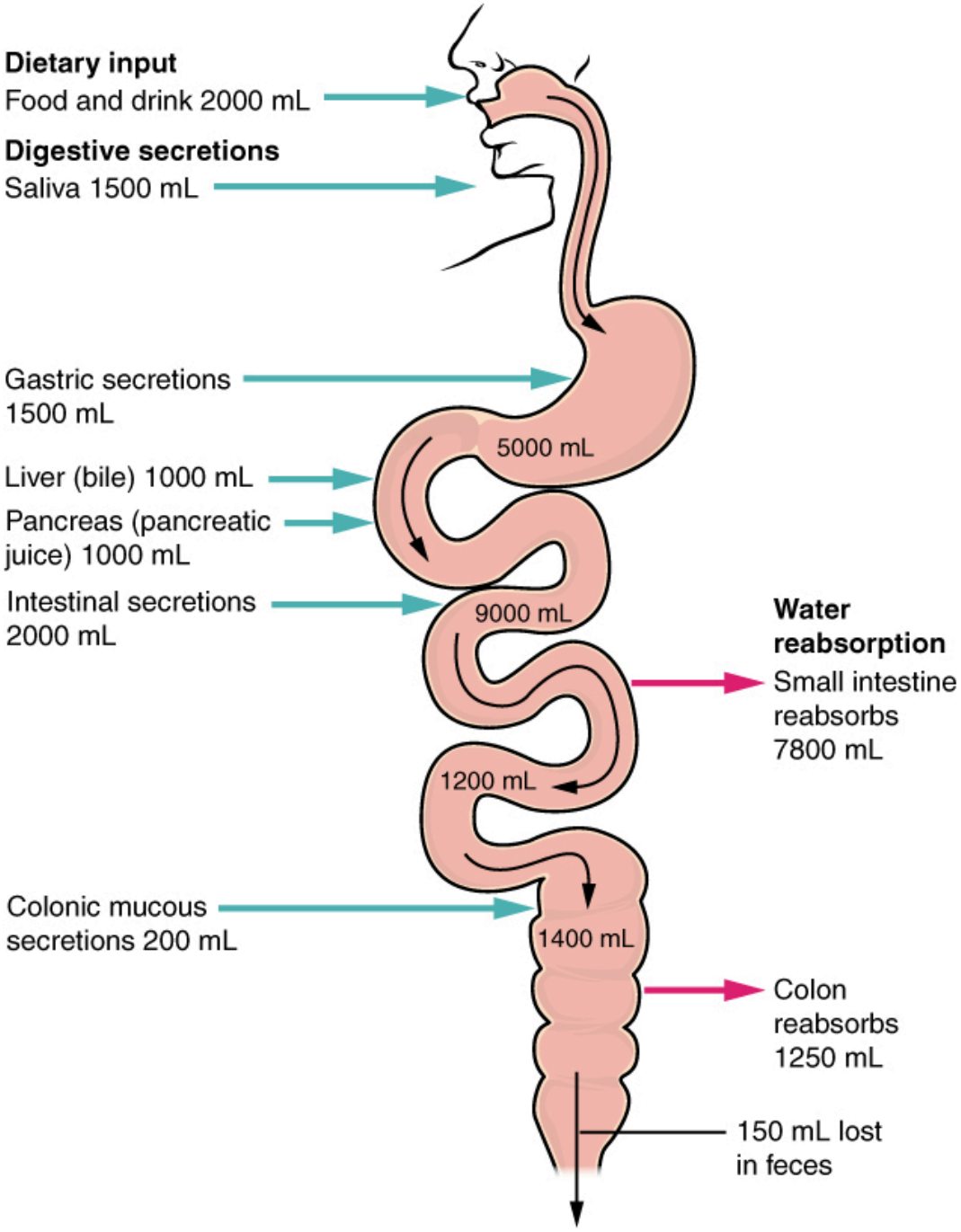


Figure 28. Digestive Secretions and Absorption of Water. *Absorption is a complex process, in which nutrients from digested food are harvested.*

Carbohydrate Absorption: All carbohydrates are absorbed in the form of monosaccharides. The small intestine is highly efficient at this, absorbing monosaccharides at an estimated rate of 120 grams per hour. All normally digested dietary carbohydrates are absorbed; indigestible fibers are eliminated in the feces. The monosaccharides glucose and galactose are transported into the epithelial cells by common protein carriers via secondary active transport (that is, co-transport with sodium ions). The monosaccharides leave these cells via facilitated diffusion and enter the capillaries through intercellular clefts. The monosaccharide fructose (which

is in fruit) is absorbed and transported by facilitated diffusion alone. The monosaccharides combine with the transport proteins immediately after the disaccharides are broken down.

Protein Absorption: Active transport mechanisms, primarily in the duodenum and jejunum, absorb most proteins as their breakdown products, amino acids. Almost all (95 to 98 percent) protein is digested and absorbed in the small intestine. The type of carrier that transports an amino acid varies. Most carriers are linked to the active transport of sodium. Short chains of two amino acids (dipeptides) or three amino acids (tripeptides) are also transported actively. However, after they enter the absorptive epithelial cells, they are broken down into their amino acids before leaving the cell via facilitated diffusion.

Table 7: Absorption in the Alimentary Canal

Food	Breakdown products	Absorption mechanism	Entry to bloodstream	Destination
Carbohydrates	Glucose	Co-transport with Na^+, facilitated diffusion out of intestinal epithelial cells	Diffusion through pores of fenestrated capillaries in villi	Liver (via hepatic portal vein)
Carbohydrates	Galactose	Co-transport with Na^+, facilitated diffusion out of intestinal epithelial cells	Diffusion through pores of fenestrated capillaries in villi	Liver (via hepatic portal vein)
Carbohydrates	Fructose	Facilitated diffusion	Diffusion through pores of fenestrated capillaries in villi	Liver (via hepatic portal vein)
Protein	Amino acids	Co-transport with Na^+, facilitated diffusion out of intestinal epithelial cells	Diffusion through pores of fenestrated capillaries in villi	Liver (via hepatic portal vein)
Lipids	Long-chain fatty acids	Simple diffusion into intestinal epithelial cells, exocytosis of chylomicrons out of intestinal epithelial cells	Paracellular transport into lacteals in villi, to left subclavian vein via lymphatic vessels	Systemic circulation via lymph entering thoracic duct
Lipids	Monoglycerides	Simple diffusion into intestinal epithelial cells, exocytosis of chylomicrons out of intestinal epithelial cells	Paracellular transport into lacteals in villi, to left subclavian vein via lymphatic vessels	Systemic circulation via lymph entering thoracic duct
Lipids	Short-chain fatty acids	Simple diffusion	Simple diffusion into, and diffusion through pores of, fenestrated capillaries in villi	Liver (via hepatic portal vein)
Lipids	Glycerol	Simple diffusion	Simple diffusion into, and diffusion through pores of, fenestrated capillaries in villi	Liver (via hepatic portal vein)
Nucleic acids	Nitrogenous bases, ribose, deoxyribose, phosphate	Active transport into intestinal epithelial cells, facilitated diffusion out of intestinal epithelial cells; also paracellular transport	Diffusion through pores of fenestrated capillaries in villi	Liver (via hepatic portal vein)

Lipid Absorption: About 95 percent of lipids are absorbed in the small intestine. Bile salts not only speed up lipid digestion, they are also essential to the absorption of the end products of lipid digestion. Short-chain fatty acids (under 6 carbon atoms in length) are relatively water soluble and can enter the absorptive cells (enterocytes) directly. Despite being hydrophobic, the small size of short-chain fatty acids enables them to be absorbed by

enterocytes via simple diffusion, and then take the same path as monosaccharides and amino acids into the blood capillary of a villus.

The large and hydrophobic long-chain fatty acids and monoacylglycerides are not so easily suspended in the watery intestinal chyme. However, bile salts and lecithin resolve this issue by enclosing them in a **micelle**, which is a tiny sphere with polar (hydrophilic) ends facing the watery environment and hydrophobic tails turned to the interior, creating a receptive environment for the long-chain fatty acids. The core also includes cholesterol and fat-soluble vitamins. Without micelles, lipids would sit on the surface of chyme and never come in contact with the absorptive surfaces of the epithelial cells. Micelles can easily squeeze between microvilli and get very near the luminal cell surface. At this point, lipid substances exit the micelle and are absorbed via simple diffusion.

The free fatty acids and monoacylglycerides that enter the epithelial cells are reincorporated into triglycerides. The triglycerides are mixed with phospholipids and cholesterol, and surrounded with a protein coat. This new complex, called a **chylomicron**, is a water-soluble lipoprotein. After being processed by the Golgi apparatus, chylomicrons are released from the cell (Figure 29). Too big to pass through the basement membranes of blood capillaries, chylomicrons instead enter the large pores of lacteals. The lacteals come together to form the lymphatic vessels. The chylomicrons are transported in the lymphatic vessels and empty through the thoracic duct into the subclavian vein of the circulatory system. Once in the bloodstream, the enzyme **lipoprotein lipase** breaks down the triglycerides of the chylomicrons into free fatty acids and glycerol. These breakdown products then pass through capillary walls to be used for energy by cells or stored in adipose tissue as fat. Liver cells combine the remaining chylomicron remnants with proteins, forming lipoproteins that transport cholesterol in the blood.

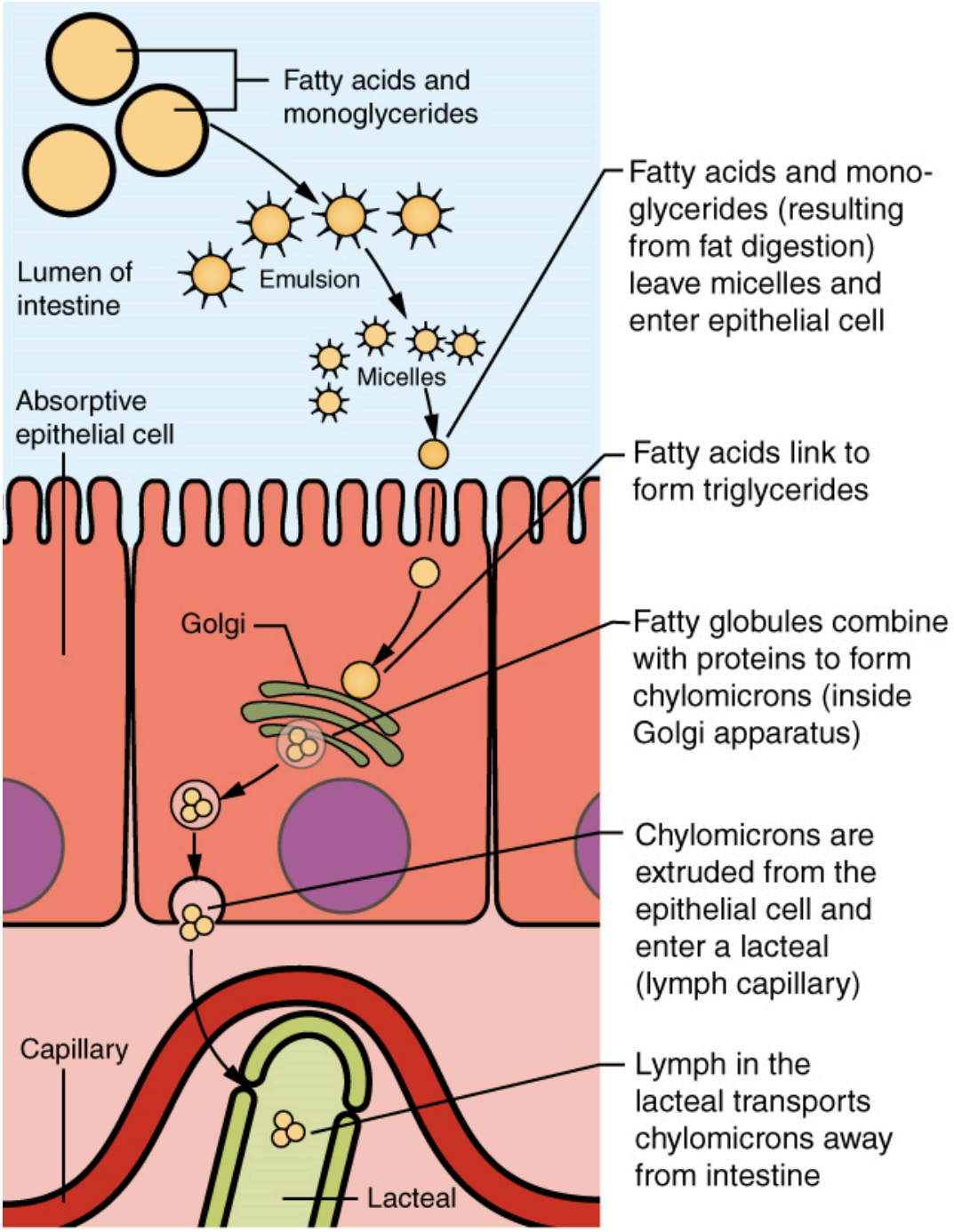


Figure 29. Lipid Absorption. *Unlike amino acids and simple sugars, lipids are transformed as they are absorbed through epithelial cells.*

Nucleic Acid Absorption: The products of nucleic acid digestion—pentose sugars, nitrogenous bases, and phosphate ions—are transported by carriers across the villus epithelium via active transport. These products then enter the bloodstream.

Mineral Absorption: The electrolytes absorbed by the small intestine are from both GI secretions and ingested foods. Since electrolytes dissociate into ions in water, most are absorbed via active transport throughout the entire small intestine. During absorption, co-transport mechanisms result in the accumulation of sodium ions inside the cells, whereas anti-port mechanisms reduce the potassium ion concentration inside

the cells. To restore the sodium-potassium gradient across the cell membrane, a sodium-potassium pump requiring ATP pumps sodium out and potassium in.

In general, all minerals that enter the intestine are absorbed, whether you need them or not. Iron and calcium are exceptions; they are absorbed in the duodenum in amounts that meet the body's current requirements.

Vitamin Absorption: The small intestine absorbs the vitamins that occur naturally in food and supplements. Fat-soluble vitamins (A, D, E, and K) are absorbed along with dietary lipids in micelles via simple diffusion. This is why you are advised to eat some fatty foods when you take fat-soluble vitamin supplements. Most water-soluble vitamins (including most B vitamins and vitamin C) are absorbed by facilitated diffusion. An exception is vitamin B12, which is a very large molecule. Intrinsic factor secreted in the stomach binds to vitamin B12, preventing its digestion and creating a complex that binds to mucosal receptors in the terminal ileum, where it is taken up by endocytosis.

Water Absorption: Each day, about nine liters of fluid enter the small intestine. About 2.3 liters are ingested in foods and beverages, and the rest is from GI secretions. About 90 percent of this water is absorbed in the small intestine. Water absorption is driven by the concentration gradient of the water: The concentration of water is higher in chyme than it is in epithelial cells. Thus, water moves down its concentration gradient from the chyme into cells. As noted earlier, much of the remaining water is then absorbed in the colon.

Unit 6: Nutrition

Unit outline

- Nutrition and Diet
- Food and metabolism
- Essential nutrients
- Essential amino acids and lipids
- Water
- Vitamins
- Minerals

Learning Objectives

At the end of this unit, you should be able to:

I. Explain the relative importance of including essential nutrients compared to nonessential nutrients in a diet.

II. Specify five essential nutritional factors.

III. Define vitamin and describe the general functions, categories and examples of vitamins.

IV. Specify six major minerals (macrominerals) required in human nutrition and one function of each.

V. Describe the general guidelines published by Health Canada for recommended types and quantities of food consumed every day.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Explain the relative importance of including essential nutrients compared to nonessential nutrients in a diet.

1. Define each of the following terms:
    - Metabolism
    - Nutrient
    - Diet
    - Recommended daily allowance
2. Clearly explain the difference between:
    - An essential nutrient and a nonessential nutrient
    - An essential amino acid and a nonessential amino acid
    - An essential fatty acid and anonessential fatty acid

II. Specify five essential nutritional factors.

1. Specify the five essential nutrient groups and describe the main function of each group of nutrients.

III. Define vitamin and describe the general functions, categories and examples of vitamins.

1. Define 'vitamin' and describe the general functions of vitamins.
2. Explain how the main difference between lipid-soluble and water-soluble vitamins leads to differences in dietary requirements for each type.
3. List all the water-soluble and fat-soluble vitamins, along with the primary function of each vitamin.

IV. Specify six major minerals (macrominerals) required in human nutrition and one function of each.

1. Specify the six major minerals in humans and describe one major function of each.
2. With the aid of specific examples, clearly distinguish between:
    - Minerals and vitamins
    - Trace minerals and major minerals

V. Describe the general guidelines published by Health Canada for recommended types and quantities of food consumed every day.

1. Describe the approximate quantity of each type of food that should be consumed every day.

Overview of Metabolic Reactions: Metabolic processes are constantly taking place in the body. **Metabolism** is the sum of all of the chemical reactions that are involved in catabolism and anabolism. The reactions governing the breakdown of food to obtain energy are called catabolic reactions. In catabolic reactions, large organic molecules are broken down to smaller molecules, releasing the energy contained in the chemical bonds. Some of this energy is used to form adenosine triphosphate (ATP). Conversely, anabolic reactions use the energy released by catabolic reactions to synthesize larger molecules from smaller ones, such as when the body forms proteins by stringing together amino acids. Both sets of reactions are critical to maintaining life.

The energy obtained from ATP drives all bodily functions, such as contracting muscles, maintaining the

electrical potential of nerve cells, and absorbing food in the gastrointestinal tract. The metabolic reactions that produce ATP come from various sources (Figure 1).

Because catabolic reactions produce energy and anabolic reactions use energy, ideally, energy usage would balance the energy released. If the net energy change is positive (catabolic reactions release more energy than the anabolic reactions use), then the body stores the excess energy by building fat molecules for long-term storage. On the other hand, if the net energy change is negative (catabolic reactions release less energy than anabolic reactions use), the body uses stored energy to compensate for the deficiency of energy released by catabolism.

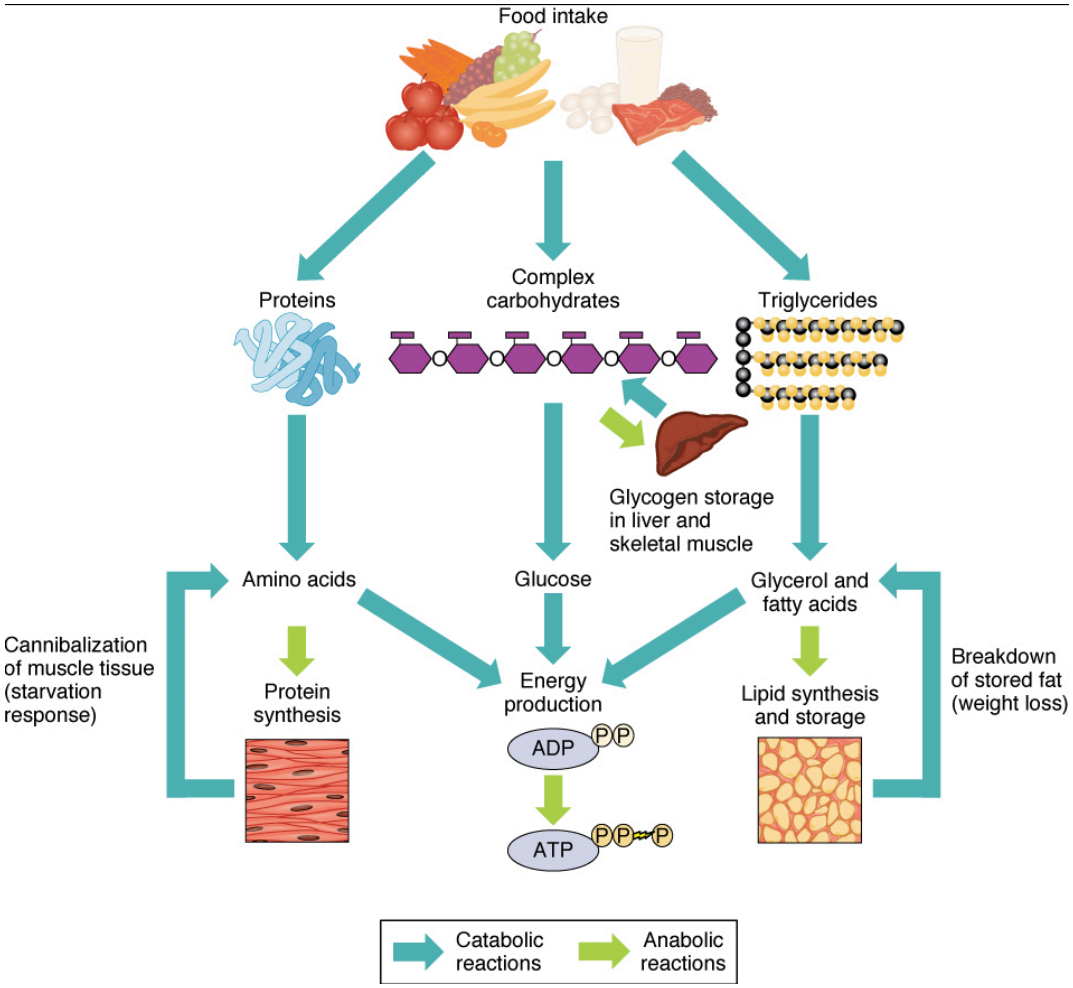


Figure 1. Sources of ATP. *During catabolic reactions, proteins are broken down into amino acids, lipids are broken down into fatty acids, and polysaccharides are broken down into monosaccharides. These building blocks are then used for the synthesis of molecules in anabolic reactions.*

Of the four major macromolecular groups (carbohydrates, lipids, proteins, and nucleic acids) that are processed by digestion, carbohydrates are the most common source of energy to fuel the body. They take the form of either complex carbohydrates, polysaccharides like starch and glycogen, or simple sugars (monosaccharides) like glucose and fructose. Among the monosaccharides, glucose is the most common fuel for ATP production in cells, and as such, there are a number of endocrine control mechanisms to regulate glucose concentration in the bloodstream. Excess glucose is either stored as an energy reserve in the liver and skeletal muscles as the complex polymer glycogen, or it is converted into fat (triglyceride) in adipose cells (adipocytes).

Among the lipids (fats), triglycerides are most often used for energy via a metabolic process called β-oxidation. About one-half of excess fat is stored in adipocytes that accumulate in the subcutaneous tissue under the skin, whereas the rest is stored in adipocytes in other tissues and organs.

Proteins, which are polymers, can be broken down into their monomers, individual amino acids. Amino acids

can be used as building blocks of new proteins or broken down further for the production of ATP. When one is chronically starving, this use of amino acids to obtain energy can lead to a wasting away of the body as more and more proteins are broken down.

Nucleic acids are present in most of the foods you eat. During digestion, nucleic acids including DNA and various RNAs are broken down into their constituent nucleotides. These nucleotides are readily absorbed and transported throughout the body to be used by individual cells during nucleic acid metabolism.

Nutrition and Diet: The carbohydrates, lipids, and proteins in the foods you eat are used for energy to power molecular, cellular, and organ system activities. Importantly, the energy is stored primarily as fats. The quantity and quality of food that is ingested, digested, and absorbed affects the amount of fat that is stored as excess calories. A dietary nutrient is a substance that must be ingested and is essential for growth and the maintenance of life. Diet—both what you eat and how much you eat—has a dramatic impact on your health. Eating too much or too little food can lead to serious medical issues, including cardiovascular disease, cancer, and diabetes, among others. Combine an unhealthy diet with unhealthy environmental conditions, such as smoking, and the potential medical complications increase significantly.

Food and Metabolism: The amount of energy that is needed or ingested per day is measured in calories. The nutritional **Calorie** (C) is the amount of heat it takes to raise 1 kg (1000 g) of water by 1 °C. This is different from the calorie (c) used in the physical sciences, which is the amount of heat it takes to raise 1 g of water by 1 °C. When we refer to "calorie," we are referring to the nutritional Calorie.

On average, a person needs 1500 to 2000 calories per day to sustain (or carry out) daily activities. The total number of calories needed by one person is dependent on their body mass, age, height, gender, activity level, and the amount of exercise per day. If exercise is regular part of one's day, more calories are required.

The Recommended Dietary Allowance (RDA) is used as a general guide for the amount of micronutrients, such as vitamins and minerals, which are required on a daily basis. According to Health Canada, the Recommended Dietary Allowance (RDA) is the average daily dietary intake level that is sufficient to meet the nutrient requirement of nearly all healthy individuals for a specific gender or age group.

The type of food ingested also affects the body's metabolism. Processing of carbohydrates requires less energy than processing of proteins. In fact, the breakdown of carbohydrates requires the least amount of energy, whereas the processing of proteins demands the most energy. In general, the amount of calories ingested and the amount of calories burned determines the overall weight.

To help provide guidelines regarding the types and quantities of food that should be eaten every day, Health Canada has published a simplified "Eat Well Plate" graphic to summarize the recommendations found in Canada's Food Guide (Figure 2). Such representations seek to put the recommended elements of a healthy meal into the context of a place setting of food.

The accompanying website food-guide.canada.ca gives clear recommendations regarding quantity and type of each food that you should consume each day, as well as identifying which foods belong in each category. The guidelines in general suggest you "Make half your plate fruits and vegetables." The other half is grains and protein, with a slightly higher quantity of grains than protein. Dairy products are represented by a drink, but the quantity can be applied to other dairy products as well. All of these foodstuffs contain the energy-containing nutrients carbohydrates, lipids, and proteins in varying amounts, as well as various vitamins, minerals, and essential nutrients. Specifics vary with particular choices within each group, but in general grain products, vegetables and fruit contain higher amounts of carbohydrates than the other groups, whereas meat and dairy products contain higher amounts of protein and lipids.

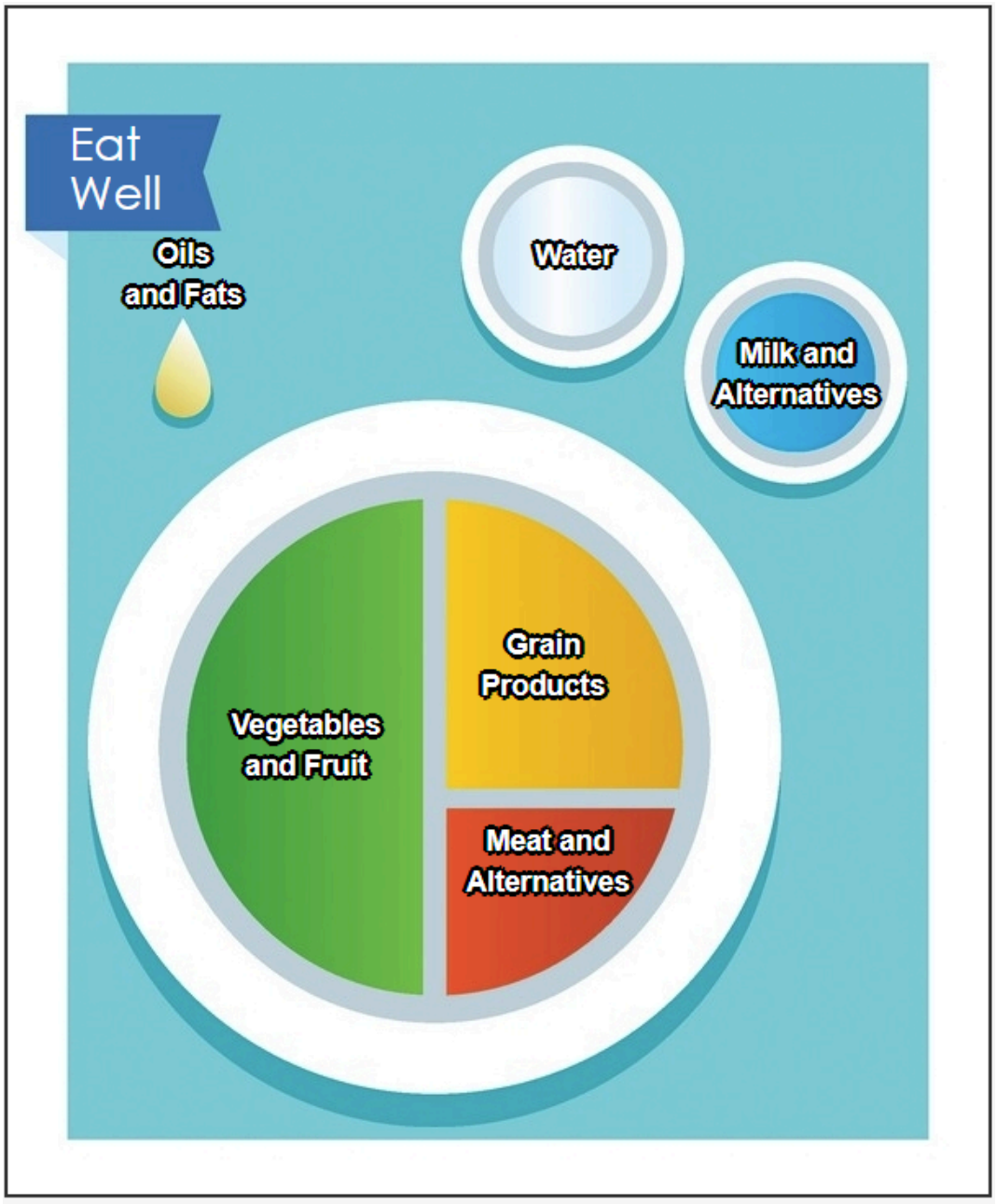


Figure 2. Health Canada's Eat Well Plate. *Health Canada has developed food guidelines to help demonstrate how to maintain a healthy lifestyle.*

Essential nutrients: In addition to providing chemical energy, ingested foodstuffs must also provide any molecules that cannot be produced fast enough (or in some cases, at all) by the body to meet the body's needs. Such molecules are referred to as essential because they must be ingested to allow normal functioning of the human body.

There are two **essential fatty acids** that humans must ingest: linoleic acid (LA), an omega-6 fatty acid, and linolenic acid (ALA), an omega-3 fatty acid. These two fatty acids serve as precursor molecules that can be modified by the body, particularly in the liver, to produce other lipid molecules. However, they cannot be created from other molecules in the human body and so must be provided by consuming an external source.

There are eight **essential amino acids** that humans must ingest from other sources: tryptophan, methionine, valine, threonine, phenylalanine, leucine, isoleucine, and lysine. An additional two – histidine and arginine – are essential for infants but not for adults. Any protein that contains in its primary structure any of these amino

acids will not be made at all in their absence. All of the essential amino acids are found in animal product proteins (e.g. eggs, milk, fish, most meats), but almost no single plant source contains all of the essential amino acids, with the exception of soybean and quinoa. However, combinations of plants can be ingested together to provide them; for example, a combination of cereal grains (e.g. corn) and legumes (e.g. beans) can provide all eight essential amino acids.

Although humans do produce it as a byproduct of cellular respiration, **water** is also an essential nutrient. We lose far more water through constant evaporation from our breath, mucous membranes, and sweat than is produced. Thus humans must ingest water regularly. Plant and animal cells consist largely of water, so a substantial amount of water can be obtained from (non-dehydrated) dietary sources. Nevertheless, humans living in all but the most comfortable of environments typically require access to a source of additional liquid water in addition to plant and animal sources. Excessive water loss (dehydration) can be fatal from a combination of an inability to sweat allowing a dangerous rise in body temperature and a dramatic drop in blood volume and increase in blood viscosity due to water loss from the blood plasma. Under extreme conditions (e.g. exercising strenuously in a hot environment) the lack of a reliable water sources can prove fatal within a few hours; an adult in comfortable surroundings could survive up to about a week without any water intake before succumbing. Generally, the lack of other dietary nutrients in an otherwise health human would not prove fatal nearly as quickly.

The other essential nutrients are the **vitamins** and **minerals**. Vitamins in general must be ingested directly or produced by modifying specific precursor molecules that can be ingested instead, but they are required and cannot be produced from other types of nutrients. Minerals are inorganic ions and as such cannot be 'produced' in the human body at all and must be ingested in an appropriate form.

Vitamins: Vitamins are organic compounds found in foods and are a necessary part of the biochemical reactions in the body. They are involved in a number of processes, including mineral and bone metabolism, and cell and tissue growth, and they act as cofactors for energy metabolism. The B vitamins play the largest role of any vitamins in metabolism (Table 1 and Table 2).

You get most of your vitamins through your diet, although some can be formed from the precursors absorbed during digestion. For example, the body synthesizes vitamin A from the β-carotene in orange vegetables like carrots and sweet potatoes. Vitamins are either fat-soluble or water-soluble. Fat-soluble vitamins A, D, E, and K are absorbed through the intestinal tract with lipids in chylomicrons. Vitamin D is also synthesized in the skin through exposure to sunlight. Because they are carried in lipids, fat-soluble vitamins can accumulate in the lipids stored in the body. If excess vitamins are retained in the lipid stores in the body, hypervitaminosis can result leading to toxic symptoms depending on the vitamin.

Water-soluble vitamins, including the eight B vitamins and vitamin C, are absorbed with water in the gastrointestinal tract. These vitamins move easily through bodily fluids, which are water based, so they are not stored in the body. Excess water-soluble vitamins are excreted in the urine. Therefore, hypervitaminosis of water-soluble vitamins rarely occurs, except with an excess of vitamin supplements.

Table 1: Fat-Soluble Vitamins

Vitamin and alternative name	Sources	Recommended daily allowance	Functions	Problems associated with deficiency
A retinal or β-carotene	Yellow/orange fruits/ vegetables, dark green leafy vegetables, eggs, milk, liver	700-900 µg	Eye & bone development, immune function	Night blindness, epithelial changes, immune system deficiency
D cholecalciferol	Dairy products, egg yolks; synthesis in skin using sunlight	5-15 µg	Aids in calcium & phosphorus absorption, thereby promoting bone growth	Rickets, bone pain, muscle weakness, increased risk of death from cardiovascular disease, cognitive impairment, asthma in children, cancer
E tocopherols	Seeds, nuts, vegetable oils, avocados, wheat germ	15 mg	Antioxidant	Anemia
K phylloquinone	Dark green leafy vegetables, broccoli, Brussels sprouts, cabbage	90-120 µg	Blood clotting, bone health	Hemorrhagic disease of newborn in infants; uncommon in adults

Table 2: Water-Soluble Vitamins

Vitamin and alternative name	Sources	Recommended daily allowance	Functions	Problems associated with deficiency
B_1 thiamine	Whole grains, enriched bread/ cereals, milk, meat	1.1-1.2 mg	Synthesis of pyruvate dehydrogenase for carbohydrate metabolism (pyruvate → acetyl CoA)	Beriberi, Wernicke-Korsikoff syndrome
B_2 riboflavin	Brewer's yeast, almonds, milk, organ meats, legumes, enriched breads/ cereals, broccoli, asparagus	1.1-1.3 mg	Synthesis of FAD for metabolism; production of erythrocytes	Fatigue, slowed growth, digestive problems, light sensitivity, epithelial problems like cracks in the corners of the mouth
B_3 niacin	Meat, fish, poultry, enriched breads/ cereals, peanuts	14-16 mg	Synthesis of NAD^+ for metabolism; nerve function, cholesterol production	Pellagra (cracked, scaly skin; mouth sores; dementia; diarrhea)
B_5 pantothenic acid	Meat, poultry, potatoes, oats, enriched breads/ cereals, tomatoes	5 mg	Synthesis of coenzyme A for metabolism	Rare; fatigue, insomnia, depression, irritability
B_6 pyridoxine	Potatoes, bananas, beans, seeds, nuts, meat, poultry, fish, eggs, dark green leafy vegetables, soy, organ meats	1.3-1.5 mg	Sodium/ potassium balance, erythrocyte synthesis, amino acid metabolism, glycogenolysis and gluconeogenesis, ceramide synthesis	Confusion, irritability, depression, mouth/tongue sores
B_7 biotin	Liver, fruits, meats	30 µg	Cell growth, fatty acid metabolism, blood cell production	Rare in developed countries; dermatitis, hair loss, loss of muscular coordination
B_9 folic acid	Liver, legumes, dark green leafy vegetables, enriched breads/ cereals, citrus fruits	400 µg	DNA/protein synthesis	Poor growth, gingivitis, appetite loss, shortness of breath, gastrointestinal problems, mental deficits

B_{12} cyanobalamin	Fish, meat, poultry, dairy products, eggs	2.4 µg	Fatty acid oxidation, nerve cell function, erythrocyte production	Pernicious anemia leading to nerve cell damage
C ascorbic acid	Citrus fruits, red berries, peppers, tomatoes, broccoli, dark green leafy vegetables	75-90 mg	Collagen production (for formation of connective tissues and teeth, and for wound healing)	Dry hair, gingivitis, bleeding gums, dry/scaly skin, slow wound healing, easy bruising, compromised immunity; can lead to scurvy

Minerals: Minerals in food are inorganic ions or compounds that work with other nutrients to ensure the body functions properly. Minerals cannot be made in the body; they come from the diet. The amount of minerals in the body is small—only 4 percent of the total body mass—and most of that consists of the minerals that the body requires in moderate quantities: potassium, sodium, calcium, phosphorus, magnesium, and chloride.

The most common minerals in the body are calcium and phosphorous, both of which are stored in the skeleton and necessary for the hardening of bones. Most minerals are ionized, and their ionic forms are used in physiological processes throughout the body. Sodium and chloride ions are electrolytes in the blood and extracellular tissues, and iron ions are critical to the formation of hemoglobin. There are additional trace minerals that are still important to the body's functions, but their required quantities are much lower.

Like vitamins, minerals can be consumed in toxic quantities (although it is rare). A healthy diet includes most of the minerals your body requires, so supplements and processed foods can add potentially toxic levels of minerals. Table 3 and Table 4 provide a summary of minerals and their function in the body.

Table 3: Major Minerals

Mineral	Sources	Recommended daily allowance	Functions	Problems associated with deficiency
Potassium (K^+)	Meats, some fish, fruits, vegetables, legumes, dairy products	4700 mg	Nerve & muscle function, electrolyte	Hypokalemia (weaknedd, fatigue, muscle cramping, gastrointestinal problems, cardiac problems)
Sodium (Na^+)	Table salt, milk, beets, celery, processed foods	2300 mg	Blood pressure, blood volume, nerve & muscle function, electrolyte	Rare
Calcium (Ca^{2+})	Dairy products, dark green leafy vegetables, blackstrap molasses, nuts, brewer's yeast, some fish	1000 mg	Bone structure & health; nerve & muscle functions, especially cardiac function, electrolyte	Slow growth, weak and brittle bones
Phosphorus (P, usually as phosphate PO_4^{3-})	Meat, milk	700 mg	Bone formation, metabolism, ATP production	Rare
Magnesium (Mg^{2+})	Whole grains, nuts, leafy green vegetables	310-420 mg	Enzyme activation, ATP production, regulation of other nutrients	Agitation, anxiety, sleep problems, nausea/vomiting, abnormal hearth rhythms, low blood pressure, muscular problems
Chloride (Cl^-)	Most foods; table salt; vegetables, especially seaweed, tomatoes, lettuce, celery, olives	2300 mg	Balance of body fluids, digestion, electrolyte	Loss of appetite, muscle cramps

Table 4: Trace Minerals

Mineral	Sources	Recommended daily allowance	Functions	Problems associated with deficiency
Iron (Fe)	Meat, poultry, fish, shellfish, legumes, nuts, seeds, whole grains, dark leafy green vegetables	8-18 mg	Transport of oxygen in blood, ATP production	Anemia, weakness, fatigue
Zinc (Zn)	Meat, fish, poultry, cheese, shellfish	8-11 mg	Immunity, reproduction, growth, blood clotting, insulin, thyroid function	Loss of appetite, poor growth, weight loss, skin problems, hair loss, vision problems, lack of taste/smell
Copper (Cu)	Seafood, organ meats, nuts, legumes, chocolate, enriched breads/ cereals, some fruits/ vegetables	900 µg	Erythrocyte production, nerve and immune system function, collagen formation, antioxidant	Anemia, low body temperature, bone fractures, low leukocyte count, irregular heartbeat, thyroid problems
Iodine (I)	Fish, shellfish, garlic, lima beans, sesame seeds, soybeans, dark green leafy vegetables	150 µg	Thyroid function	Hypothyroidism (fatigue, weight gain, dry skin, temperature sensitivity); may cause a goiter
Sulfur (S) (as sulfate SO_4^{2-}, or in S-containing amino acids)	Eggs, meat, poultry, fish, legumes	14 mg/kg body weight, as sulfur-containing amino acids	Component of some amino acids required for protein synthesis	Protein deficiency
Fluoride (Fl)	Fluoridated water	3-4 mg	Maintenance of bone and tooth structure	Increased cavities, weak bones and teeth
Manganese (Mn)	Nuts, seeds, whole grains, legumes	1.8-2.3 mg	Formation of connective tissues, blood clotting, sex hormone production, metabolism, brain & nerve function	Infertility, bone malformation, weakness, seizures

Cobalt (Co)	Fish, nuts, leafy green vegetables, whole grains	Not set	Component of B_{12}	None in isolation (same as B_{12} deficiency)
Selenium (Se)	Brewer's yeast, wheat germ, liver, butter, fish, shellfish, whole grains	55 µg	Antioxidant, thyroid function, immune system function	Muscle pain
Chromium (Cr)	Whole grains, lean meats, cheese, black pepper, thyme, brewer's yeast	25-35 µg	Insulin formation	High levels in the blood of glucose, triglycerides, cholesterol
Molybdenum (Mo)	Legumes, whole grains, nuts	45 µg	Cofactor for enzymes	Rare

Unit 7: Cellular Respiration and Energy Metabolism

Unit outline

Part 1: Carbohydrate Metabolism

- Glycolysis
- Krebs Cycle
- Oxidative Phosphorylation
- Gluconeogenesis

Part 2: Lipid Metabolism

- Lipolysis
- Ketogenesis
- Ketone Body Oxidation
- Lipogenesis

Part 3: Protein Metabolism

Part 4: Metabolic States of the Body

- The Absorptive State
- The Postabsorptive State
- Starvation

Learning Objectives

At the end of this unit, you should be able to:

I. Describe the process of cellular respiration in general terms.

II. Describe the roles of ATP, NAD, and FAD in energy metabolism in the cell.

III. Describe the process of glycolysis.

IV. Describe the formation of acetyl coenzyme A from pyruvic acid.

V. Explain the role of the Krebs cycle in cellular respiration.

VI. Describe the role of the electron transport chain in cellular respiration.

VII. Describe the major steps in the generation of ATP by chemiosmosis.

VIII. Summarize the ATP produced from the breakdown of a single glucose molecule.

IX. Describe the importance of oxygen (O_2) in cellular respiration and compare aerobic respiration with lactic acid fermentation.

X. Describe the importance of carbohydrates, lipids and proteins in energy storage and energy availability, and their use during starvation conditions.

XI. Describe the importance of glucose in cellular respiration and ATP production.

XII. Describe the role of lipids and amino acids in ATP production.

XIII. Describe the role of ketone bodies in energy metabolism.

XIV. Describe the relationship between gluconeogenesis, lipid metabolism, and protein catabolism.

XV. Describe the fate of amino acids that are metabolized for ATP production.

XVI. Explain the importance of appropriate nutrient intake for maintaining homeostasis of the body.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Describe the process of cellular respiration in general terms.

1. Define the term “cellular respiration”.
2. What is the main biological function of cellular respiration?
3. Determine and write out the overall chemical equation for aerobic cellular respiration.

II. Describe the roles of ATP, NAD, and FAD in energy metabolism in the cell.

1. Use complete sentences to describe how cells produce:
    - ATP
    - NADH
    - $FADH_2$
2. Use complete sentences to describe the biological purpose of a cell producing:
    - ATP
    - NADH
    - $FADH_2$

III. Describe the process of glycolysis.

IV. Describe the formation of acetyl coenzyme A from pyruvic acid.

V. Explain the role of the Krebs cycle in cellular respiration.

VI. Describe the role of the electron transport chain in cellular respiration.

VII. Describe the major steps in the generation of ATP by chemiosmosis.

1. Write a single-sentence summary of the chemical events that occur during each of the following processes:
    - Glycolysis
    - Pyruvic acid oxidation
    - The Krebs (citric acid) cycle
    - The electron transport chain
    - Substrate-level phosphorylation
    - Oxidative phosphorylation
2. Specify the molecules that are required, consumed, and produced during each of the following processes:
    - Glycolysis
    - Pyruvic acid oxidation
    - The Krebs (citric acid) cycle
    - The electron transport chain
3. Starting with the arrival of NADH and $FADH_2$ at the electron transport chain, thoroughly describe how the electron transport chain is used to generate ATP.

VIII. Summarize the ATP produced from the breakdown of a single glucose molecule.

1. At which point(s) during aerobic cellular respiration of one glucose molecule are ATP molecules produced by each of the following processes, and how many ATP molecules are produced by each process?
    - Substrate-level phosphorylation
    - Oxidative phosphorylation

IX. Describe the importance of oxygen (O_2) in cellular respiration and compare aerobic respiration and lactic acid fermentation.

1. For what single main function is oxygen required during cellular respiration?
2. In the absence of oxygen, how many molecules of ATP can be produced from a single glucose molecule?
3. Explain why, in the absence of oxygen, the continued generation of ATP from glucose requires the conversion of pyruvic acid to lactic acid.

X. Describe the importance of carbohydrates, lipids and proteins in energy storage and energy availability, and their use during starvation conditions.

1. Describe and explain the use of carbohydrates, lipids, and proteins for ATP production when in:
    - An absorptive (fed) state.
    - A postabsorptive (fasting) state.
    - Starvation conditions.

2. Protein molecules contain approximately the same amount of energy per gram as carbohydrates and are found extensively throughout the human body. Explain why it is physiologically important that proteins are used as major sources of chemical energy only after other energy-containing molecules (*i.e.*, carbohydrates and lipids) have been depleted.

XI. Describe the importance of glucose in cellular respiration and ATP production.

1. Which specific nutrient molecule are all human body cells normally capable of breaking down to generate ATP?

XII. Describe the role of lipids and amino acids in ATP production.

1. What other nutrient molecules are at least some human body cells capable of breaking down to generate ATP? For each of these nutrient molecules, which body cell types can (or cannot) break it down?

XIII. Describe the role of ketone bodies in energy metabolism.

1. What types of molecules can be used to produce ketone bodies?
2. Under what conditions should ketone bodies be produced?
3. What function do ketone bodies serve in the human body?

XIV. Describe the relationship between gluconeogenesis, lipid metabolism, and protein catabolism.

XV. Describe the fate of amino acids that are metabolized for ATP production.

1. Name and describe with a one-sentence summary the mechanism(s) that are used to allow body cells to continue generating ATP in the event that:
    - Blood glucose levels decline
    - Glycogen stores in the body decline
    - Lipid stores in the body decline
    - Oxygen is unavailable
2. Explain the functional reason why, under conditions of low oxygen availability, lactic acid (or lactate) must be produced to allow glycolysis to continue.
3. Clearly define each of the following terms:
    - Glycolysis
    - Glycogenesis
    - Gluconeogenesis
    - Glycogenolysis
4. Describe the process in the human body by which some of the energy present in lipid molecules can be used to generate ATP.
    - In which organ(s) and/or cell type(s) can this process occur?
    - Which major steps are involved?
    - Can any of the intermediate molecules be transported to other tissues in a form that will allow the receiving tissues to generate ATP in the absence of glucose?
    - At what stage(s) of cellular respiration can the breakdown products of lipid molecules be

used?

5. Describe the process in the human body by which some of the energy present in amino acids can be used to generate ATP.
    - In which organ(s) and/or cell type(s) can this process occur?
    - Which major step are involved?
    - At what stage(s) of cellular respiration can the breakdown products of amino acids be used?
    - In breaking down amino acids, what potentially toxic chemical is produced that is not produced when a lipid or carbohydrate is broken down? What is the fate of this product?
    - What are the potentially detrimental physiological consequences of breaking down amino acids, rather than glucose, to produce ATP?

XVI. Explain the importance of appropriate nutrient intake for maintaining homeostasis of the body.

1. List the classes of nutrients that can be broken down to release energy that can be used to produce ATP.
2. For each of the following chemicals, describe its function in metabolism and name the specific nutrient(s) that must be ingested to produce it:
    - Pyruvate dehydrogenase
    - Nicotinamide adenine dinucleotide (NAD^+)
    - Flavin adenine dinucleotide (FAD)
    - Coenzyme A

Part 1: Carbohydrate Metabolism

Carbohydrates are organic molecules composed of carbon, hydrogen, and oxygen atoms. The family of carbohydrates includes both sugars (i.e. monosaccharides and disaccharides) and polysaccharides. Glucose and fructose are examples of sugars, and starch, glycogen, and cellulose are all examples of polysaccharides. Polysaccharides are made of multiple monosaccharide molecules. Polysaccharides serve as energy storage (e.g., starch and glycogen) and as structural components (e.g., chitin in insects and cellulose in plants).

During digestion, carbohydrates are broken down into simple, soluble sugars that can be transported across the intestinal wall into the circulatory system to be transported throughout the body. Carbohydrate digestion begins in the mouth with the action of salivary amylase on starches and ends with monosaccharides being absorbed across the epithelium of the small intestine. Once the absorbed monosaccharides are transported to the tissues, the process of **cellular respiration** begins (Figure 1). This section will focus first on glycolysis, a process where the monosaccharide glucose is oxidized, releasing the energy stored in its bonds to produce ATP.

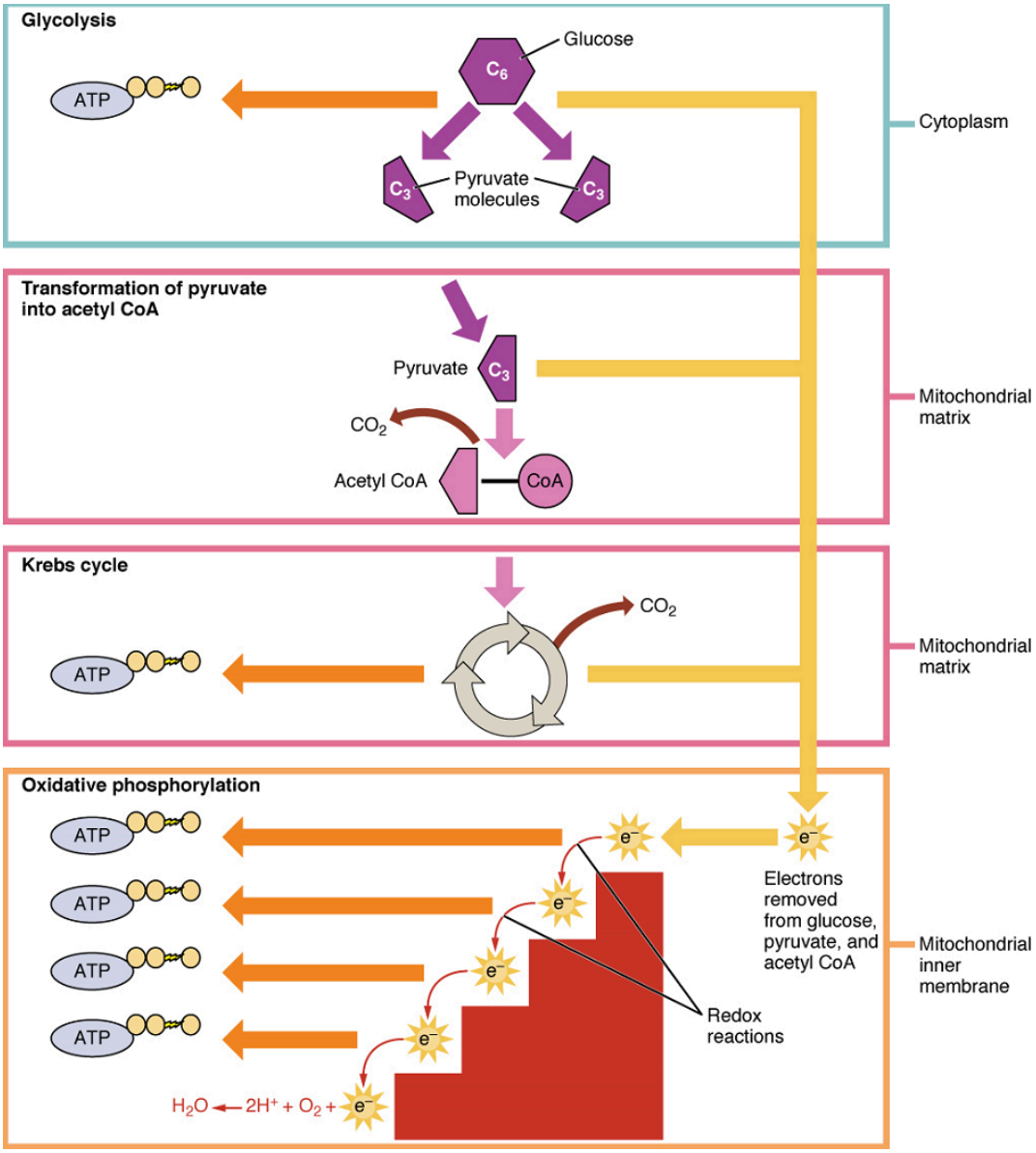


Figure 1. Cellular Respiration. *Cellular respiration oxidizes glucose molecules through glycolysis, the Krebs cycle, and oxidative phosphorylation to produce ATP.*

Glycolysis: Glucose is the body's most readily available source of energy. After digestive processes break polysaccharides down into monosaccharides, including glucose, the monosaccharides are transported across the wall of the small intestine and into the circulatory system, which transports them to the liver. In the liver, hepatocytes either pass the glucose on through the circulatory system or store excess glucose as glycogen. Cells in the body take up the circulating glucose in response to insulin and, through a series of reactions called **glycolysis**, transfer some of the energy in glucose to a new bond between adenosine diphosphate (ADP) and a third phosphate group to form adenosine triphosphate (ATP) (Figure 2). The last step in glycolysis produces the product **pyruvate**.

Glycolysis can be expressed as the following equation:

$$\text{Glucose} + 2\text{ATP} + 2\text{NAD}^+ + 4\text{ADP} + 2\text{P}_i \rightarrow 2\ \text{Pyruvate} + 4\text{ATP} + 2\text{NADH} + 2\text{H}^+$$

This equation states that glucose – in combination with ATP (a source of chemical energy), nicotinamide adenine dinucleotide (NAD^+, a coenzyme that serves as an electron acceptor), and inorganic phosphate – breaks down into two pyruvate molecules, generating four ATP molecules – for a net yield of two ATP – and two

energy-containing NADH coenzyme molecules (resulting from adding a hydrogen atom and an extra electron to NAD^+). The NADH that is produced in this process will be used later to produce ATP in the mitochondria. Importantly, by the end of this process, one glucose molecule generates two pyruvate molecules, two high-energy ATP molecules, and two electron-carrying NADH molecules.

Glycolysis can be divided into two phases: energy consuming (also called chemical priming) and energy yielding. The first phase is the **energy-consuming phase**, so it requires two ATP molecules to start the reaction for each molecule of glucose. At the end of this phase, the six-carbon sugar is split to form two phosphorylated three-carbon sugars, glyceraldehyde-3-phosphate (G3P) and dihydroxyacetone phosphate (DHAP). DHAP is then converted into glyceraldehyde-3-phosphate.

The second phase of glycolysis, the **energy-yielding phase**, harvests the energy contained in G3P, which is further phosphorylated and oxidized. During this step an electron is released that is then picked up by NAD^+ to create an NADH molecule. NADH is a high-energy molecule, like ATP, but unlike ATP, it is not used as energy currency by the cell. Because there are two glyceraldehyde-3-phosphate molecules, two NADH molecules are synthesized during this step. In a series of reactions leading to pyruvate, the two phosphate groups are then transferred from the molecule to which they are attached to two ADPs to form two ATPs by the process of **substrate-level phosphorylation** (direct phosphorylation). Thus, glycolysis uses two ATPs but generates four ATPs, yielding a net gain of two ATPs and two molecules of pyruvate. In the presence of oxygen, pyruvate continues on to the **Krebs cycle** (also called the **citric acid cycle** or **tricarboxylic acid cycle (TCA)**, where additional energy is extracted and passed on, converted into lactic acid by **fermentation**; or used later for the synthesis of glucose through **gluconeogenesis**.

Anaerobic Conditions: When oxygen (O_2) is limited or absent, pyruvate enters an anaerobic pathway. In these reactions, pyruvate can be converted into lactic acid. This pathway serves to oxidize NADH into the NAD+ needed by glycolysis. In this reaction, pyruvate replaces oxygen as the final electron acceptor. It accepts the electrons from the NADH produced from glycolysis, regenerating NAD^+, and is reduced to form lactic acid. This lactic acid fermentation occurs in most cells of the body when oxygen is limited or mitochondria are absent or nonfunctional. For example, because erythrocytes (red blood cells) lack mitochondria, they must produce their ATP from lactic acid fermentation. This is an effective pathway of ATP production for short periods of time, ranging from seconds to a few minutes. The lactic acid produced diffuses into the plasma and is carried to the liver, where it is converted back into pyruvate or glucose. Similarly, when a person exercises, muscles use ATP faster than oxygen can be delivered to them. They depend on glycolysis and lactic acid production for rapid ATP production.

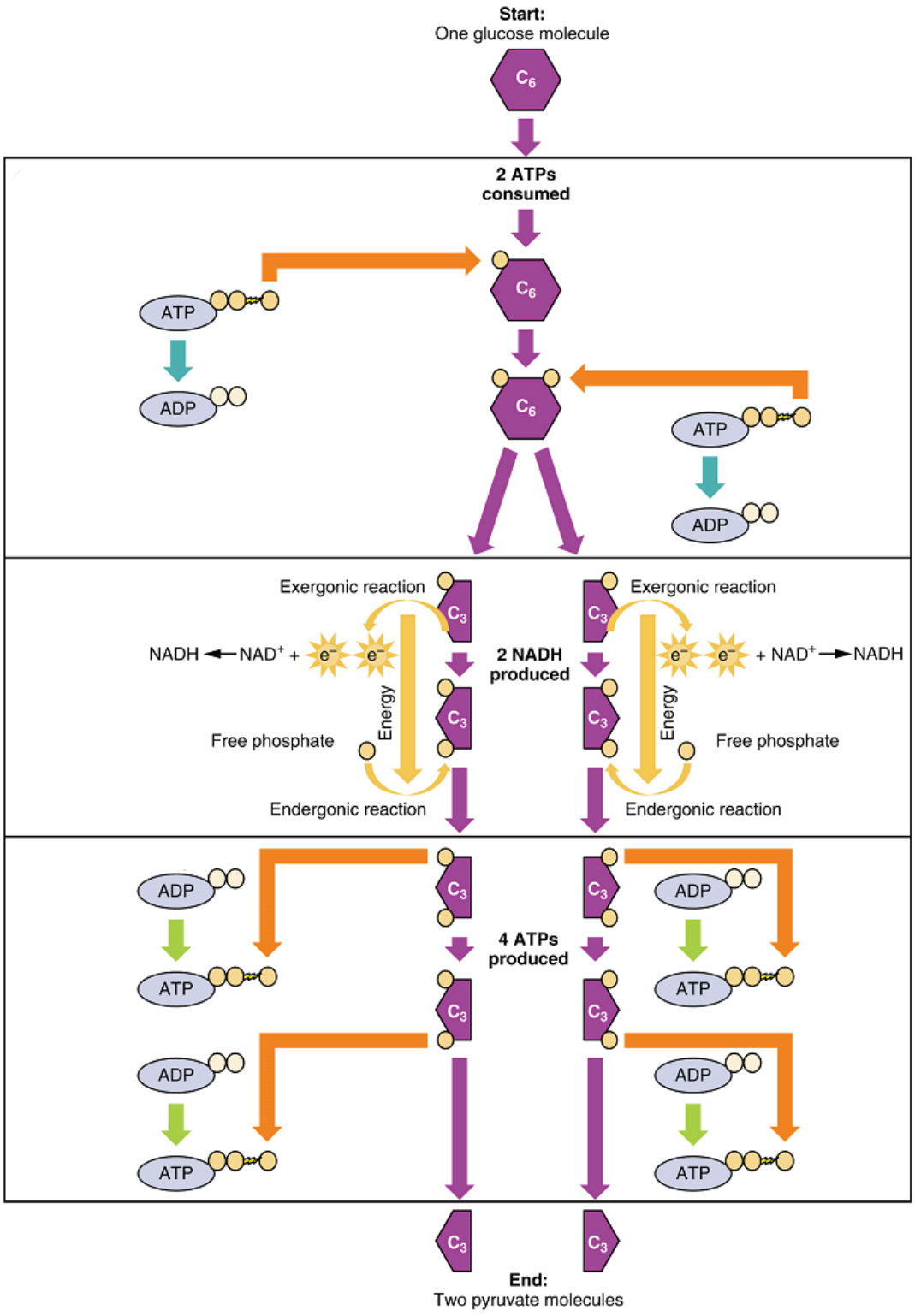


Figure 2. Glycolysis Overview. *During the energy-consuming phase of glycolysis, two ATPs are consumed, transferring two phosphates to the glucose molecule. The glucose molecule then splits into two three-carbon compounds, each containing a phosphate. During the second phase, an additional phosphate is added to each of the three-carbon compounds. The energy for this endergonic reaction is provided by the removal of two electrons (oxidation) from each three-carbon compound. During the energy-yielding phase, the phosphates are removed from both three-carbon compounds and used to produce four ATP molecules.*

Aerobic Respiration: In the presence of oxygen, pyruvate can enter the Krebs cycle where additional energy is extracted as electrons are transferred from the pyruvate to the acceptors NAD+ and flavin adenine dinucleotide (FAD), with carbon dioxide released as a waste product (Figure 3). The NADH and $FADH_2$ (resulting from the addition of two hydrogen atoms to FAD) pass electrons on to the electron transport chain, which uses the

transferred energy to produce ATP by oxidative phosphorylation. As the last step in the electron transport chain, oxygen is the terminal electron acceptor, combining with electrons and hydrogen ions to produce water inside the mitochondria.

Krebs Cycle (Citric Acid Cycle or Tricarboxylic Acid Cycle): The pyruvate molecules generated during glycolysis are transported across the mitochondrial membrane into the inner mitochondrial matrix, where they are metabolized by enzymes in a pathway called the Krebs cycle (Figure 4). The Krebs cycle is also commonly called the citric acid cycle or the tricarboxylic acid (TCA) cycle. During the Krebs cycle, high-energy molecules, including ATP, NADH, and $FADH_2$, are created. NADH and $FADH_2$ then pass electrons through the electron transport chain in the mitochondria to generate more ATP molecules.

The three-carbon pyruvate molecule generated during glycolysis moves from the cytoplasm into the mitochondrial matrix, where it is converted into a two-carbon acetyl group and bound to coenzyme A to form an **acetyl coenzyme A (acetyl CoA)** molecule. This reaction is an oxidative decarboxylation that releases carbon dioxide and transfers two electrons to NAD+ to form NADH. Acetyl CoA enters the Krebs cycle by combining with a four-carbon molecule, oxaloacetate, to form the six-carbon molecule citrate, or citric acid, at the same time releasing the coenzyme A molecule.

The six-carbon citrate molecule is then converted to a five-carbon molecule and then a four-carbon molecule, ending with oxaloacetate, the beginning of the cycle. Along the way, each citrate molecule will produce one ATP, one $FADH_2$, and three NADH. The $FADH_2$ and NADH will enter the oxidative phosphorylation system located in the inner mitochondrial membrane. In addition, the Krebs cycle supplies the starting materials to process and break down proteins and fats.

Oxidative Phosphorylation: Oxidative phosphorylation is made up of two closely tied components, the electron transport chain and chemiosmosis. The **electron transport chain (ETC)** uses the NADH and $FADH_2$ produced by the Krebs cycle to generate a proton gradient. Electrons from NADH and $FADH_2$ are transferred through protein complexes embedded in the inner mitochondrial membrane by a series of enzymatic reactions. The electron transport chain consists of a series of four enzyme complexes (Complex I – Complex IV) and two mobile electron shuttles (ubiquinone and Cytochrome c), which act as electron carriers and proton pumps used to transfer H^+ ions into the space between the inner and outer mitochondrial membranes (Figure 5). The ETC couples the transfer of electrons between a donor (like NADH) and an electron acceptor (O_2) with the transfer of protons (H^+ ions) across the inner mitochondrial membrane. In the presence of oxygen, energy is passed, stepwise, through the electron carriers to collect gradually the energy needed to attach a phosphate to ADP and produce ATP. The role of molecular oxygen, O_2, is as the terminal electron acceptor for the ETC. This means that once the electrons have passed through the entire ETC, they must be passed to another, separate molecule. These electrons, O_2, and H^+ ions from the matrix combine to form new water molecules. This is the basis for your need to breathe in oxygen. Without oxygen, electron flow through the ETC ceases.

The electrons released from NADH and $FADH_2$ are passed along the chain by each of the carriers, which are reduced when they receive the electron and oxidized when passing it on to the next carrier. Each of these reactions releases a small amount of energy, which is used to pump H+ ions across the inner membrane. The accumulation of these protons in the space between the membranes creates a proton gradient with respect to the mitochondrial matrix.

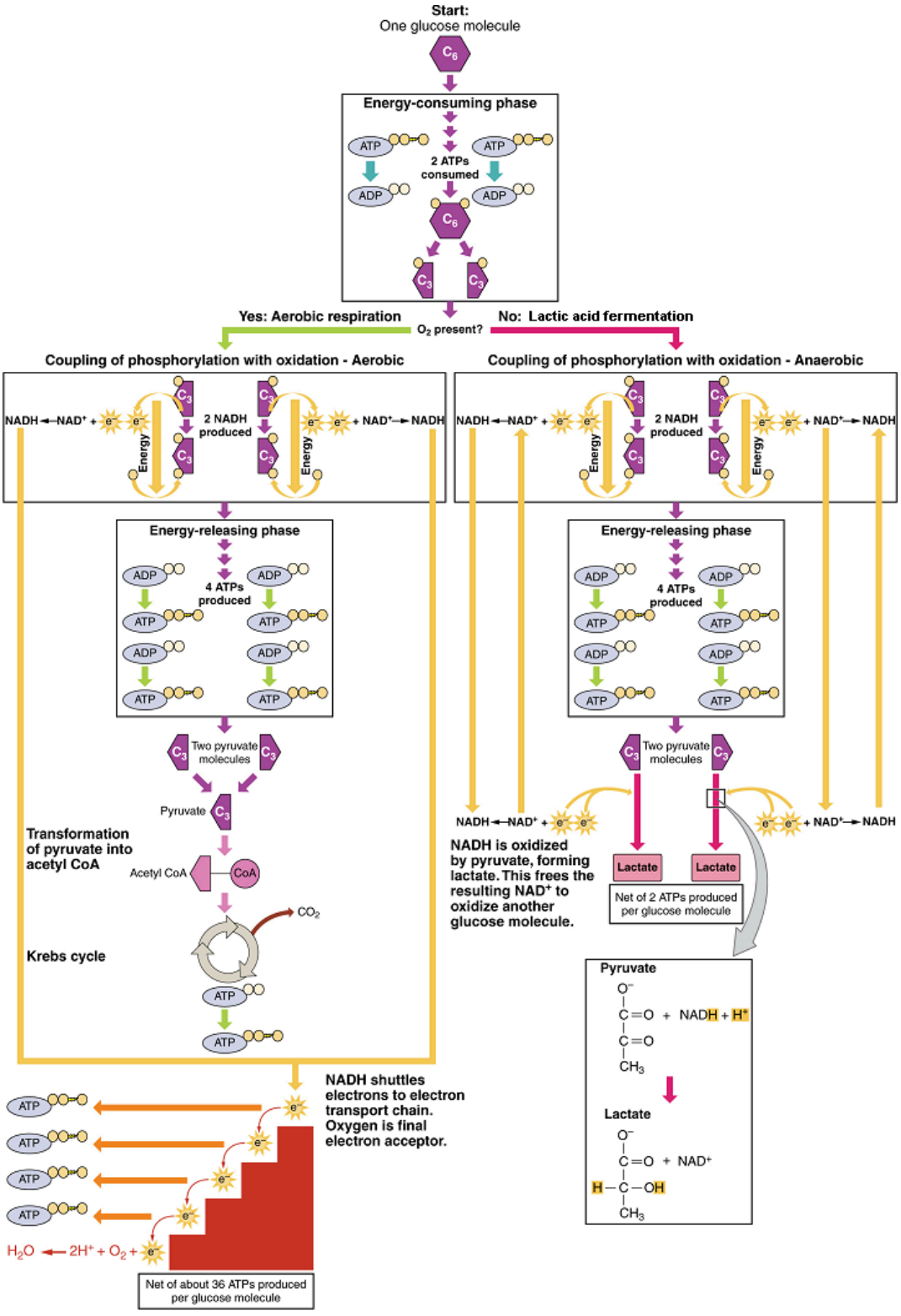


Figure 3. Aerobic Respiration Versus Lactic Acid Production. *The process of lactic acid fermentation converts glucose into two lactate molecules in the absence of oxygen or within erythrocytes that lack mitochondria. During aerobic respiration, glucose is oxidized into two pyruvate molecules.*

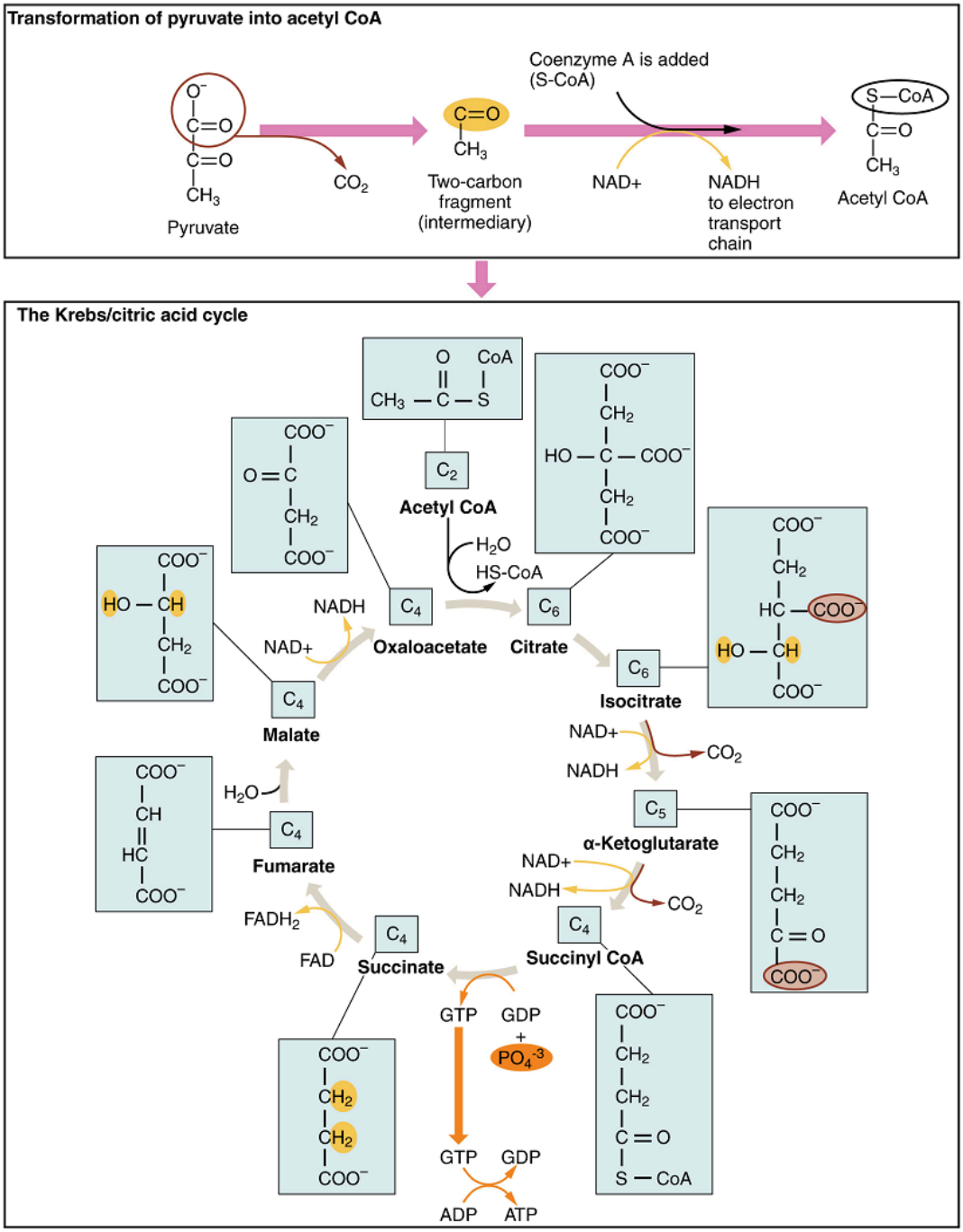


Figure 4. Krebs Cycle. *During the Krebs cycle, each pyruvate that is generated by glycolysis is converted into a two-carbon acetyl CoA molecule. The acetyl CoA is systematically processed through the cycle and produces high-energy NADH, $FADH_2$, and ATP molecules. (Not all material in this figure is examinable.)*

In chemiosmosis, the energy stored in the proton gradient generated by the electron transport chain is used to generate ATP. Embedded in the inner mitochondrial membrane is an amazing protein pore complex called ATP synthase. Effectively, it is a turbine that is powered by the flow of H^+ ions across the inner membrane down a gradient and into the mitochondrial matrix. As the H^+ ions traverse the complex, the shaft of the complex rotates. This rotation enables other portions of ATP synthase to encourage ADP and Pi to create ATP.

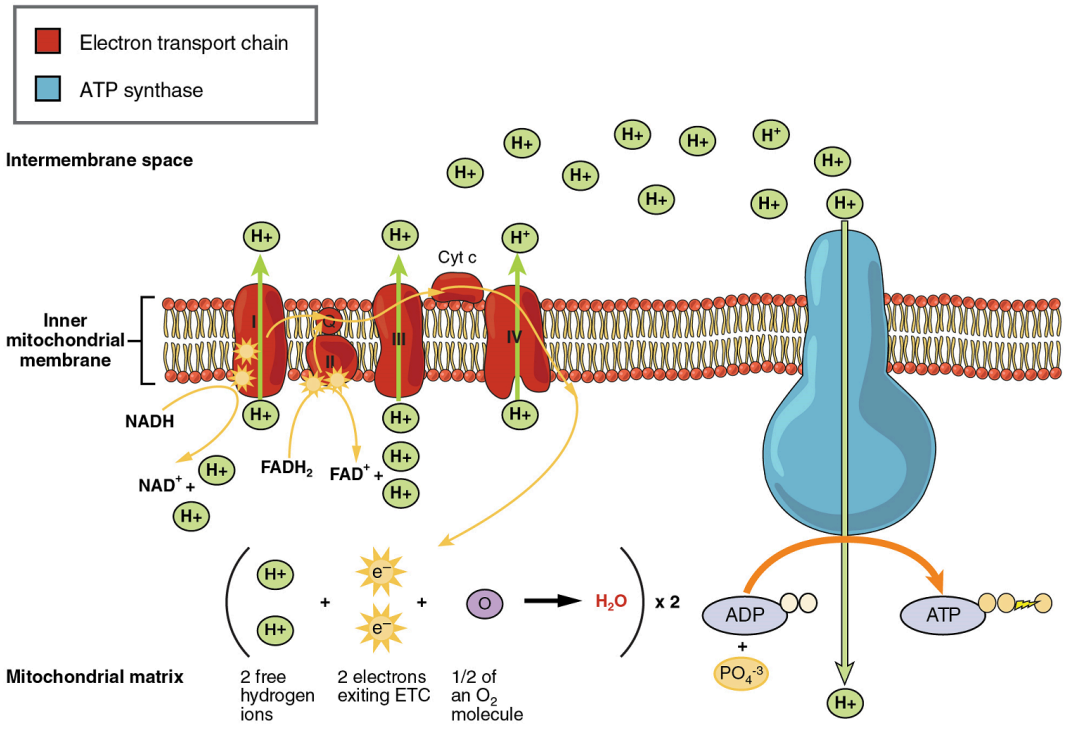


Figure 5. Oxidative Phosphorylation. *The electron transport chain is a series of electron carriers and ion pumps that are used to pump H^+ ions out of the inner mitochondrial matrix. The resulting proton gradient then drives ATP production by ATP synthase.*

In accounting for the total number of ATP produced per glucose molecule through aerobic respiration, it is important to remember the following points:

A net of two ATP are produced through glycolysis (four produced and two consumed during the energy-consuming stage).

In all phases after glycolysis, the number of ATP, NADH, and $FADH_2$ produced must be multiplied by two to reflect how each glucose molecule produces two pyruvate molecules.

In the ETC, about 2.5 ATP are produced for every oxidized NADH. However, only about 1.5 ATP are produced for every oxidized $FADH_2$. The electrons from $FADH_2$ produce less ATP, because they start at a lower point in the ETC (Complex II) compared to the electrons from NADH (Complex I) (see Figure 5)

Therefore, for every glucose molecule that enters aerobic respiration, a possible net total of 32 ATPs are produced (Figure 6). This total represents the maximum potential ATP production per glucose molecule from aerobic cellular respiration.

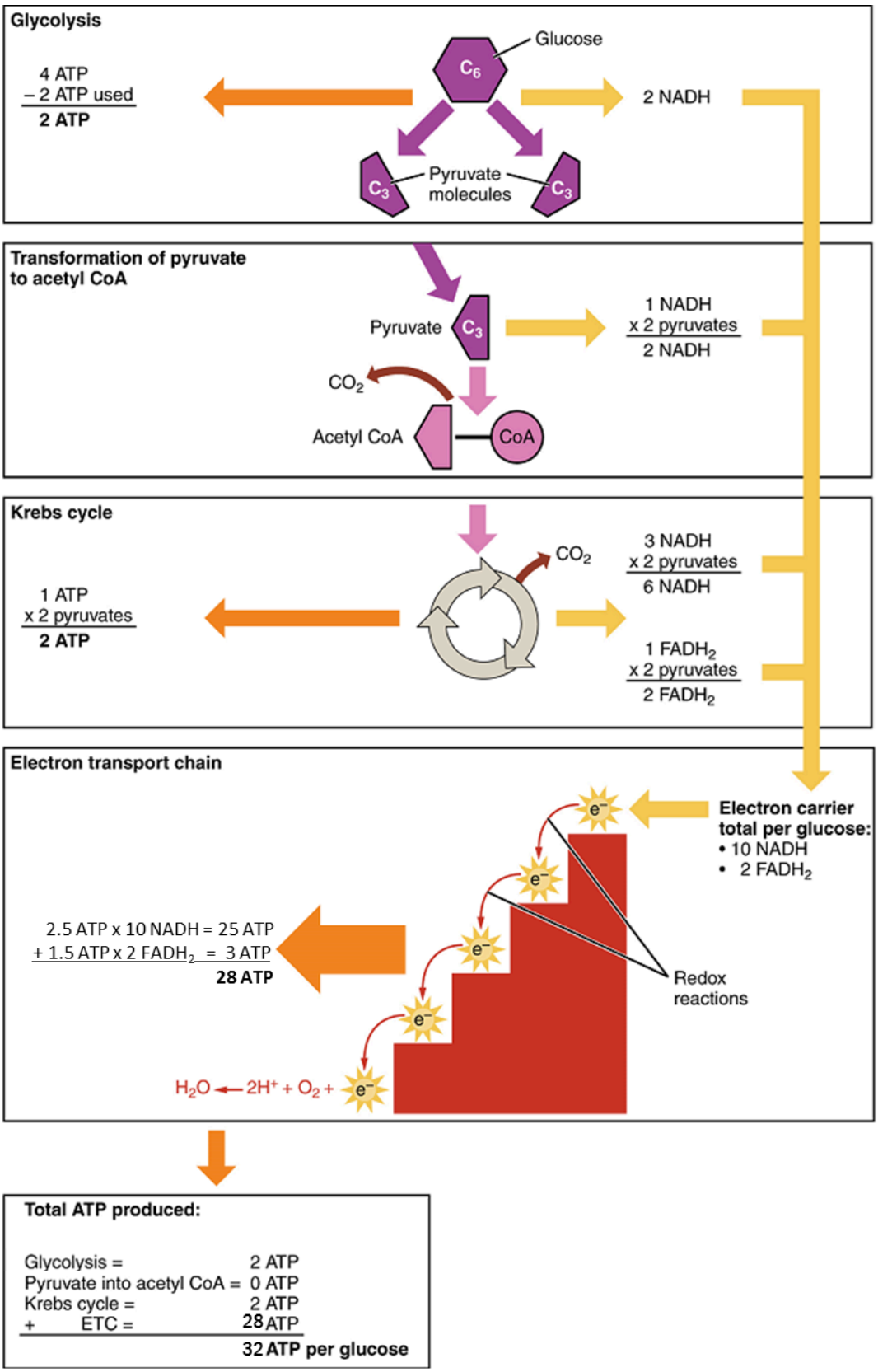


Figure 6. Carbohydrate Metabolism. *Carbohydrate metabolism involves glycolysis, the Krebs cycle, and the electron transport chain.*

Gluconeogenesis: Gluconeogenesis is the synthesis of new glucose molecules from pyruvate, lactate, glycerol,

or some amino acids. This process takes place primarily in the liver during periods of low glucose, that is, under conditions of fasting, starvation, and low carbohydrate diets. So, the question can be raised as to why the body would create something it has just spent a fair amount of effort to break down? Certain key organs, including the brain, can use only glucose as an energy source; therefore, it is essential that the body maintain a minimum blood glucose concentration. When the blood glucose concentration falls below that certain point, new glucose is synthesized by the liver to raise the blood concentration to normal.

As will be discussed as part of lipolysis, fats can be broken down into glycerol, which can be phosphorylated to form dihydroxyacetone phosphate or DHAP. DHAP can either enter the glycolytic pathway or be used by the liver as a substrate for gluconeogenesis.

Part 2: Lipid Metabolism

Fats (or triglycerides) within the body are ingested as food or synthesized by adipocytes or hepatocytes from carbohydrate precursors (Figure 8). Lipid metabolism entails the oxidation of fatty acids to either generate energy or synthesize new lipids from smaller constituent molecules. Lipid metabolism is associated with carbohydrate metabolism, as products of glucose (such as acetyl CoA) can be converted into lipids.

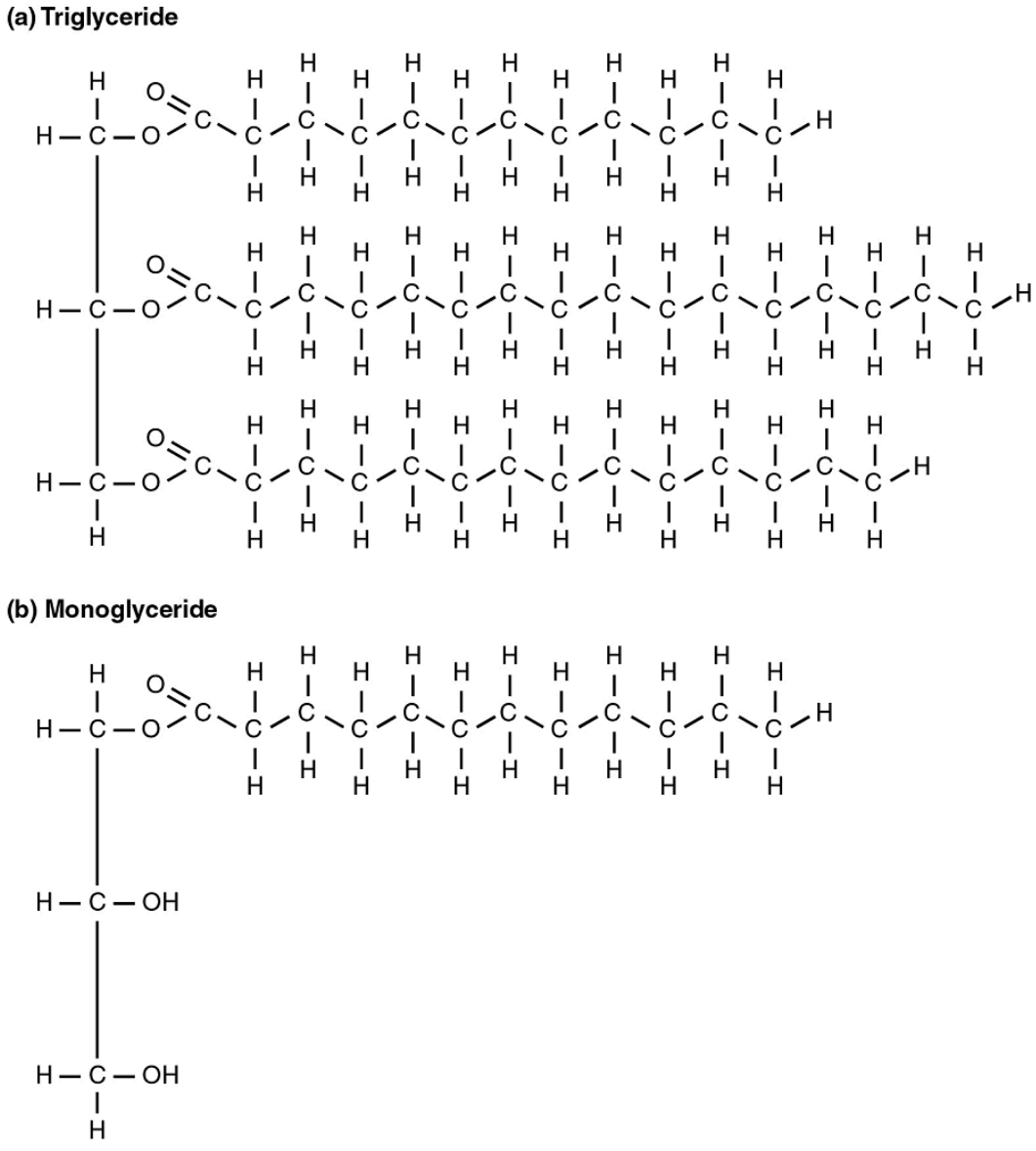


Figure 8. Triglyceride Broken Down into a Monoglyceride. *A triglyceride molecule (a) breaks down into a monoglyceride (b).*

Lipolysis: To obtain energy from fat, triglycerides must first be broken down by hydrolysis into their two principal components, fatty acids and glycerol. This process, called **lipolysis**, takes place in the cytoplasm of adipocytes. Subsequently, the fatty acids and glycerol are released into the bloodstream, to be taken up by tissues such as the muscle, heart and liver. The resulting fatty acids are oxidized by β-oxidation into acetyl CoA, which is used by the Krebs cycle. The glycerol that is released from triglycerides after lipolysis directly enters the glycolysis pathway as DHAP. Because one triglyceride molecule yields three fatty acid molecules with as much as 16 or more carbons in each one, fat molecules yield more energy than carbohydrates and are an important source of energy for the human body. Triglycerides yield more than twice the energy per unit mass when compared to carbohydrates and proteins. Therefore, when glucose levels are low, triglycerides can be converted into acetyl CoA molecules and used to generate ATP through aerobic respiration.

The breakdown of fatty acids begins in the cytoplasm, where fatty acids are converted into fatty acyl CoA molecules. This fatty acyl CoA is transported to the mitochondrial matrix, where it is broken down and oxidized to acetyl CoA in a process called **fatty acid oxidation** or **beta (β)-oxidation** (Figure 10). The newly formed acetyl CoA enters the Krebs cycle and is used to produce ATP in the same way as acetyl CoA derived from pyruvate.

Ketogenesis: If excessive acetyl CoA is created from the oxidation of fatty acids and the Krebs cycle is overloaded and cannot handle it, the acetyl CoA, in the liver, is diverted to create **ketone bodies** (Fig. 11).

Two of these ketone bodies (β-hydroxybutyrate and acetoacetate, and their acid forms β-hydroxybutyric acid and acetoacetatic acid) can serve as a fuel source if glucose levels are too low in the body. Ketone bodies serve as fuel in times of prolonged starvation or when patients suffer from uncontrolled diabetes and cannot utilize most of the circulating glucose. The third ketone body, acetone, is removed by exhalation. One symptom of ketogenesis is that the patient's breath smells sweet like alcohol. This effect provides one way of telling if a diabetic is properly controlling the disease.

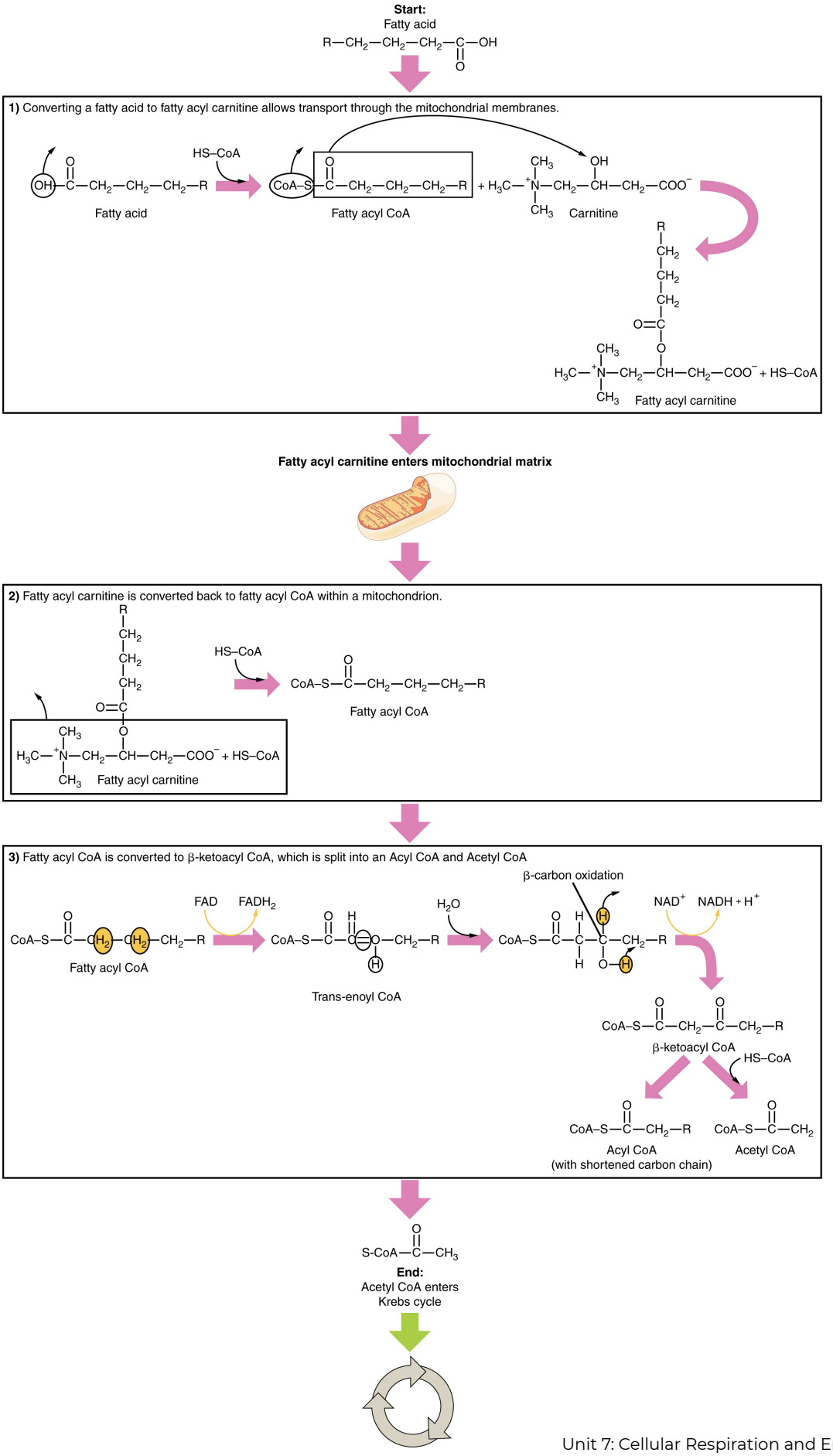


Figure 10. Breakdown of Fatty Acids. *During fatty acid oxidation, triglycerides can be broken down into acetyl CoA molecules and used for energy when glucose levels are low.*

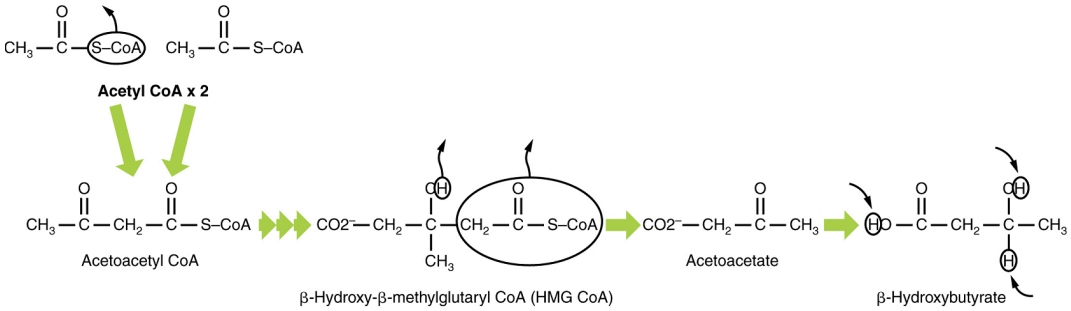


Figure 11. Ketogenesis. *Excess acetyl CoA is diverted from the Krebs cycle to the ketogenesis pathway. This reaction occurs in the mitochondria of liver cells. The result is the production of β-hydroxybutyrate, the primary ketone body found in the blood.*

Ketone Body Oxidation: Organs that have classically been thought to be dependent solely on glucose, such as the brain, can actually use ketone bodies as an alternative energy source. This keeps the brain and other organs, such as the heart, functioning when glucose is limited. Since both β-hydroxybutyric acid and acetoacetatic acid are acids, their presence in blood, can cause acidosis (ketoacidosis), a dangerous condition in diabetics.

In these organs, ketone bodies are converted to two acetyl CoA molecules each. These acetyl CoA molecules are then processed through the Krebs cycle to generate energy (Figure 12).

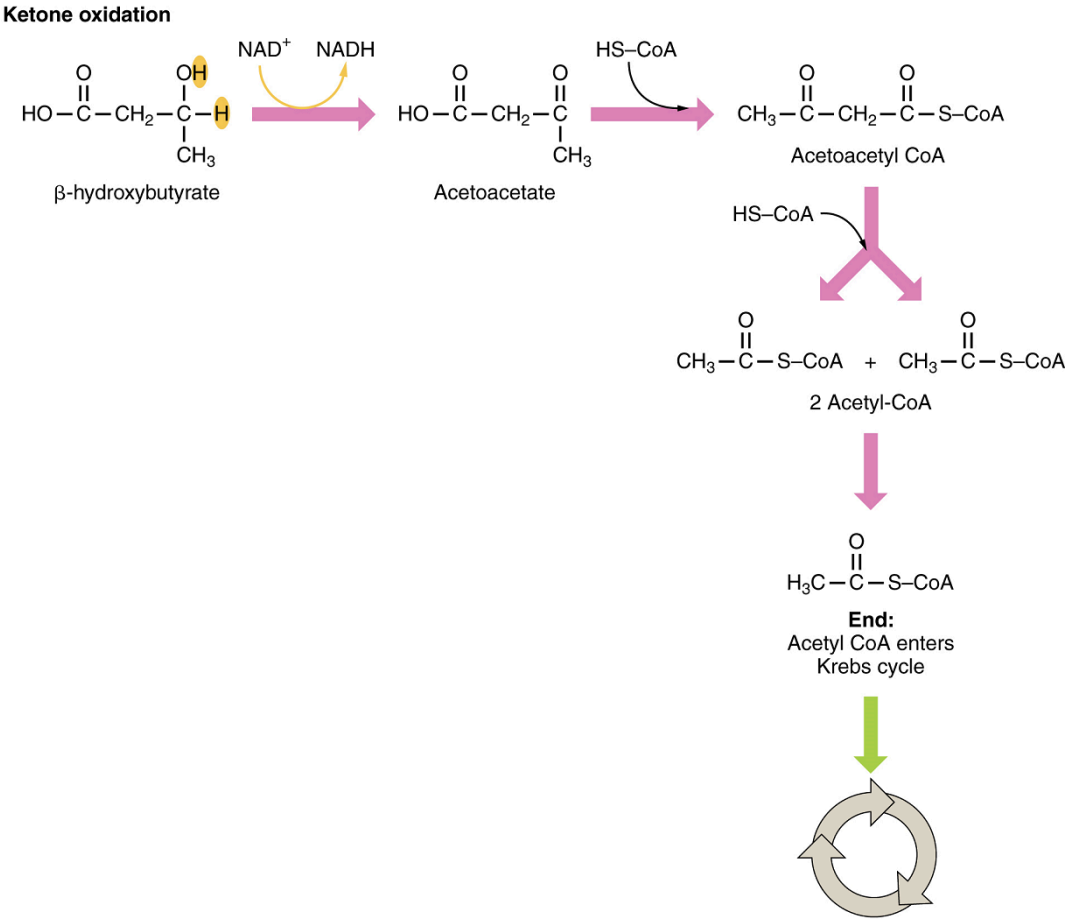


Figure 12. Ketone Oxidation. *When glucose is limited, ketone bodies can be oxidized to produce acetyl CoA to be used in the Krebs cycle to generate energy.*

Lipogenesis: When glucose levels are plentiful, the excess acetyl CoA generated by glycolysis and pyruvate oxidation can be converted into fatty acids, triglycerides, cholesterol, steroids, and bile salts. This process, called **lipogenesis**, creates lipids (fat) from the acetyl CoA and takes place in the cytoplasm of adipocytes (fat cells) and hepatocytes (liver cells) (Figure 13). When you eat more glucose or carbohydrates than your body needs, acetyl CoA is turned into fat. Although there are several metabolic sources of acetyl CoA, it is most commonly derived from glycolysis. Acetyl CoA availability is significant, because it initiates lipogenesis. Lipogenesis begins with acetyl CoA and advances by the subsequent addition of two carbon atoms from another acetyl CoA; this process is repeated until fatty acids are the appropriate length. Because this is a bond-creating anabolic process, ATP is consumed. However, the creation of triglycerides and lipids is an efficient way of storing the energy available in carbohydrates. Triglycerides and lipids, high-energy molecules, are stored in adipose tissue until they are needed.

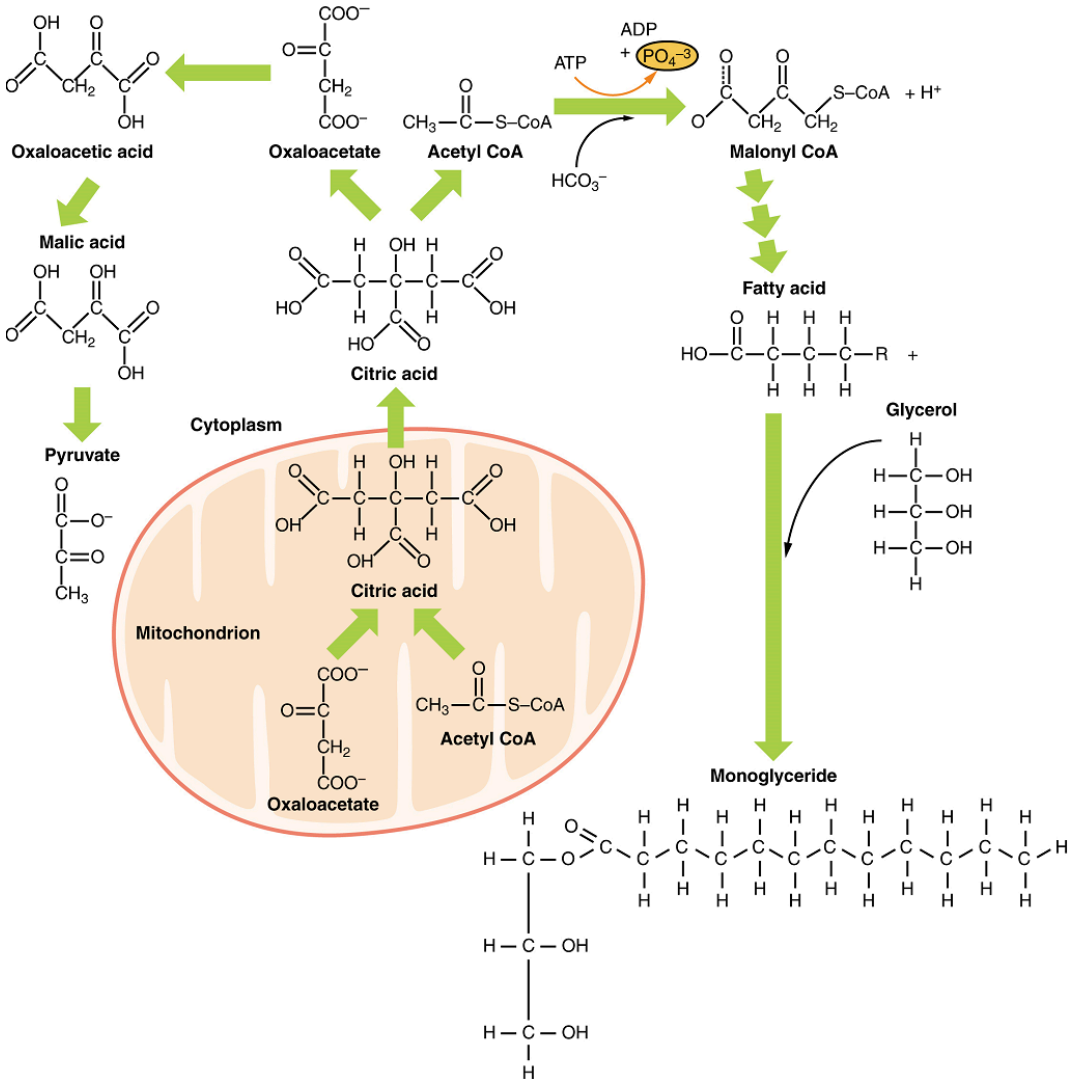


Figure 13. Lipid Metabolism. *Lipids may follow one of several pathways during metabolism. Glycerol and fatty acids follow different pathways.*

Part 3: Protein Metabolism

Freely available amino acids are used to create proteins. If amino acids exist in excess, the body has no capacity or mechanism for their storage; thus, they are converted into glucose or ketone bodies. Amino acid breakdown results in hydrocarbons, which are converted to glucose through gluconeogenesis, and nitrogenous

waste, due to the removal of the amino group via deamination (i.e. ammonium, NH_4^+). However, high concentrations of nitrogen are toxic. The **urea cycle**, a liver process, converts ammonium into urea, facilitating the excretion of excess nitrogen from the body.

In the urea cycle, ammonium is combined with CO_2, resulting in urea and water. The urea is eliminated through the kidneys in the urine.

Amino acids can also be used as a source of energy, especially in times of starvation. Because the processing of amino acids results in the creation of metabolic intermediates, including pyruvate, acetyl CoA, acetoacyl CoA, oxaloacetate, and α-ketoglutarate, amino acids can serve as a source of energy production through the Krebs cycle (Figure 16). Figure 17 summarizes the pathways of catabolism and anabolism for carbohydrates, lipids, and proteins.

Part 4: Metabolic States of the Body

You eat periodically throughout the day; however, your organs, especially the brain, need a continuous supply of glucose. How does the body meet this constant demand for energy? Your body processes the food you eat both to use immediately and, importantly, to store as energy for later demands. If there were no method in place to store excess energy, you would need to eat constantly in order to meet energy demands. Distinct mechanisms are in place to facilitate energy storage, and to make stored energy available during times of fasting and starvation.

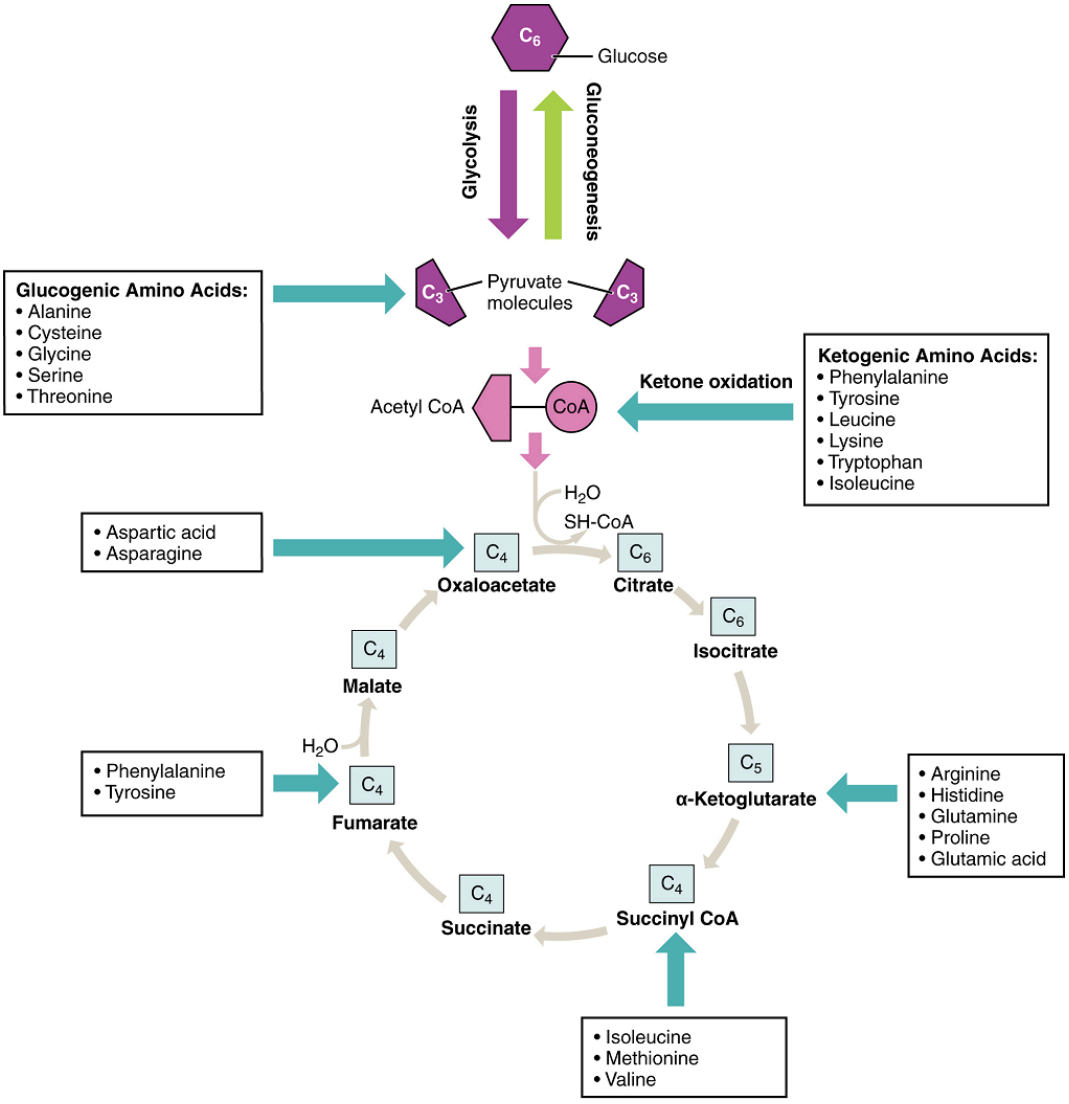


Figure 16. Accessing the Energy in Amino Acids. *Amino acids can be broken down into precursors for glycolysis or the Krebs cycle. Amino acids (in bold) can enter the cycle through more than one pathway. The points of entry of all the amino acids are not examinable.*

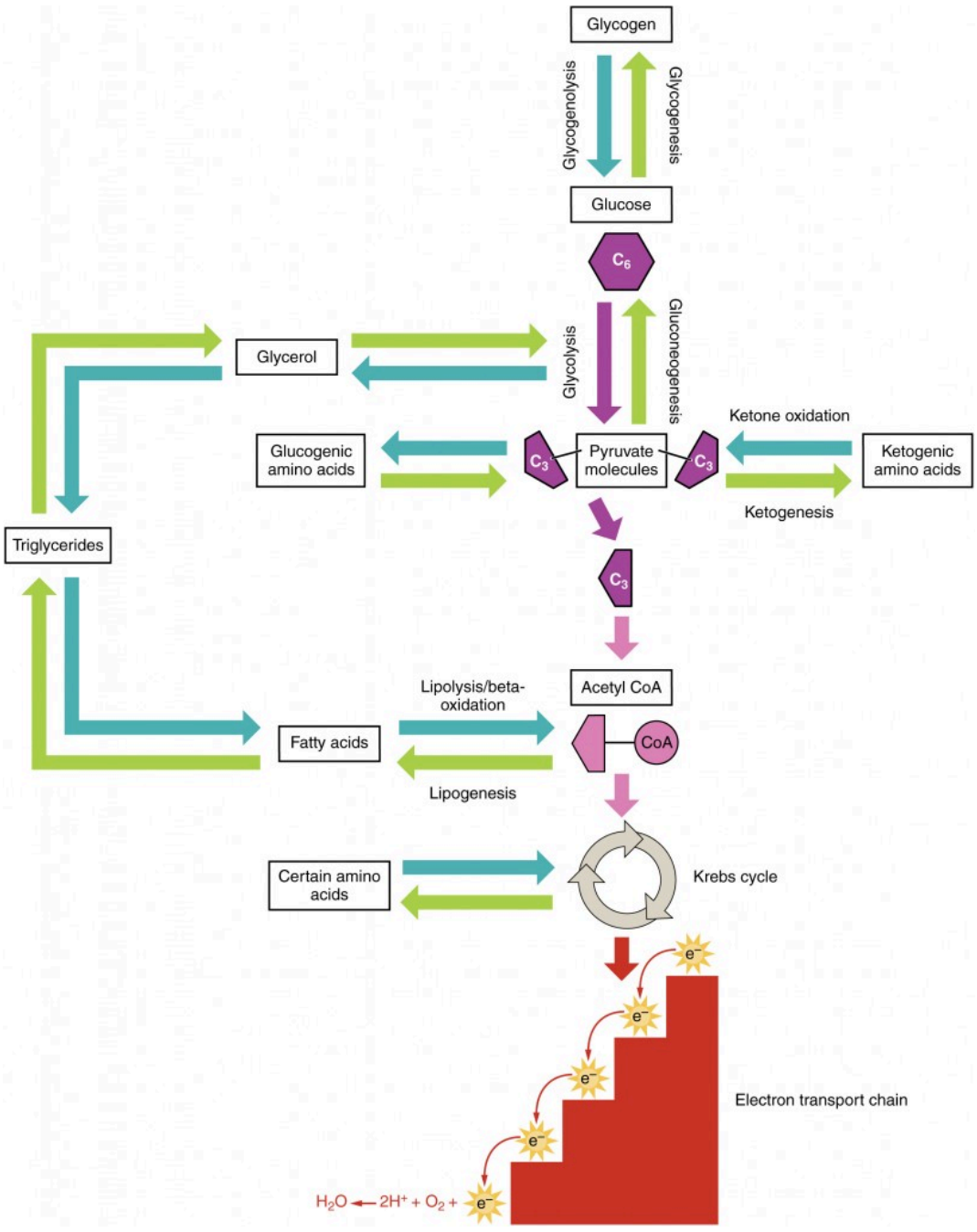


Figure 17. Catabolic and Anabolic Pathways. *Nutrients follow a complex pathway from ingestion through anabolism and catabolism to energy production.*

The Absorptive State: The absorptive state, or the fed state, occurs after a meal when your body is digesting the food and absorbing the nutrients (anabolism exceeds catabolism) (Figure 18). Digestion begins the moment you put food into your mouth, as the food is broken down into its constituent parts to be absorbed through the intestine. The digestion of carbohydrates begins in the mouth, whereas the digestion of proteins and fats begins in the stomach and small intestine. The constituent parts of these carbohydrates, fats, and proteins are transported across the intestinal wall and enter the bloodstream (sugars and amino acids) or the lymphatic system (fats). From the intestines, these systems transport them to the liver, adipose tissue, or muscle cells that will process and use, or store, the energy.

Depending on the amounts and types of nutrients ingested, the absorptive state can linger for up to 4 hours.

The ingestion of food and the rise of glucose concentrations in the bloodstream stimulate pancreatic beta cells to release **insulin** into the bloodstream, where it initiates the absorption of blood glucose by liver hepatocytes, and by adipose and muscle cells. Insulin also stimulates **glycogenesis**, the storage of glucose as glycogen, in the liver and muscle cells where it can be used for later energy needs of the body. Insulin also promotes the synthesis of protein in muscle. As you will see, muscle protein can be catabolized and used as fuel in times of starvation.

If energy is exerted shortly after eating, the dietary fats and sugars that were just ingested will be processed and used immediately for energy. If not, the excess glucose is stored as glycogen in the liver and muscle cells, or as fat in adipose tissue; excess dietary fat is also stored as triglycerides in adipose tissues.

The Postabsorptive State: The postabsorptive state, or the fasting state, occurs when the food has been digested, absorbed, and stored (Figure 19). You commonly fast overnight, but skipping meals during the day puts your body in the postabsorptive state as well. During this state, the body must rely initially on stored **glycogen**. Glucose levels in the blood begin to drop as it is absorbed and used by the cells. In response to the decrease in glucose, insulin levels also drop. Glycogen and triglyceride storage slows. However, due to the demands of the tissues and organs, blood glucose levels must be maintained in the normal range of 80–120 mg/dL. In response to a drop in blood glucose concentration, the hormone glucagon is released from the alpha cells of the pancreas. Glucagon acts upon the liver cells, where it inhibits glycogenesis and stimulates **glycogenolysis**, the breakdown of stored glycogen back into glucose. The glucose is released from the liver to be used by the peripheral tissues and the brain. As a result, blood glucose levels begin to rise. The stored glycogen in a well-fed human typically is sufficient to meet the energy needs of the body for several hours. **Gluconeogenesis**, the production of glucose from non-carbohydrates, will also begin in the liver to replace the glucose that has been used by the peripheral tissues.

Starvation: When the body is deprived of nourishment for an extended period of time, it goes into "survival mode." The first priority for survival is to provide enough glucose or fuel for the brain. The second priority is the conservation of amino acids for proteins. Therefore, when glucose is no longer available, the use of ketone bodies as an energy source helps to decrease the demand for glucose, thus minimizing gluconeogenesis in order to maintain body proteins.

Because glucose levels are very low during starvation, glycolysis will shut off in cells that can use alternative fuels. For example, muscles will switch from using glucose to fatty acids as fuel. As previously explained, fatty acids can be converted into acetyl CoA and processed through the Krebs cycle to make ATP. Pyruvate, lactate, and alanine from muscle cells are not converted into acetyl CoA and used in the Krebs cycle, but are exported to the liver to be used in the synthesis of glucose. As starvation continues, and more glucose is needed, glycerol from fatty acids can be liberated and used as a source for gluconeogenesis.

After several days of starvation, ketone bodies become the major source of fuel for the heart and other organs. As starvation continues, fatty acids and triglyceride stores are oxidized to create these molecules. This prevents the continued breakdown of proteins that serve as carbon sources for gluconeogenesis, helping to maintain the proper functioning of the body's muscles. Once these lipid stores are fully depleted, proteins from muscles are released and broken down for glucose synthesis. This leads to muscle wasting, as the body is forced to cannibalize the tissue for survival. Overall survival is dependent on the amount of fat and protein stored in the body.

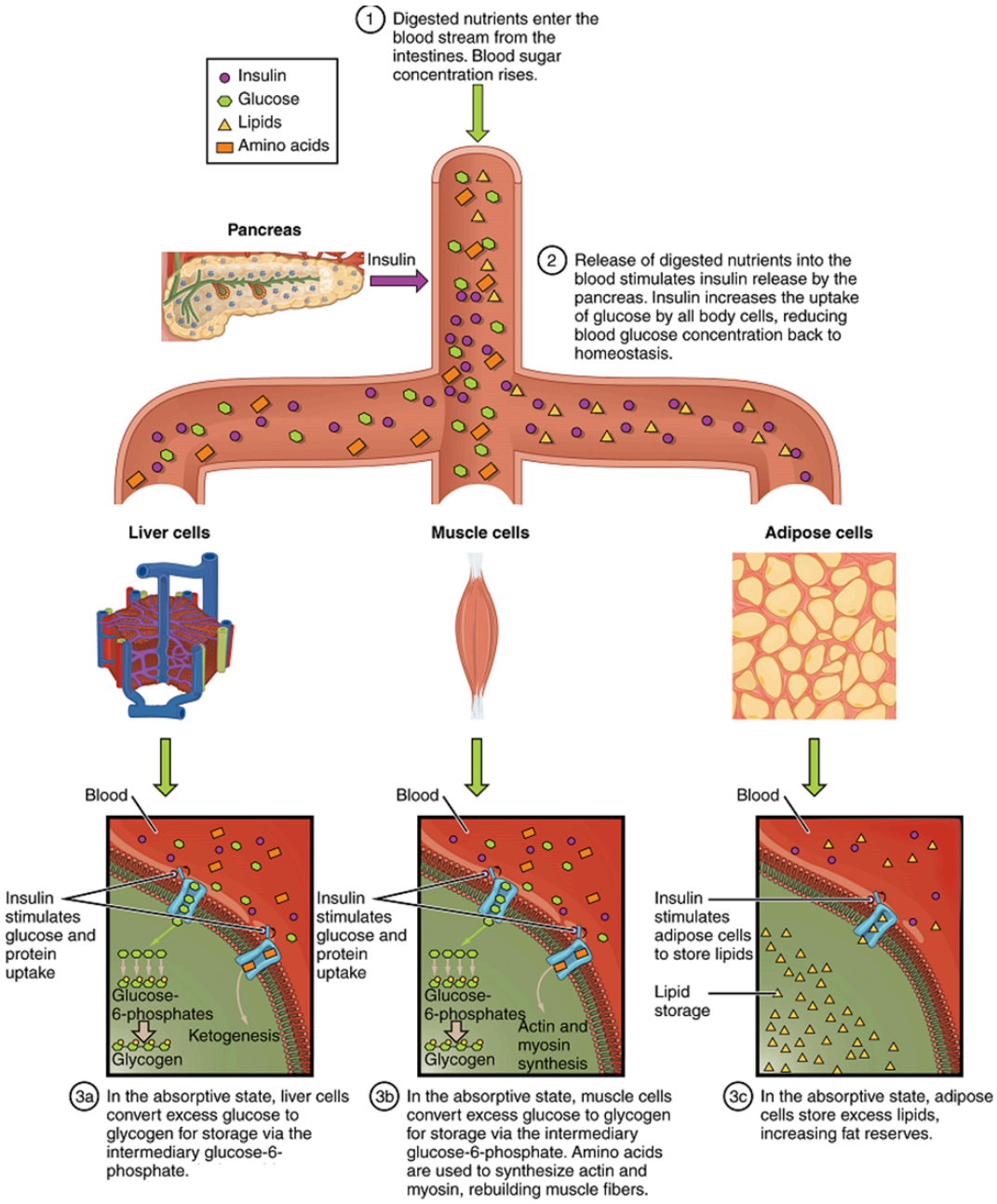


Figure 18. Absorptive State. *During the absorptive state, the body digests food and absorbs the nutrients.*

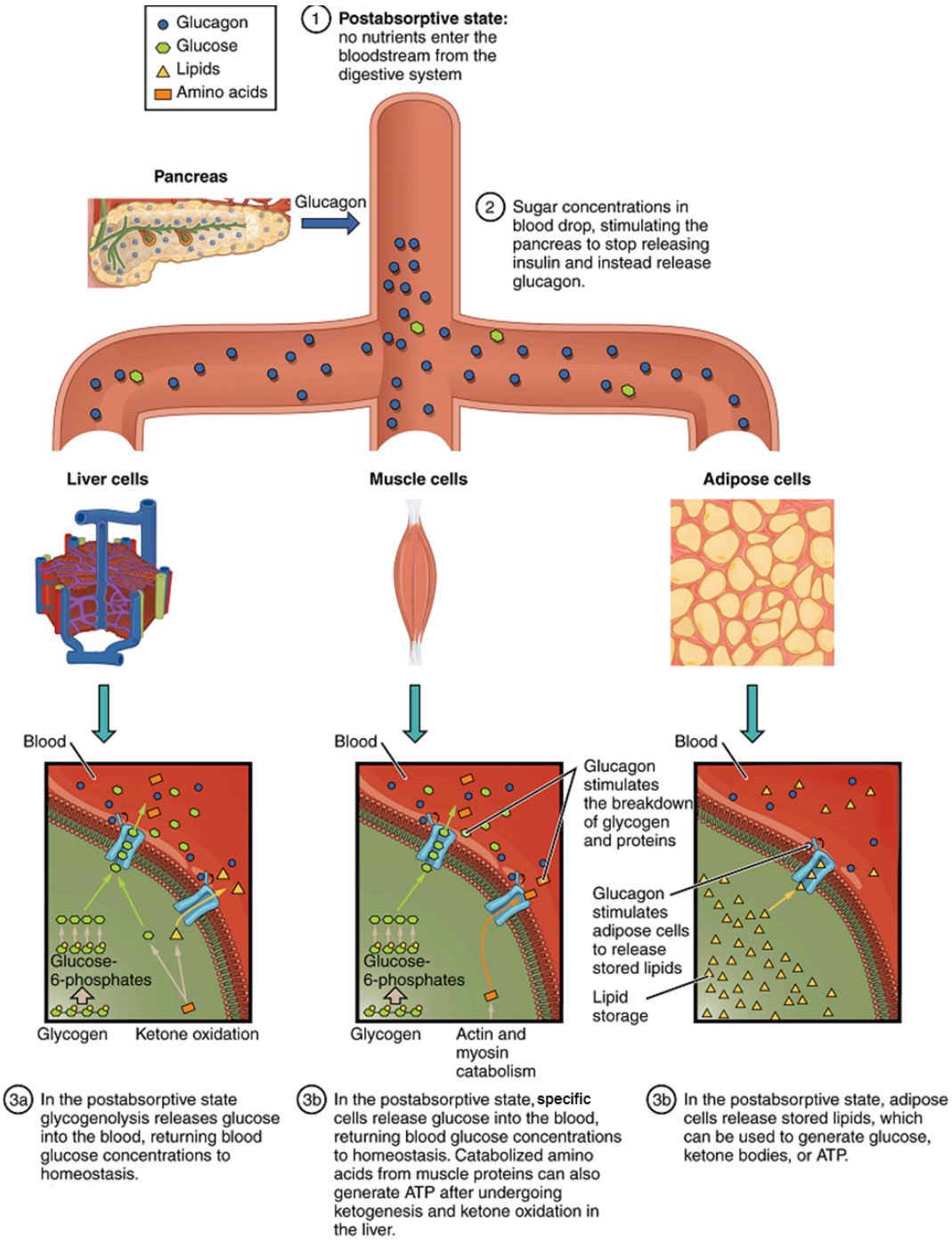


Figure 19. Postabsorptive State. *During the postabsorptive state, the body must rely on stored glycogen for energy.*

Unit 8: The Urinary System

Unit outline

Part 1: The Urinary Tract

- Gross Anatomy of the Kidney
- External Anatomy
- Internal Anatomy
- Renal Hilum
- Nephrons and Vessels
- Ureters
- Bladder
- Urethra

Part 2: Microscopic Anatomy of the Kidney

- Cortex
- Nephrons
- Renal Corpuscle
- Proximal Convoluted Tubule (PCT)
- Loop of Henle
- Distal Convoluted Tubule (DCT)
- Collecting Ducts

Part 3: Physiology of Urine Formation

- Filtration, Tubular Reabsorption and Secretion
- Reabsorption and Secretion in the Proximal Convoluted Tubule
- Descending Loop
- Ascending Loop
- Reabsorption and Secretion in the Distal Convoluted Tubule
- Collecting Ducts and Recovery of Water

Part 4: Micturition Reflex

- Micturition

Part 5: Physical Characteristics of Urine

- Urinalysis

Learning Objectives

At the end of this unit, you should be able to:

I. Specify the location and describe the function(s) of each of the organs of the renal system.

II. Describe the gross anatomy of the kidney.

III. Describe the formation and composition of urine.

IV. Describe the hormonal control of urine production.

V. Describe the nervous control of micturition.

VI. Describe the physical properties and chemical composition of urine.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Specify the location and describe the function(s) of each of the organs of the renal system.

1. Sketch a diagram with annotations specifying the location (using correct anatomical terms) and function (one-sentence summary for each structure) of the following structures:
    - Kidneys
    - Ureters
    - Urinary bladder
    - Urethra

II. Describe the gross anatomy of the kidney.

1. Sketch a diagram of the kidneys showing its gross anatomy and blood supply. Your diagram should include all the following structures:
    - Renal cortex
    - Renal medulla
    - Renal pelvis
    - Renal artery
    - Renal vein
2. Describe in layman's terms the appearance (size, shape, colour) of a human kidney.
3. Describe the structure of a nephron and explain how the nephrons are arranged in the kidney.
4. Use correct anatomical terms to describe the location and function (one-sentence summary for each structure) of all the following structures of/at a nephron:
    - Afferent arteriole
    - Proximal convoluted tubule

- Glomerulus
- Descending Loop of Henle
- Efferent arteriole
- Ascending Loop of Henle
- Peritubular capillaries
- Distal convoluted tubule
- Glomerular capsule
- Collecting duct

III. Describe the formation and composition of urine.

1. Describe each of the three main processes that occur in a nephron that result in urine production from blood:
    - Glomerular filtration
    - Tubular reabsorption
    - Tubular secretion
2. Compare the composition of glomerular filtrate and urine in humans with respect to the following: glucose, proteins, salts, urea, uric acid, water.

IV. Describe the hormonal control of urine production.

1. For aldosterone and anti-diuretic hormone, describe (where applicable):
    - The stimulus or stimuli that causes their release (or increased release)
    - The effect of the hormone on urine composition and volume
    - The effect of the hormone on blood composition and volume

V. Describe the nervous control of micturition.

1. Describe the involuntary and voluntary pathways that control the process of micturition, and the interaction between the two pathways.

VI. Describe the physical properties and chemical composition of urine.

1. Describe normal urine in terms of the following:
    - Volume voided in single day
    - Specific gravity
    - pH
    - Chemical composition

Part 1: The Urinary Tract

Urine is a fluid of variable composition that requires specialized structures to remove it from the body safely and efficiently. Blood is filtered, and the filtrate is transformed into urine at a relatively constant rate throughout the day. This processed liquid is stored until a convenient time for excretion. All structures involved in the transport and storage of the urine are large enough to be visible to the naked eye.

Gross Anatomy of the Kidney: The kidneys lie on either side of the spine in the retroperitoneal space between

the parietal peritoneum and the posterior abdominal wall, well protected by muscle, fat, and ribs. They are roughly the size of your fist. The kidneys are well vascularized, receiving about 25 percent of the cardiac output at rest.

External Anatomy: The left kidney is located at about the T12 to L3 vertebrae, whereas the right is lower due to slight displacement by the liver. Upper portions of the kidneys are somewhat protected by the eleventh and twelfth ribs (Figure 1). Each kidney weighs about 125–175 g in males and 115–155 g in females. They are about 11–14 cm in length, 6 cm wide, and 4 cm thick, and are directly covered by a fibrous capsule composed of dense, irregular connective tissue that helps to hold their shape and protect them.

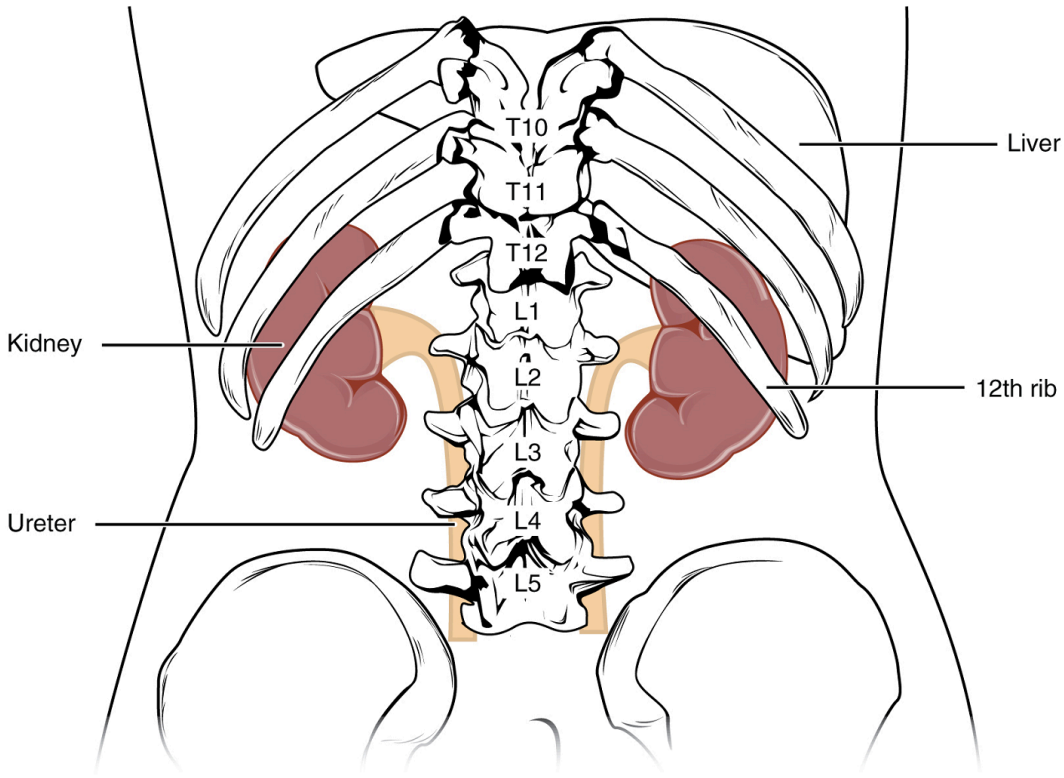


Figure 1. Kidneys. *The kidneys are slightly protected by the ribs and are surrounded by fat for protection (not shown).*

This capsule is covered by a shock-absorbing layer of adipose tissue called the **renal fat pad.** On the superior aspect of each kidney is the adrenal gland.

Internal Anatomy: A frontal section through the kidney reveals an outer region called the **renal cortex** and an inner region called the **medulla** (Figure 2). The **renal columns** are connective tissue extensions that radiate downward from the cortex through the medulla to separate the most characteristic features of the medulla, the **renal pyramids** and **renal papillae**. The papillae are bundles of collecting ducts that transport urine made by nephrons to the calyces of the kidney for excretion. The renal columns also serve to divide the kidney into 6–8 lobes and provide a supportive framework for vessels that enter and exit the cortex. The pyramids and renal columns taken together constitute the kidney lobes.

Renal Hilum: The **renal hilum** is the entry and exit site for structures serving the kidneys: blood vessels, nerves, lymphatics, and ureters. Emerging from the hilum is the renal pelvis, which is formed from the major and minor calyces in the kidney. The smooth muscle in the renal pelvis funnels urine via peristalsis into the ureter. The renal arteries form directly from the descending aorta, whereas the renal veins return cleansed blood directly to the inferior vena cava. The artery, vein, and renal pelvis are arranged in an anterior-to-posterior order.

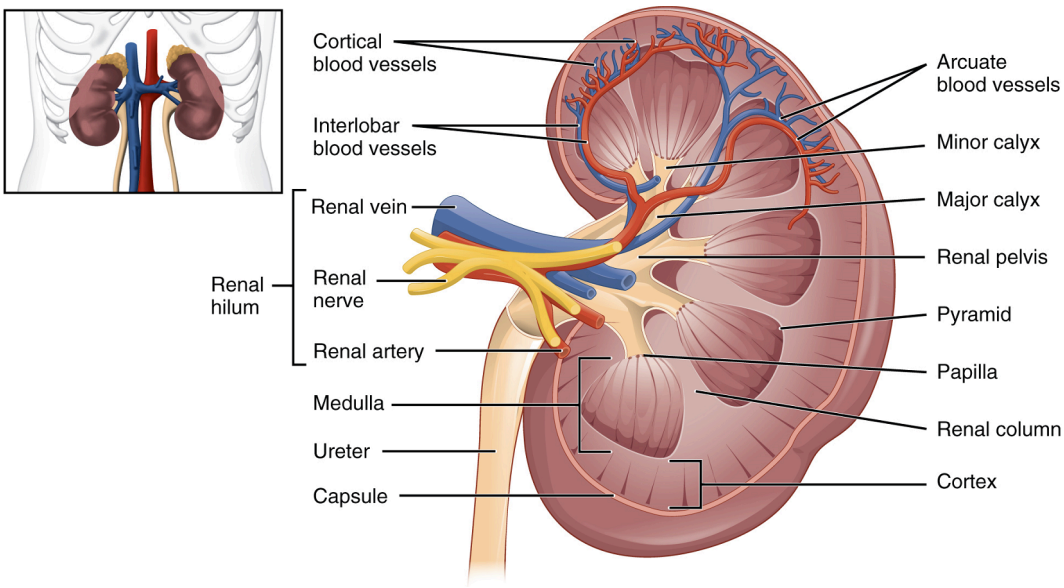


Figure 2. Left Kidney.

Nephrons and Vessels: The renal artery first divides into segmental arteries, followed by further branching to form interlobar arteries that pass through the renal columns to reach the cortex (Figure 3). The interlobar arteries, in turn, branch into arcuate arteries, cortical radiate arteries, and then into afferent arterioles. The afferent arterioles supply blood to about 1.3 million nephrons in each kidney.

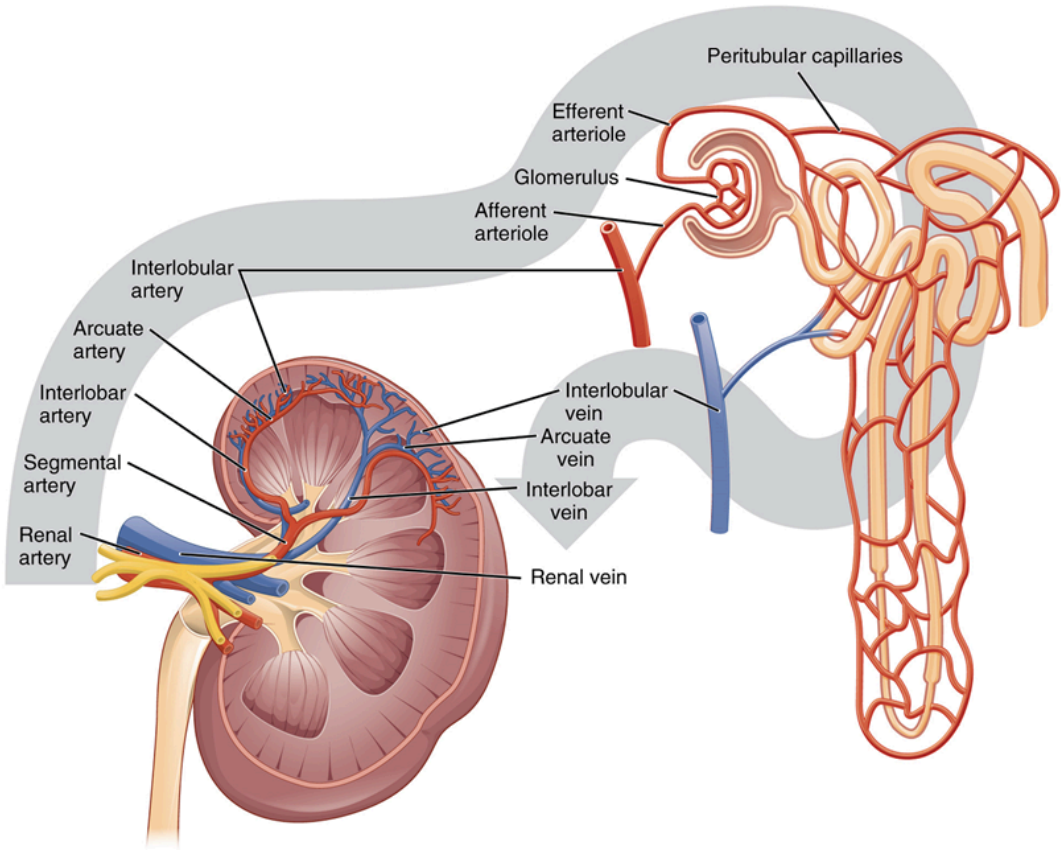


Figure 3. Blood Flow in the Kidney.

Ureters: The kidneys and ureters are completely retroperitoneal, and the bladder has a peritoneal covering only over the dome. As urine is formed, it drains into the calyces of the kidney, which merge to form the funnel-shaped renal pelvis in the hilum of each kidney. The hilum narrows to become the ureter of each kidney. As urine passes through the ureter, it does not passively drain into the bladder but rather is propelled by waves of peristalsis.

As they approach the bladder, they turn medially and pierce the bladder wall obliquely. This is important because it creates a one-way valve (a **physiological sphincter** rather than an **anatomical sphincter**) that allows urine into the bladder but prevents reflux of urine from the bladder back into the ureter.

The ureters are approximately 30 cm long. The inner mucosa is lined with transitional epithelium (Figure 5) and scattered goblet cells that secrete protective mucus. The muscular layer of the ureter consists of longitudinal and circular smooth muscles that create the peristaltic contractions to move the urine into the bladder without the aid of gravity (Figure 4).

Bladder: The urinary bladder collects urine from both ureters (Figure 5). The bladder lies anterior to the uterus in females, posterior to the pubic bone and anterior to the rectum. In males, the anatomy is similar, minus the uterus, and with the addition of the prostate inferior to the bladder. The bladder is partially **retroperitoneal** (outside the peritoneal cavity).

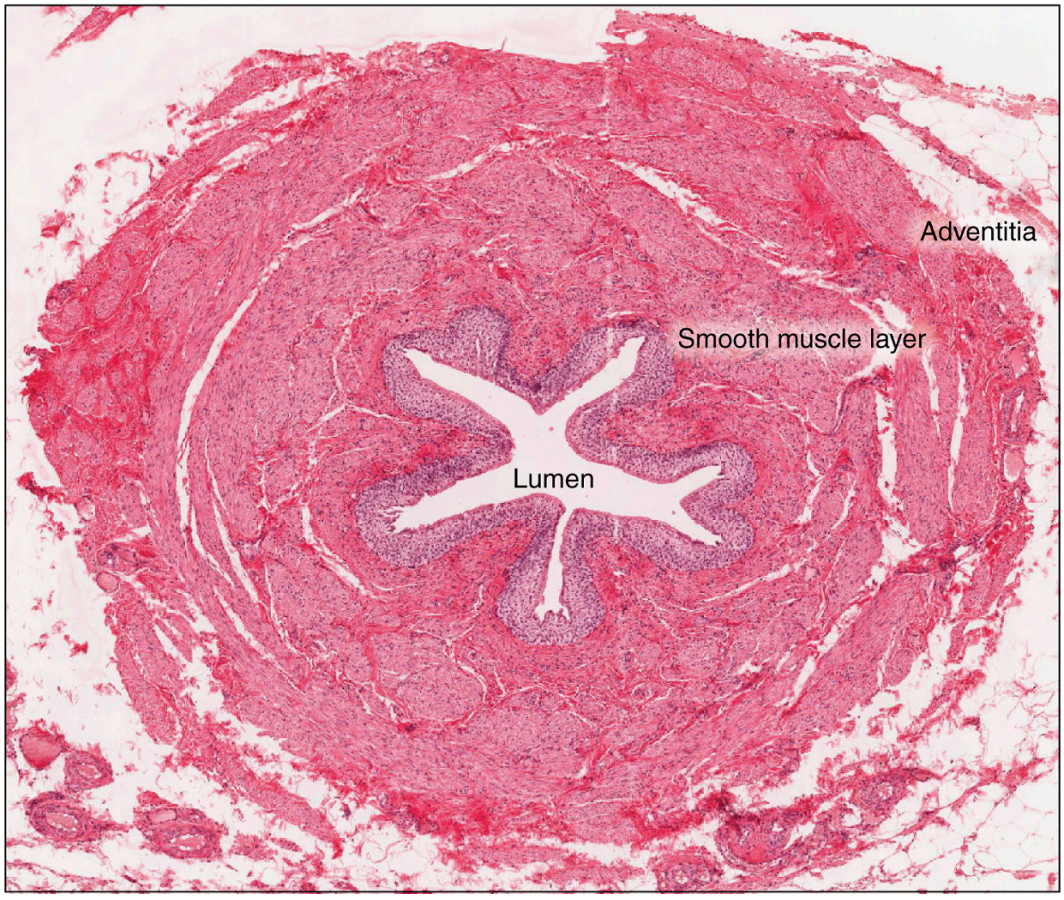


Figure 4. Ureter. *Peristaltic contractions help to move urine through the lumen with contributions from fluid pressure and gravity. LM × 128. (Micrograph provided by the Regents of the University of Michigan Medical School © 2012)*

The bladder is a highly distensible organ comprised of irregular crisscrossing bands of smooth muscle collectively called the **detrusor muscle**. The interior surface is made of transitional epithelium that is structurally suited for the large volume fluctuations of the bladder. Volumes in adults can range from nearly zero to 500–600 mL.

The detrusor muscle contracts with significant force in the young. The bladder's strength diminishes with

age, but voluntary contractions of abdominal skeletal muscles can increase intra-abdominal pressure to promote more forceful bladder emptying.

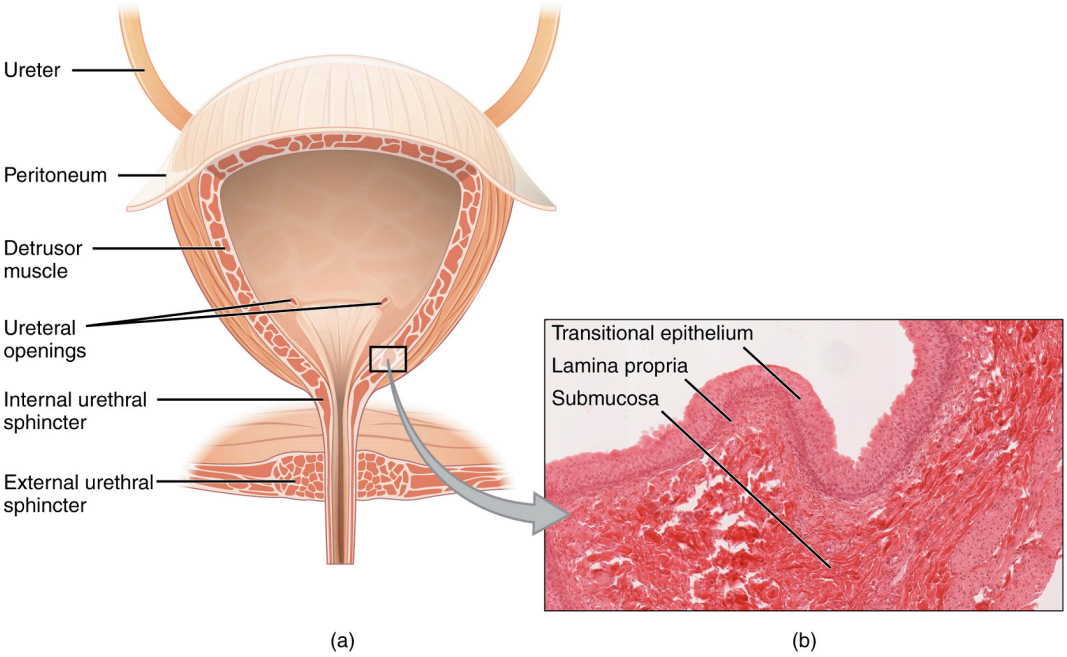


Figure 5. Urinary Bladder. *(a) Anterior cross section of the urinary bladder. (b) The detrusor muscle of the urinary bladder (source: monkey tissue) LM × 448. (Micrograph provided by the Regents of the University of Michigan Medical School © 2012)*

Urethra: The urethra transports urine from the bladder to the outside of the body for disposal. The urethra shows significant anatomic difference between males and females; all other urine transport structures are identical (Figure 6).

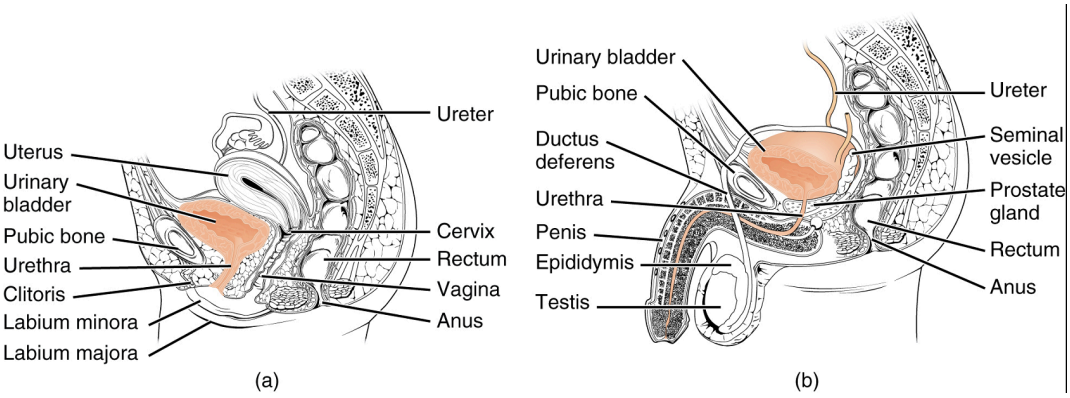


Figure 6. Female and Male Urethras. *The urethra transports urine from the bladder to the outside of the body. This image shows (a) a female urethra and (b) a male urethra.*

The urethra in both males and females begins inferior and central to the two ureteral openings. (Figure 5a). In both males and females, the proximal urethra is lined by transitional (stratified epithelium which can contract and expand) epithelium, whereas the terminal portion is a nonkeratinized, stratified squamous epithelium. Voiding is regulated by an involuntary autonomic nervous system-controlled **internal urinary sphincter**, consisting of smooth muscle and voluntary skeletal muscle that forms the **external urinary sphincter** below it.

Female Urethra: Its short length, about 4 cm, is less of a barrier to fecal bacteria than the longer male urethra and the best explanation for the greater incidence of urinary tract infections in women.

Male Urethra: The male urethra passes through the prostate gland immediately inferior to the bladder before passing below the pubic symphysis (Figure 6b). The urethra passes through the deep muscles of the perineum

and exits at the tip (external urethral orifice) of the penis. Mucous glands are found along much of the length of the urethra and protect the urethra from extremes of urine pH. Male urethra also serves as the duct through which semen (fluid containing sperms) is discharged.

Part 2: Microscopic Anatomy of the Kidney

The renal structures that conduct the essential work of the kidney cannot be seen by the naked eye.

Cortex: In a dissected kidney, it is easy to identify the cortex; it appears lighter in color compared to the rest of the kidney. All of the renal corpuscles as well as both the **proximal convoluted tubules** and **distal convoluted tubules** are found here. Some nephrons have a short **loop of Henle** that does not dip beyond the cortex. These nephrons are called **cortical nephrons**. About 15 percent of nephrons have long loops of Henle that extend deep into the medulla and are called **juxtamedullary nephrons**.

Nephrons: Nephrons are the functional unit of the kidney; they take a simple filtrate of the blood and modify it into urine. They cleanse the blood and maintain the levels of blood chemical components within physiological values. The afferent arterioles form a tuft of high-pressure capillaries, the **glomerulus**. The rest of the nephron consists of a continuous tubule, the proximal end of which surrounds the glomerulus—this is **Bowman's capsule** or **glomerular capsule**. As mentioned earlier, these glomerular capillaries filter the blood based on particle size.

Renal Corpuscle: The glomerulus and Bowman's capsule together form the **renal corpuscle**. After passing through the renal corpuscle, the capillaries form a second arteriole, the efferent arteriole (Figure 7). These will next form a capillary network around the more distal portions of the nephron tubule, the peritubular capillaries and vasa recta, before returning to the venous system. As the glomerular filtrate progresses through the nephron, these capillary networks regain most of the solutes and water and return them to the circulation.

The glomerulus is a high-pressure capillary bed between afferent and efferent arterioles. Bowman's capsule surrounds the glomerulus to form a lumen, and captures and directs this filtrate to the proximal convoluted tubule. As blood passes through the glomerulus, 10 to 20 percent of the fluid that moves into the glomerulus is captured by Bowman's capsule and funneled to the proximal convoluted tubule.

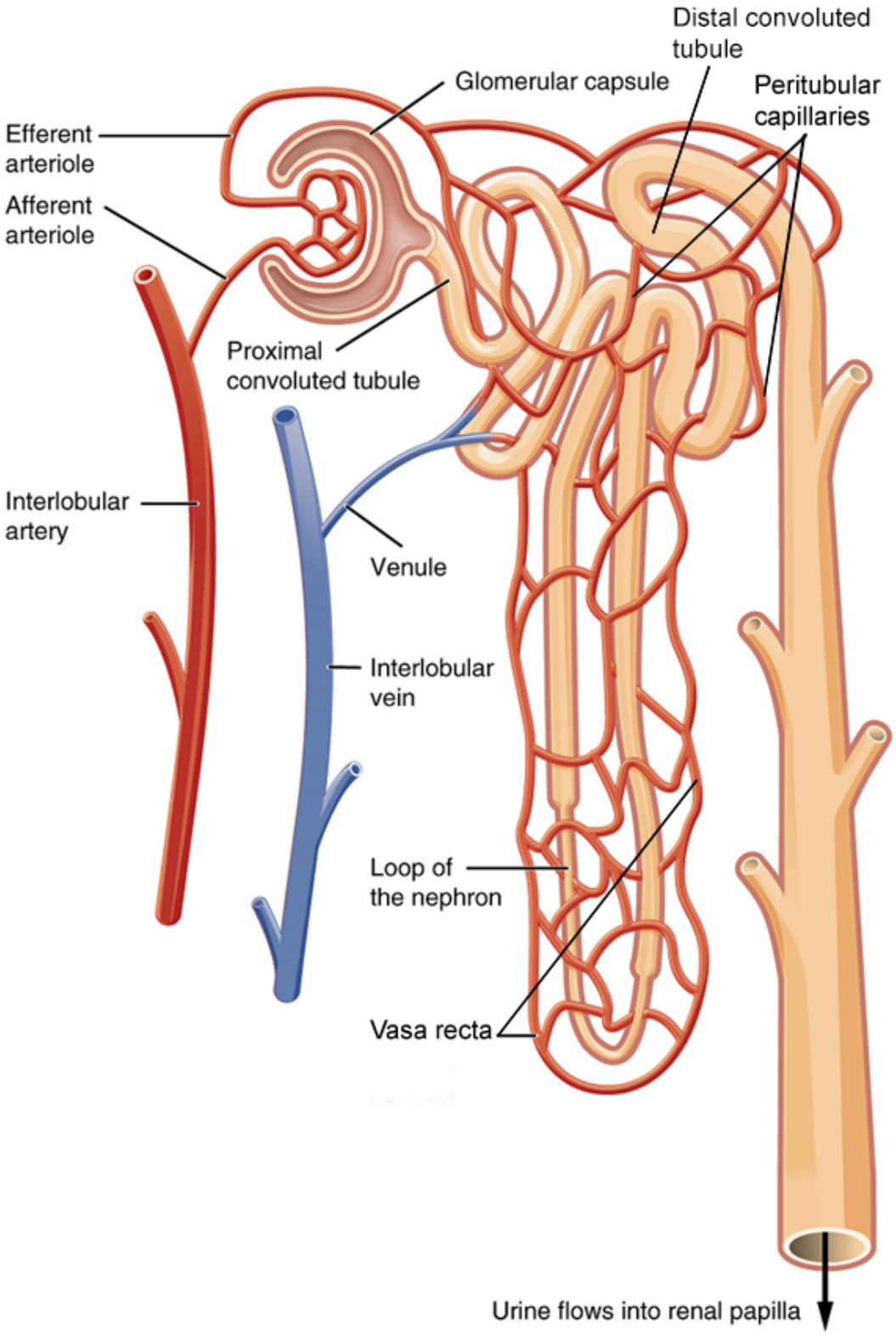


Figure 7. Blood Flow in the Nephron. *The two capillary beds are clearly shown in this figure. The efferent arteriole is the connecting vessel between the glomerulus and the peritubular capillaries and vasa recta.*

Proximal Convoluted Tubule (PCT): Filtered fluid collected by Bowman's capsule enters into the proximal convoluted tubule. It is called convoluted due to its tortuous path. Simple cuboidal cells form this tubule with prominent microvilli on the luminal surface, forming a brush border. These microvilli create a large surface area

to maximize the absorption and secretion of solutes (Na+, Cl–, glucose, etc.), the most essential function of this portion of the nephron. These cells actively transport ions across their membranes.

Loop of Henle: The descending and ascending portions of the loop of Henle (nephron loop) are, the continuations of the same tubule. The descending loop of Henle consists of an initial short, thick portion and long, thin portion, whereas the ascending loop consists of an initial short, thin portion followed by a long, thick portion. The descending thick portion consists of simple cuboidal epithelium similar to that of the proximal convoluted tubule. The descending and ascending thin portions consist of simple squamous epithelium. these are important differences, since different portions of the loop have different permeabilities for solutes and water. The ascending thick portion consists of simple cuboidal epithelium similar to the distal convoluted tubule.

Distal Convoluted Tubule (DCT): The distal convoluted tubule, like the proximal convoluted tubule, is very tortuous and formed by simple cuboidal epithelium, but it is shorter than the proximal convoluted tubule.

Collecting Ducts: The collecting ducts are continuous with the nephron but not generally considered part of it. In fact, each duct collects filtrate from several nephrons for final modification. Collecting ducts merge as they descend deeper in the medulla to form about 30 terminal ducts, which empty at a papilla. They are lined with simple squamous epithelium with receptors for antidiuretic hormone (ADH). When stimulated by antidiuretic hormone, these cells will allow water to pass from the duct lumen through the cells and into the interstitial spaces to be reabsorbed by the vasa recta. This process allows for the recovery of large amounts of water from the filtrate back into the blood.

Watch this CrashCourse video to learn about the production of urine! Direct link: https://youtu.be/l128tW1H5a8

Part 3: Physiology of Urine Formation – Filtration, Tubular Reabsorption, and Secretion

Different parts of the nephron utilize specific processes to produce urine: filtration, reabsorption, and secretion. The volume of filtrate formed by both kidneys per minute is termed the **glomerular filtration rate (GFR)**. The heart pumps about 5 L blood per min under resting conditions. Approximately 20 percent or one liter enters the kidneys to be filtered. Ninety-nine percent of this filtrate is returned to the circulation by reabsorption so that only about 1–2 liters of urine are produced per day. The glomerular filtration rate is influenced by the hydrostatic pressure and colloid osmotic pressure on either side of the capillary membrane of the glomerulus (Figure 8). Filtration occurs as pressure forces fluid and solutes through a semipermeable barrier with the solute movement constrained by particle size. Hydrostatic pressure is the pressure produced by a fluid against a surface. If you have a fluid on both sides of a barrier, both fluids exert a pressure in opposing directions. Net fluid movement will be in the direction of the lower pressure.

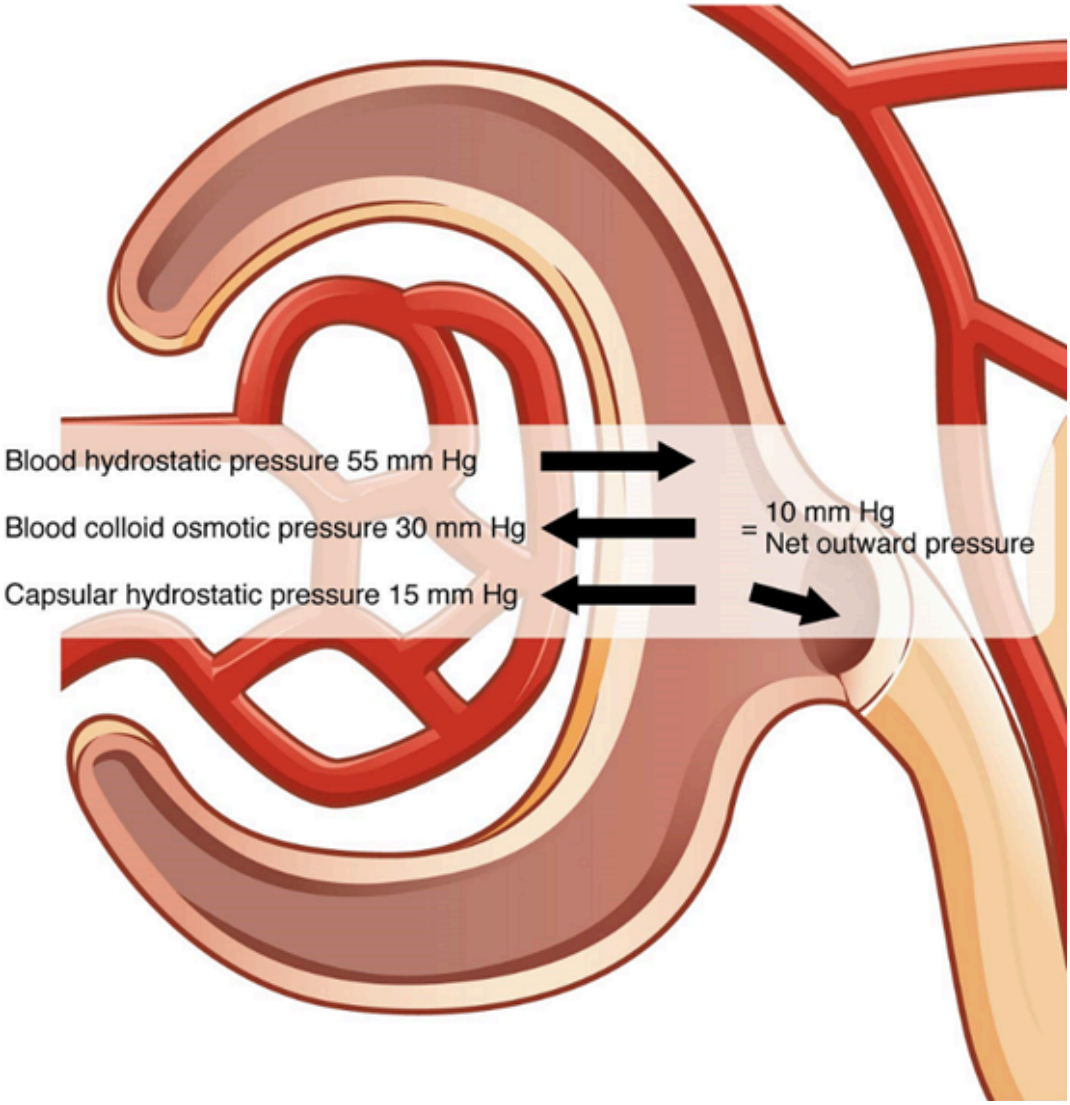


Figure 8. Pressures in the Glomerulus. *Hydrostatic and osmotic pressures interaction drives the fluid out of the capillary.*

Up to 180 liters per day passes through the nephrons of the kidney. The renal corpuscle filters the blood to create a filtrate that differs from blood mainly in the absence of cells and large proteins. From this point to the ends of the collecting ducts, the filtrate or forming urine is undergoing modification through secretion and reabsorption before true urine is produced. The first point at which the forming urine is modified is in the proximal convoluted tubule. Here, some substances are reabsorbed, whereas others are secreted. Note the use of the term "reabsorbed." All of these substances were "absorbed" in the digestive tract—99 percent of the water and most of the solutes filtered by the nephron must be reabsorbed. Water and substances that are reabsorbed are returned to the circulation by the peritubular and vasa recta capillaries (Figure 7). Movement of water into the peritubular capillaries and vasa recta will be influenced primarily by osmolarity and concentration gradients. Sodium is actively pumped out of the proximal convoluted tubule into the interstitial spaces between cells and diffuses down its concentration gradient into the peritubular capillaries. As it does so, water will follow passively to maintain an isotonic fluid environment inside the capillary. Most of that fluid and its contents are reabsorbed. That recovery occurs in the proximal convoluted tubule, loop of Henle, distal convoluted tubule, and the collecting ducts (Table 1 and Figure 9). Various portions of the nephron differ in their capacity to reabsorb water and specific solutes. While much of the reabsorption and secretion occur passively based on concentration gradients, the amount of water that is reabsorbed or lost is tightly regulated. This control is

exerted directly by antidiuretic hormone and another hormone aldosterone. Most water is recovered in the proximal convoluted tubule, loop of Henle, and distal convoluted tubule. About 10 percent (about 18 L) reaches the collecting ducts. The collecting ducts, under the influence of antidiuretic hormone, can recover almost all of the water passing through them.

Mechanisms by which substances move across membranes for reabsorption or secretion include active transport, diffusion, facilitated diffusion, secondary active transport, and osmosis. These were discussed in an earlier chapter, and you may wish to review them.

Most of the Ca^{2+}, Na^{+}, glucose, and amino acids must be reabsorbed by the nephron to maintain homeostatic plasma concentrations. Other substances, such as urea, K^{+}, ammonia (NH_3), creatinine, and some drugs are secreted into the filtrate as waste products. Acid–base balance is maintained through actions of the lungs and kidneys. In the case of urea, about 50 percent is passively reabsorbed by the proximal convoluted tubule.

A few of the substances that are transported with Na^{+} include Cl^{-}, Ca^{2+}, amino acids, glucose, and $PO4^{3-}$. Sodium is actively exchanged for K^{+} using ATP on the basal membrane.

About 67 percent of the water, Na^{+}, and K^{+} entering the nephron is reabsorbed in the proximal convoluted tubule and returned to the circulation. Almost 100 percent of glucose, amino acids, and other organic substances such as vitamins are normally recovered here.

More substances move across the membranes of the proximal convoluted tubule than any other portion of the nephron.

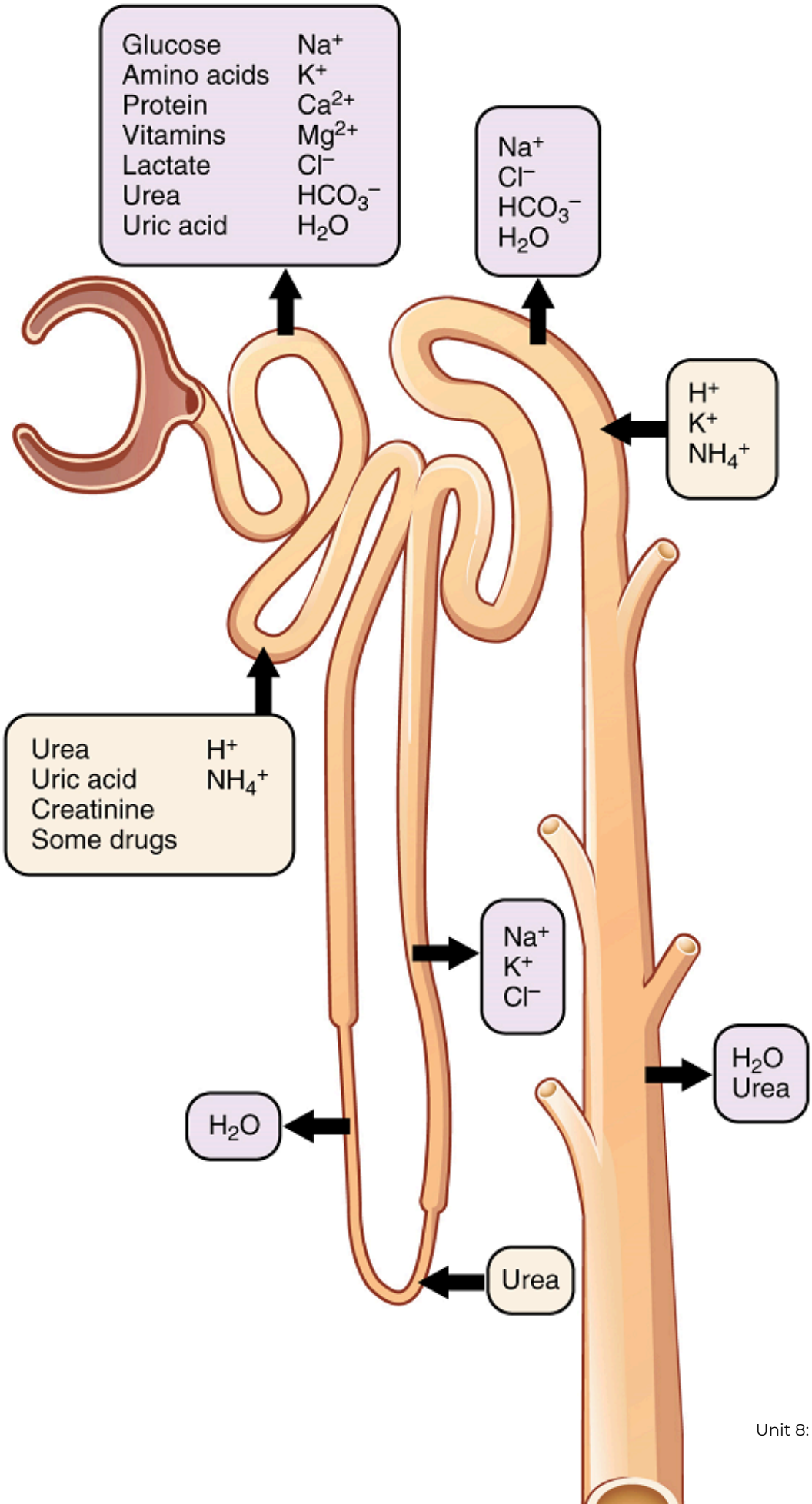


Figure 9. Locations of Secretion and Reabsorption in the Nephron.

Table 1: Substances Secreted or Reabsorbed in the Nephron

Substance	Proximal convoluted tubule	Loop of Henle	Distal convoluted tubule	Collecting ducts
Glucose	~100% reabsorbed; secondary active transport with Na^+			
Amino acids, oligopeptides, proteins	~100% reabsorbed; symport with Na^+			
Vitamins	Reabsorbed			
Lactate	Reabosorbed			
Creatinine	Secreted			
Urea	50% reabsorbed by diffusion; also secreted	Secretion by diffusion in descending limb		Reabsorption by diffusion in medullary collecting ducts
Sodium	65% reabsorbed	25% reabsorbed by active transport in thick ascending limb	5% reabsorbed by active transport	5% reabsorbed by active transport, stimulated by aldosterone
Chloride	Reabsorbed by symport with Na^+ and diffusion	Reabsorbed by diffusion in thin & thick ascending limb	Reabsorbed by diffusion	Reabsorbed by symport
Water	67% reabsorbed by osmosis (follows solutes)	15% reabsorbed by osmosis in descending limb	In the presence of antidiuretic hormone, 8% reabsorbed by osmosis	Variable amounts reabsorbed by osmosis; regulated by antidiuretic hormone
Bicarbonate	80-90% reabsorbed by symport with Na^+	Reabsorbed in ascending limb by symport with Na^+ and antiport with Cl^-		Reabsorbed by antiport with Cl^-
Hydrogen ions (H^+)	Secreted by diffusion		Secreted by active transport	Secreted by active transport
Ammonium (NH_4^+)	Secreted by diffusion		Secreted by diffusion	Secreted by diffusion
Bicarbonate (HCO_3^-)	Reabsorbed by diffusion	Reabsorbed by diffusion in ascending limb	Reabsorbed by diffusion	Reabsorbed by antiport with Na^+
Some drugs	Secreted		Secreted by active transport	Secreted by active transport
Potassium	65% reabsorbed by diffusion	20% reabsorbed by symport in thick ascending limb	Secreted by active transport	Secreted by active transport, regulated by aldosterone
Calcium	Reabsorbed by diffusion	Reabsorbed by diffusion in thick ascending limb		In the presence of parathyroid hormone, reabsorbed by active transport
Magnesium	Reabsorbed by diffusion	Reabsorbed by diffusion in thick ascending limb	Reabsorbed	
Phosphate	85% reabsorbed by diffusion; inhibited by parathyroid hormone		Reabsorbed by diffusion	

Reabsorption and Secretion in the Loop of Henle: The descending and ascending portions of the loop are

highly specialized to enable recovery of much of the Na^+ and water that were filtered by the glomerulus. As the forming urine moves through the loop, the osmolarity will change from isosmotic with blood to both a very hypertonic solution and a very hypotonic solution. These changes are accomplished by osmosis in the descending limb and active transport in the ascending limb. Solutes and water recovered from these loops are returned to the circulation by way of the vasa recta.

Descending Loop: The majority of the descending loop is comprised of simple squamous epithelial cells; to simplify the function of the loop, this discussion focuses on these cells. These membranes have permanent aquaporin channel proteins that allow unrestricted movement of water from the descending loop into the surrounding interstitial fluid. Approximately 15 percent of the water and modest amounts of urea, Na^+, and other ions are recovered here.

Most of the solutes that were filtered in the glomerulus have now been recovered along with a majority of water, about 82 percent. As the forming urine enters the ascending loop, major adjustments will be made to the concentration of solutes to create urine.

Ascending Loop: The thick portion is lined with simple cuboidal epithelium It is completely impermeable to water due to the absence of aquaporin proteins, but ions, mainly Na^+, are actively pumped out of the loop by large numbers of the Na^+/K^+ ATPase pump. This causes the removal of Na^+ while retaining water leading to a hypotonic filtrate by the time it reaches the distal convoluted tubule.

The Na^+/K^+ ATPase pumps allow Na^+ to be actively pumped into the interstitial fluid, Cl^- follows the Na^+ from the lumen into the interstitial fluid.

In addition, collecting ducts have urea pumps that actively pump urea into the interstitial spaces. This results in the recovery of Na^+ to the circulation via the vasa recta and creates a high osmolar environment in the depths of the medulla.

At the same time that water is freely diffusing out of the descending loop through aquaporin channels into the interstitial spaces of the medulla, urea freely diffuses into the lumen of the descending loop as it descends deeper into the medulla, much of it to be reabsorbed from the forming urine when it reaches the collecting duct. Thus, the movement of Na^+ and urea into the interstitial spaces by these mechanisms creates the hyperosmotic environment of the medulla. The net result of this is to recover both water from the descending limb of the loop of Henle and Na^+ in the circulation (Figure 10).

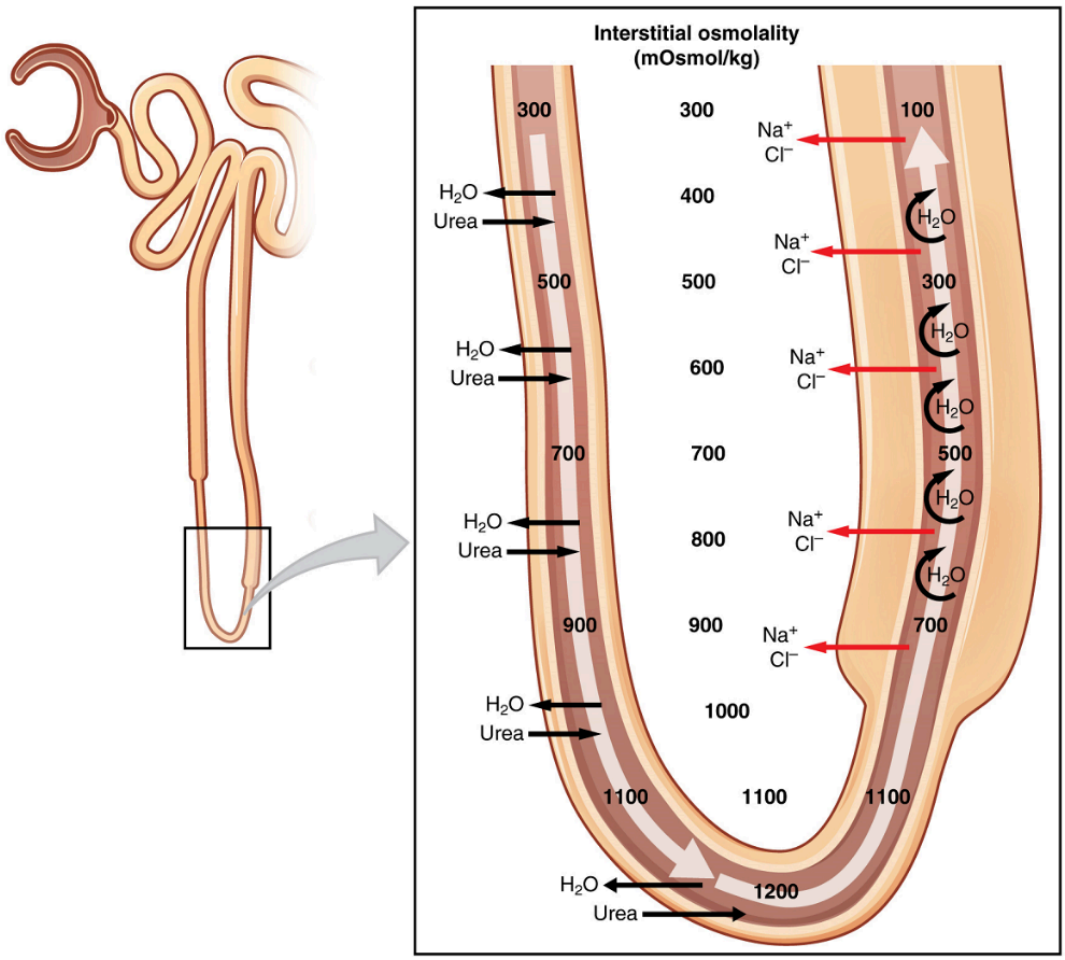


Figure 10. Water and Solute Reabsorption from the Loop of Henle.

Reabsorption and Secretion in the Distal Convoluted Tubule: Approximately 80 percent of filtered water has been recovered by the time the dilute forming urine enters the distal convoluted tubule. The distal convoluted tubule will recover another 10–15 percent before the forming urine enters the collecting ducts. Aldosterone increases the amount of Na^+/K^+ ATPase in the basal membrane of the distal convoluted tubule and collecting duct. The movement of Na^+ out of the lumen of the collecting duct creates a negative charge that promotes the movement of Cl^- out of the lumen into the interstitial space. Peritubular capillaries receive the solutes and water, returning them to the circulation.

Cells of the distal convoluted tubule also recover Ca^{2+} from the filtrate. In addition, as Na^+ is pumped out of the cell, the resulting electrochemical gradient attracts Ca^{2+} into the cell. Any Ca^{2+} not reabsorbed at this point is lost in the urine.

Collecting Ducts and Recovery of Water: Regulation of urine volume and osmolarity are major functions of the collecting ducts. By varying the amount of water that is recovered, the collecting ducts play a major role in maintaining the body's normal osmolarity. If the blood becomes hyperosmotic, the collecting ducts recover more water to dilute the blood; if the blood becomes hyposmotic, the collecting ducts recover less of the water, leading to concentration of the blood. Another way of saying this is: If plasma osmolarity rises, more water is recovered and urine volume decreases; if plasma osmolarity decreases, less water is recovered and urine volume increases. This function is regulated by antidiuretic hormone (vasopressin), a hypothalamic hormone that is stored and released by the posterior pituitary. With mild dehydration, plasma osmolarity rises slightly. This increase is detected by osmoreceptors in the hypothalamus, which stimulates the release of antidiuretic hormone from the posterior pituitary. If plasma osmolarity decreases slightly, the opposite occurs.

When stimulated by antidiuretic hormone, aquaporin channels are inserted into the apical membrane of the cells lining the collecting ducts. As the ducts descend through the medulla, the osmolarity surrounding them increases. If aquaporin water channels are present, water will be osmotically pulled from the collecting duct into the surrounding interstitial space and into the peritubular capillaries. Therefore, the final urine will be more concentrated. If less antidiuretic hormone is secreted, fewer aquaporin channels are inserted and less water is recovered, resulting in dilute urine. By altering the number of aquaporin channels, the volume of water recovered or lost is altered. This, in turn, regulates the blood osmolarity, blood pressure, and osmolarity of the urine.

Aldosterone is a hormone produced by the cortex of the adrenal glands. The control of aldosterone is complex (Figure 11). One mechanism involves detection of a decrease in blood pressure by certain cells in the kidney, resulting in the release of the enzyme renin. This ultimately leads to the production of Angiotensin II, which stimulates the adrenal cortex to produce aldosterone. A second mechanism involves potassium ion concentration. An increased in K^+ concentration in extracellular fluid directly stimulates aldosterone secretion by the adrenal cortex.

Aldosterone acts on the cells lining the distal convoluted tubule and the collecting duct to promote, simultaneously, active reabsorption of Na^+ from the tubules and active secretion of K^+ into the tubules. As Na^+ is pumped from the forming urine, water is passively recaptured for the circulation; this preservation of vascular volume is critically important for the maintenance of a normal blood pressure. As an extremely potent vasoconstrictor, angiotensin II functions immediately to increase blood pressure. It also stimulates aldosterone production, which provides a longer-lasting mechanism to support blood pressure by maintaining vascular volume (water recovery).

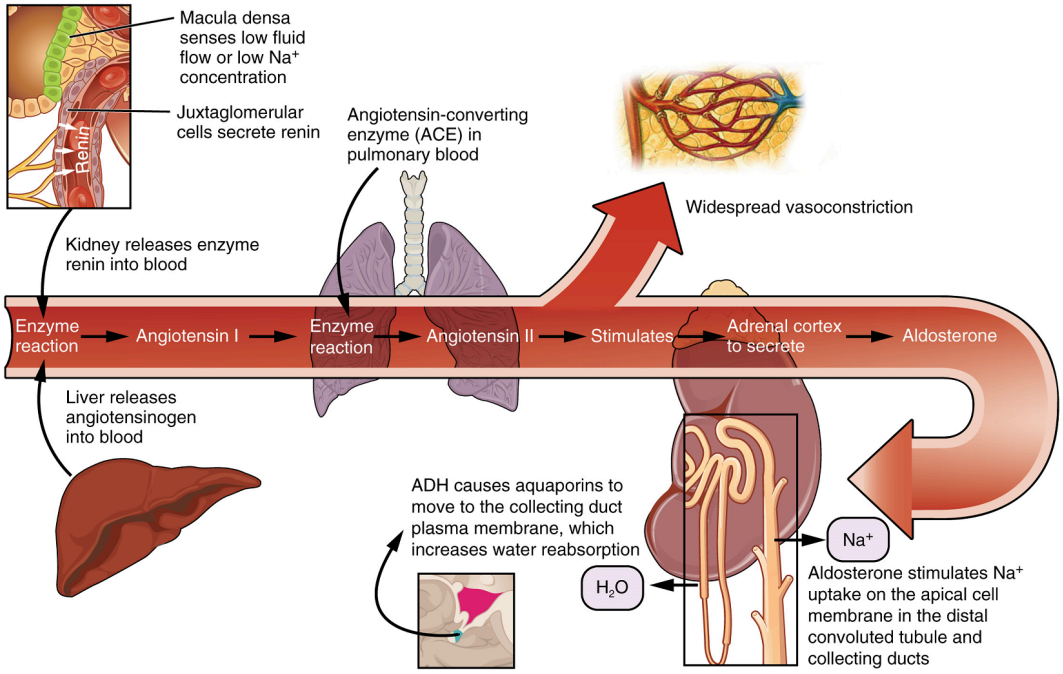


Figure 11. Conversion of Angiotensin I to Angiotensin II. *The enzyme renin converts the pro-enzyme angiotensin I; the lung-derived enzyme ACE converts angiotensin I into active angiotensin II.*

While antidiuretic hormone is primarily involved in the regulation of water recovery, aldosterone regulates Na^+ recovery. When aldosterone output increases, more Na+ is recovered from the forming urine and water follows the Na^+ passively. As the pump recovers Na^+ for the body, it is also pumping K^+ into the forming urine, since the pump moves K^+ in the opposite direction. When aldosterone decreases, more Na^+ remains in the forming urine and more K^+ is recovered in the circulation.

Part 4: Micturition Reflex

Micturition is a term for urination or voiding. It results from an interplay of involuntary and voluntary actions by the internal and external urethral sphincters. When bladder volume reaches about 150 mL, an urge to void is sensed but is easily overridden. Voluntary control of urination relies on consciously preventing relaxation of the external urethral sphincter to maintain urinary continence. As the bladder fills, subsequent urges become harder to ignore. Ultimately, voluntary constraint fails with resulting **incontinence**, which will occur as bladder volume approaches 300 to 400 mL.

Normal micturition is a result of stretch receptors in the bladder wall that transmit nerve impulses to the sacral region of the spinal cord to generate a spinal reflex. The resulting parasympathetic neural outflow causes contraction of the detrusor muscle and relaxation of the involuntary internal urethral sphincter. At the same time, the spinal cord inhibits somatic motor neurons, resulting in the relaxation of the skeletal muscle of the external urethral sphincter. The micturition reflex is active in infants but with maturity, children learn to override the reflex by asserting external sphincter control, thereby delaying voiding.

Voluntary micturition requires an intact spinal cord and functional pudendal nerve arising from the **sacral micturition center**. Since the external urinary sphincter is voluntary skeletal muscle, it remains contracted (and thereby maintains continence) during filling of the bladder. At the same time, sympathetic nervous activity suppresses contraction of the detrusor muscle. With further bladder stretch, afferent signals traveling over sacral pelvic nerves activate parasympathetic neurons. This activates efferent neurons to release acetylcholine at the neuromuscular junctions, producing detrusor contraction and bladder emptying.

Part 5: Physical Characteristics of Urine

Characteristics of the urine change, depending on influences such as water intake, exercise, environmental temperature, nutrient intake, and other factors (Table 2). Some of the characteristics such as color and odor are rough descriptors of your state of hydration. For example, if you exercise or work outside, and sweat a great deal, your urine will turn darker and produce a slight odor, even if you drink plenty of water.

Urinalysis (urine analysis) often provides clues to renal disease. Normally, only traces of protein are found in urine, and when higher amounts are found, damage to the glomeruli is the likely basis. Unusually large quantities of urine may point to diseases like diabetes mellitus or hypothalamic tumors that cause diabetes insipidus. The color of urine is determined mostly by the breakdown products of red blood cell destruction (Figure 12). The “heme” of hemoglobin is converted by the liver into water-soluble forms that can be excreted indirectly into the urine. Urine color may also be affected by certain foods like beets, berries, and fava beans. A kidney stone or a cancer of the urinary system may produce enough bleeding to manifest as pink or even bright red urine.

Table 2: Normal Urine Characteristics

Characteristic	Normal values
Colour	Pale yellow to deep amber
Odour	Odourless
Volume	750-2000 mL / 24 h
pH	4.5-8.0
Specific gravity	1.003-1.032
Osmolarity	40-1350 mOsmol / kg
Urobilinogen	0.2-1.0 mg / 100 mL
Leukocytes	0-2 HPF (per “high-power field” of microscope)
Leukocyte esterase	None
Protein	None or trace
Bilirubin	< 0.3 mg / 100 mL
Ketones	None
Nitrites	None
Erythrocytes	None
Glucose	None

Diseases of the liver or obstructions of bile drainage from the liver impart a dark “tea” or “cola” hue to the urine. Dehydration produces darker, concentrated urine that may also possess the slight odor of ammonia. Most of the ammonia produced from protein breakdown is converted into urea by the liver, so ammonia is rarely detected in fresh urine. Certain foods such as onions, garlic, and fish can impart odor to the urine.

Urine volume varies considerably. The normal range is one to two liters per day (Table 3). The kidneys must produce a minimum urine volume of about 500 mL/day to rid the body of wastes. Output below this level may be caused by severe dehydration or renal disease and is termed **oliguria**. The virtual absence of urine production is termed **anuria**.

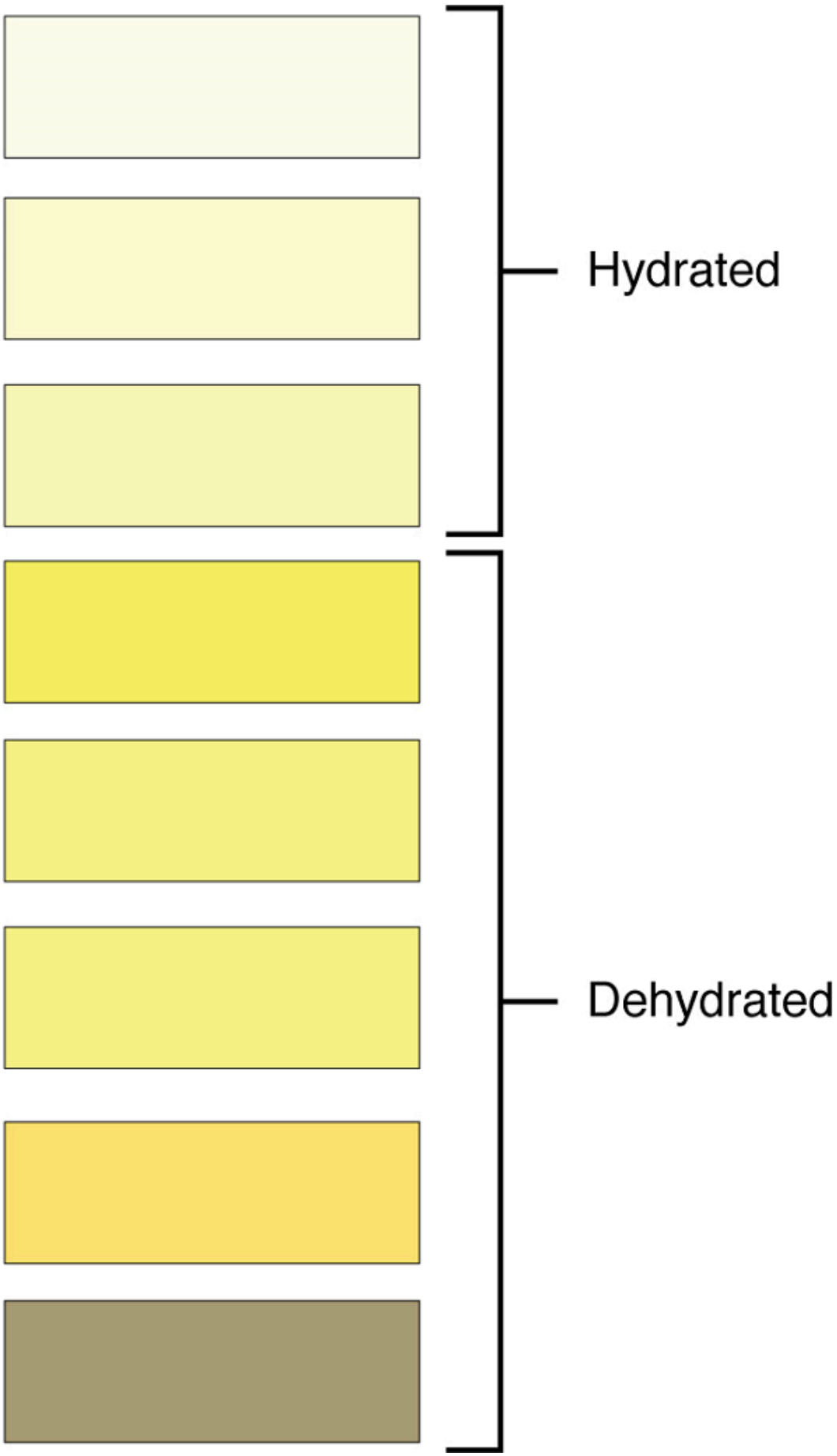


Figure 12. Urine Color.

Excessive urine production is **polyuria**, which may be due to diabetes mellitus or diabetes insipidus. In diabetes mellitus, blood glucose levels are high resulting in the appearance of glucose in the urine. Glucose attracts water osmotically, leading to its loss in the urine. In the case of diabetes insipidus, defficiency of antidiuretic hormone (ADH) leads to high volumes of very dilute urine.

Table 3: Urine Volumes

Volume condition	Volume	Causes
Normal	1-2 L/day	
Polyuria	>2.5 L/day	Diabetes mellitus; diabetes insipidus; excess caffeine or alcohol; kidney disease; certain drugs, e.g. diuretics; sickle cell anemia, excessive water intake
Oliguria	300-500 mL/day	Dehydration; blood loss; diarrhea; cardiogenic shock; kidney disease; enlarged prostate
Anuria	<50 mL/day	Kidney failure; obstruction, e.g. kidney stone or tumor; enlarged prostate

The pH (hydrogen ion concentration) of the urine can vary more than 1000-fold, from a normal low of 4.5 to a maximum of 8.0. Diet can influence pH; meats lower the pH, whereas citrus fruits, vegetables, and dairy products raise the pH. Chronically high or low pH can lead to disorders, such as the development of kidney stones or osteomalacia (softening of the bones).

Specific gravity is a measure of the quantity of solutes per unit volume of a solution. Urine will always have a specific gravity greater than pure water (water = 1.0) due to the presence of solutes.

Cells are not normally found in the urine. The presence of leukocytes may indicate a urinary tract infection.

Protein does not normally leave the glomerular capillaries, so only trace amounts of protein should be found in the urine, approximately 10 mg/100 mL in a random sample. If excessive protein is detected in the urine, it usually means that the glomerulus is damaged and is allowing protein to "leak" into the filtrate.

Ketones are byproducts of fat metabolism. Finding ketones in the urine suggests that the body is using fat as an energy source in preference to glucose. In diabetes mellitus when there is not enough insulin (type I diabetes mellitus) or because of insulin resistance (type II diabetes mellitus), there is plenty of glucose, but without the action of insulin, the cells cannot take it up, so it remains in the bloodstream. Instead, the cells are forced to use fat as their energy source, and fat consumed at such a level produces excessive ketones as byproducts. These excess ketones will appear in the urine. Ketones may also appear if there is a severe deficiency of proteins or carbohydrates in the diet. Nitrates (NO_3^-) occur normally in the urine.

Watch this Crash Course video to learn about the urinary system and the characteristics of urine! Direct link: https://youtu.be/DlqyyyvTI3k

Unit 9: Fluids & Electrolytes

Unit outline

Part 1: Body Fluids and Fluid Compartments

- Body water content
- Fluid compartments
    - Intracellular fluid
    - Extracellular fluid
- Electrolyte balance
- Roles of electrolytes

Part 2: Acid-Base Balance

- Acids
- Bases
- The concept of pH
- Buffers
- Homeostatic imbalance of acids and bases
    - Acid-base balance
    - Buffer systems in the body
    - Respiratory regulation of acid-base balance
    - Renal regulation of acid-base balance
- Disorders of acid-base balance
    - Metabolic acidosis: primary bicarbonate deficiency
    - Respiratory acidosis: primary carbonic acid/CO_2 excess
    - Respiratory alkalosis: primary carbonic acid/CO_2 deficiency
- Compensation Mechanisms
    - Respiratory compensation
    - Metabolic compensation

Learning Objectives

At the end of this unit, you should be able to:

I. Define and describe the measurement of pH.

II. Describe the electrolyte composition of the body.

III. Describe the mechanisms for the control of electrolytes in the body.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Define and describe the measurement of pH.

II. Describe the electrolyte composition of the body.

III. Describe the mechanisms for the control of electrolytes in the body.

1. Describe the difference between compounds that are electrolytes and those that are non-electrolytes.
2. Specify four functions of ions in the body.
3. Compare and contrast the electrolyte composition of each of the following pairs of body fluid compartments:
    - Intracellular fluid and blood plasma
    - Intracellular fluid and interstitial fluid
    - Blood plasma and interstitial fluid
4. Of the three major fluid compartments in the human body (intracellular fluid, blood plasma, and interstitial fluid), which two are most similar to each other? Explain why this grouping is not unexpected.
5. Define the following terms: electrolyte, acid, base, buffer.
6. Define and clearly distinguish between an electrolyte, an acid, and a base.
7. Explain the chemical nature of a buffer solution. Describe the function of such solutions in the human body and explain how they perform that function.
8. Describe pH and the pH scale.
9. Describe how to use the pH scale to determine hydrogen ion concentration, and how to use it to compare two solutions, by completing the following questions:
    - Explain in words how to calculate the difference in hydrogen ion concentration between a solution with a pH of 4 and one with a pH of 9.
    - State whether each solution described above is acidic, basic or neutral.
    - For each of the following solutions, state (a) whether it is acidic or basic, AND (b) whether it is “more acidic” or “less acidic” than a solution with a pH of 7.3:
        - A solution with a pH of 7.2

- A solution with a pH of 6.4
- A solution with a pH of 7.8

10. Describe the sources of acids and bases in the body.
11. Describe the mechanism(s) by which acids and/or bases arise in the human body as a result of each of the following:

    - Introduction of regular foodstuffs into the digestive tract
    - Ingestion of an antacid tablet
    - Anaerobic conditions (lactic acid fermentation)
    - Aerobic cellular respiration
    - Amino acid catabolism
    - Triglyceride catabolism
    - Gluconeogenesis

12. Describe three mechanisms by which the body regulates its pH.
13. Describe the carbonic acid-bicarbonate buffer system and specify two other buffer systems.
14. Define acidosis and specify two general causes of acidosis.
15. Define alkalosis and specify two general causes of alkalosis.
16. Compare and contrast (describe similarities *and* differences between):

    - Acidosis and alkalosis
    - Respiratory acidosis and respiratory alkalosis
    - Metabolic acidosis and metabolic alkalosis
    - Respiratory acidosis and metabolic acidosis
    - Metabolic alkalosis and respiratory alkalosis

17. Specify at least two causative disorders or circumstances that would cause each of the following and explain how they do so.

    - Respiratory alkalosis
    - Respiratory acidosis
    - Metabolic alkalosis
    - Metabolic acidosis

Part 1: Body Fluids and Fluid Compartments

The chemical reactions of life take place in aqueous solutions. The dissolved substances in a solution are called solutes. In the human body, solutes vary in different parts of the body, but may include proteins—including those that transport lipids, carbohydrates, and very importantly, electrolytes.

In the body, water moves through semi-permeable membranes of cells and from one compartment of the body to another by osmosis. An appropriate balance of solutes inside and outside of cells must be maintained to ensure normal function.

Body Water Content: Human beings are mostly water, ranging from about 75 percent of body mass in infants to about 50–60 percent in adult men and women, to as low as 45 percent in old age.

Fluid Compartments: Body fluids can be discussed in terms of their specific fluid compartment, a location that is largely separate from another compartment by some form of a physical barrier. The intracellular fluid

(ICF) compartment is the system that includes all fluid enclosed in cells by their plasma membranes. Extracellular fluid (ECF) surrounds all cells in the body. Extracellular fluid has two primary constituents: the fluid component of the blood (called plasma) and the interstitial fluid (IF) that surrounds all cells not in the blood (Figure 1).

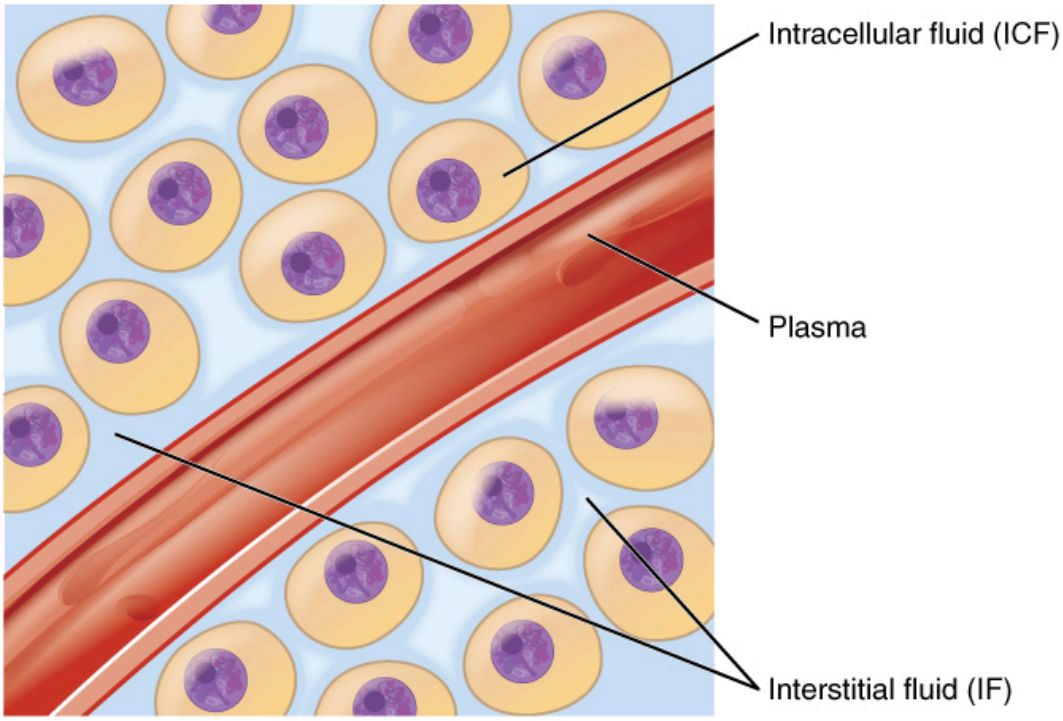


Figure 1. Fluid Compartments in the Human Body. *The intracellular fluid (ICF) is the fluid within cells. The interstitial fluid (IF) is part of the extracellular fluid (ECF) between the cells. Blood plasma is the second part of the extracellular fluid. Materials travel between cells and the plasma in capillaries through the intracellular fluid.*

Intracellular Fluid: The intracellular fluid lies within cells and is the principal component of the cytosol. The intracellular fluid makes up about 60 percent of the total water in the human body (Figure 2). The amount of water in living cells is closely regulated. If the amount of water inside a cell falls to a value that is too low, the cytosol becomes too concentrated with solutes to carry on normal cellular activities; if too much water enters a cell, the cell may burst and be destroyed.

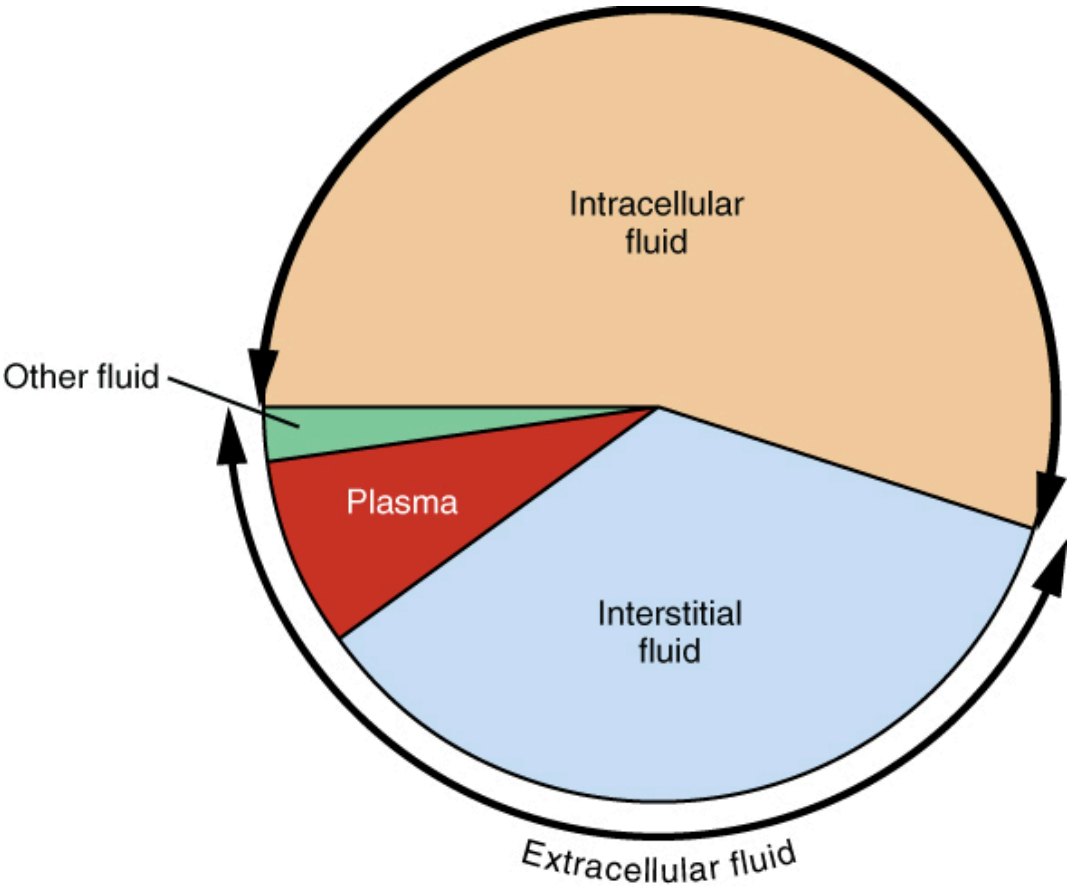


Figure 2. Proportions of Total Body Fluid in Each of the Body's Fluid Compartments. *Most of the water in the body is intracellular fluid. The second largest volume is the interstitial fluid, which surrounds cells that are not blood cells.*

Extracellular Fluid: The extracellular fluid accounts for the other one-third of the body's water content. Approximately 20 percent of the extracellular fluid is found in plasma. Plasma travels through the body in blood vessels and transports a range of materials, including blood cells, proteins (including clotting factors and antibodies), electrolytes, nutrients, gases, and wastes. Gases, nutrients, and waste materials travel between capillaries and cells through the interstitial fluid. Cells are separated from the interstitial fluid by a selectively permeable cell membrane that helps regulate the passage of materials between the interstitial fluid and the interior of the cell.

The body has other water based extracellular fluid. These include the cerebrospinal fluid that bathes the brain and spinal cord, lymph, the synovial fluid in joints, the pleural fluid in the pleural cavities, the pericardial fluid in the cardiac sac, the peritoneal fluid in the peritoneal cavity, and the aqueous humor of the eye. Because these fluids are outside of cells, these fluids are also considered components of the extracellular fluid compartment.

Composition of Body Fluids: The compositions of the two components of the extracellular fluid—plasma and interstitial fluid—are more similar to each other than either is to the intracellular fluid (Figure 3). Blood plasma has high concentrations of sodium, chloride, bicarbonate, and protein. The interstitial fluid has high concentrations of sodium, chloride, and bicarbonate, but a relatively lower concentration of protein. In contrast, the intracellular fluid has elevated amounts of potassium, phosphate, magnesium, and protein. Overall, the intracellular fluid contains high concentrations of potassium and phosphate (HPO_4^{2-}), whereas both plasma and the extracellular fluid contain high concentrations of sodium and chloride.

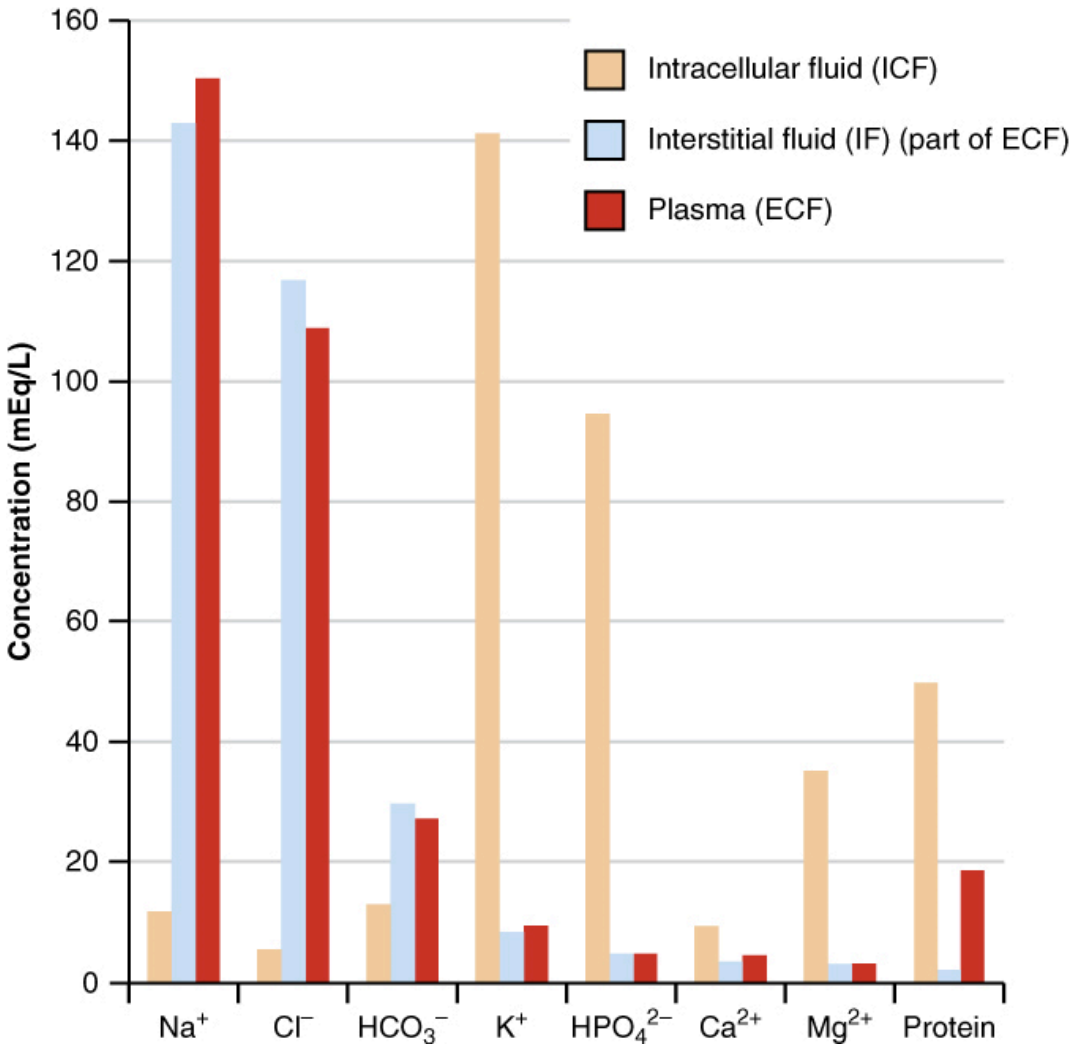


Figure 3. The Concentrations of Different Ions in Key Bodily Fluids. *The graph shows the composition of the intracellular fluid and extracellular fluid (interstitial fluid and plasma). The compositions of plasma and interstitial fluid are similar to one another but are quite different from the composition of the intracellular fluid.*

Body fluids are neutral in charge. Thus, cations, or positively charged ions, and anions, or negatively charged ions, are balanced in fluids. As seen in the previous graph, sodium (Na^+) ions and chloride (Cl^-) ions are concentrated in the extracellular fluid of the body, whereas potassium (K^+) ions are concentrated inside cells. Although sodium and potassium can "leak" through "pores" into and out of cells, respectively, the high levels of potassium and low levels of sodium in the intracellular fluid are maintained by sodium-potassium pumps in the cell membranes. These pumps use the energy supplied by ATP to pump sodium out of the cell and potassium into the cell (Figure 4).

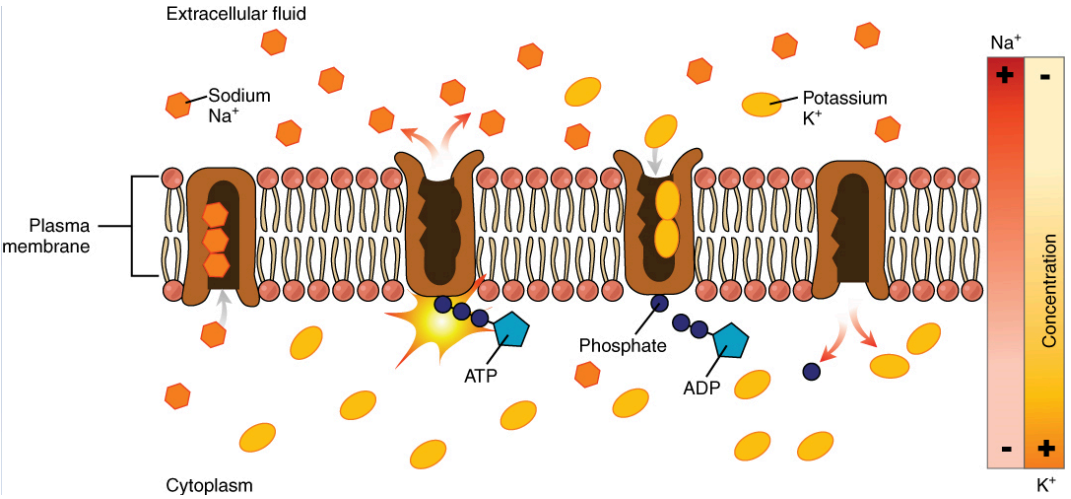


Figure 4. The Sodium-Potassium Pump. *The sodium-potassium pump is powered by ATP to transfer sodium out of the cytoplasm and into the extracellular fluid. The pump also transfers potassium out of the extracellular fluid and into the cytoplasm. (credit: modification of work by Mariana Ruiz Villarreal).*

Roles of Electrolytes: The body contains a large variety of ions, or electrolytes, which perform a variety of functions. Some ions assist in the transmission of electrical impulses along cell membranes in neurons and muscles. Other ions help to stabilize protein structures in enzymes. Still others aid in releasing hormones from endocrine glands. All of the ions in plasma contribute to the osmotic balance that controls the movement of water between cells and their environment.

Electrolytes in living systems include sodium, potassium, chloride, bicarbonate, calcium, phosphate, magnesium, copper, zinc, iron, manganese, molybdenum, copper, and chromium. In terms of body functioning, six electrolytes are most important: sodium, potassium, chloride, bicarbonate, calcium, and phosphate.

These six ions aid in nerve excitability, endocrine secretion, membrane permeability, buffering body fluids, and controlling the movement of fluids between compartments. These ions enter the body through the digestive tract. More than 90 percent of the calcium and phosphate that enters the body is incorporated into bones and teeth, with bone serving as a mineral reserve for these ions. In the event that calcium and phosphate are needed for other functions, bone tissue can be broken down to supply the blood and other tissues with these minerals. Phosphate is a normal constituent of nucleic acids; hence, blood levels of phosphate will increase whenever nucleic acids are broken down.

Excretion of ions occurs mainly through the kidneys, with lesser amounts lost in sweat and in feces. Excessive sweating may cause a significant loss, especially of sodium and chloride. Severe vomiting or diarrhea will cause a loss of chloride and bicarbonate ions. Adjustments in respiratory and renal functions allow the body to regulate the levels of these ions in the extracellular fluid.

Part 2: Acid-Base Balance

Acids and bases, like salts, dissociate in water into electrolytes. Acids and bases can very much change the properties of the solutions in which they are dissolved.

Acids: An acid is a substance that releases hydrogen ions (H^+) in solution (Figure 5a). Because an atom of hydrogen has just one proton and one electron, a positively charged hydrogen ion is simply a proton. This solitary proton is highly likely to participate in chemical reactions. Strong acids are compounds that release all of their H^+ in solution; that is, they ionize completely. Hydrochloric acid (HCl), which is released from cells in the lining of the stomach, is a strong acid because it releases all of its H^+ in the stomach's watery environment. This strong acid aids in digestion and kills ingested microbes. Weak acids do not ionize completely; that is, some of their Hydrogen ions remain bonded within a compound in solution. An example of a weak acid is vinegar, or acetic acid; it is called acetate after it gives up a proton. Other common examples used in cellular metabolism

include pyruvic acid, citric acid, and oxaloacetic acid which in water may release a proton to become pyruvate, citrate, and oxaloacetate respectively.

Bases: A base is a substance that releases hydroxyl ions (OH^-) in solution, or one that accepts H^+ already present in solution (Figure 5b). The hydroxyl ions (also known as hydroxide ions) or other basic substances combine with H^+ present to form a water molecule, thereby removing H^+ and reducing the solution's acidity. Strong bases release most or all of their hydroxyl ions; weak bases release only some hydroxyl ions or absorb only a few H^+

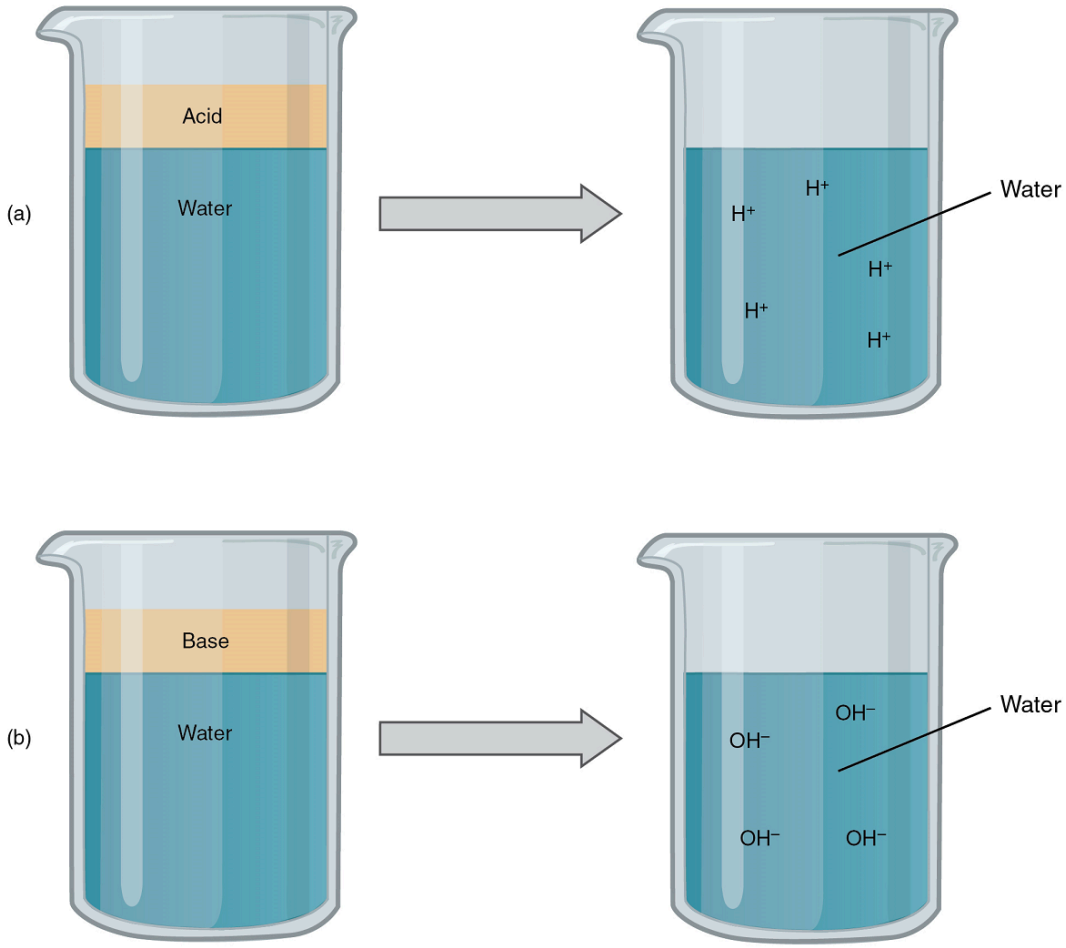


Figure 5. Acids and Bases. *(a) In aqueous solution, an acid dissociates into hydrogen ions (H^+) and anions. Nearly every molecule of a strong acid dissociates, producing a high concentration of H^+. (b) In aqueous solution, a base dissociates into hydroxyl ions (OH^-) and cations. Nearly every molecule of a strong base dissociates, producing a high concentration of OH^-.*

The Concept of pH: The relative acidity or alkalinity of a solution can be indicated by its pH. pH literally means the "potential of hydrogen". It is a measure of the amount of hydrogen ions present per litre of a solution, expressed in grams. In technical terms, pH is the logarithm of the reciprocal of the hydrogen ion concentration of a solution.

As an example, a pH 4 solution has an H^+ concentration that is ten times greater than that of a pH 5 solution. That is, a solution with a pH of 4 is ten times more acidic than a solution with a pH of 5. The concept of pH will begin to make more sense when you study the pH scale (Figure 6). The scale consists of a series of increments ranging from 0 to 14. A solution with a pH of 7 is considered neutral—neither acidic nor basic. Pure water has a pH of 7. The lower the number below 7, the more acidic the solution, or the greater the concentration of H^+. The concentration of hydrogen ions at each pH value is 10 times different than the next pH. The higher the number above 7, the more basic (alkaline) the solution, or the lower the concentration of H^+.

Buffers: The pH of human blood normally ranges from 7.35 to 7.45. At this slightly basic pH, blood can reduce

the acidity resulting from the carbon dioxide (CO_2) constantly being released into the bloodstream by the trillions of cells in the body. Homeostatic mechanisms (along with exhaling CO_2 while breathing) normally keep the pH of blood within this narrow range. This is critical, because fluctuations—either too acidic or too alkaline—can lead to life-threatening disorders.

All cells of the body depend on homeostatic regulation of acid–base balance at a very narrow range of pH between 7.35 to 7.45. The body therefore has several mechanisms for this regulation, involving breathing, the excretion of chemicals in urine, and the internal release of chemicals collectively called buffers into body fluids. A buffer is a solution of a weak acid and its conjugate base. A buffer can resist sudden changes in the acidity and alkalinity of the body fluids. For example, if there is even a slight decrease below 7.35 in the pH of a bodily fluid, the buffer in the fluid—in this case, acting as a weak base—will bind the excess hydrogen ions.

In contrast, if pH rises above 7.45, the buffer will act as a weak acid and contribute hydrogen ions.

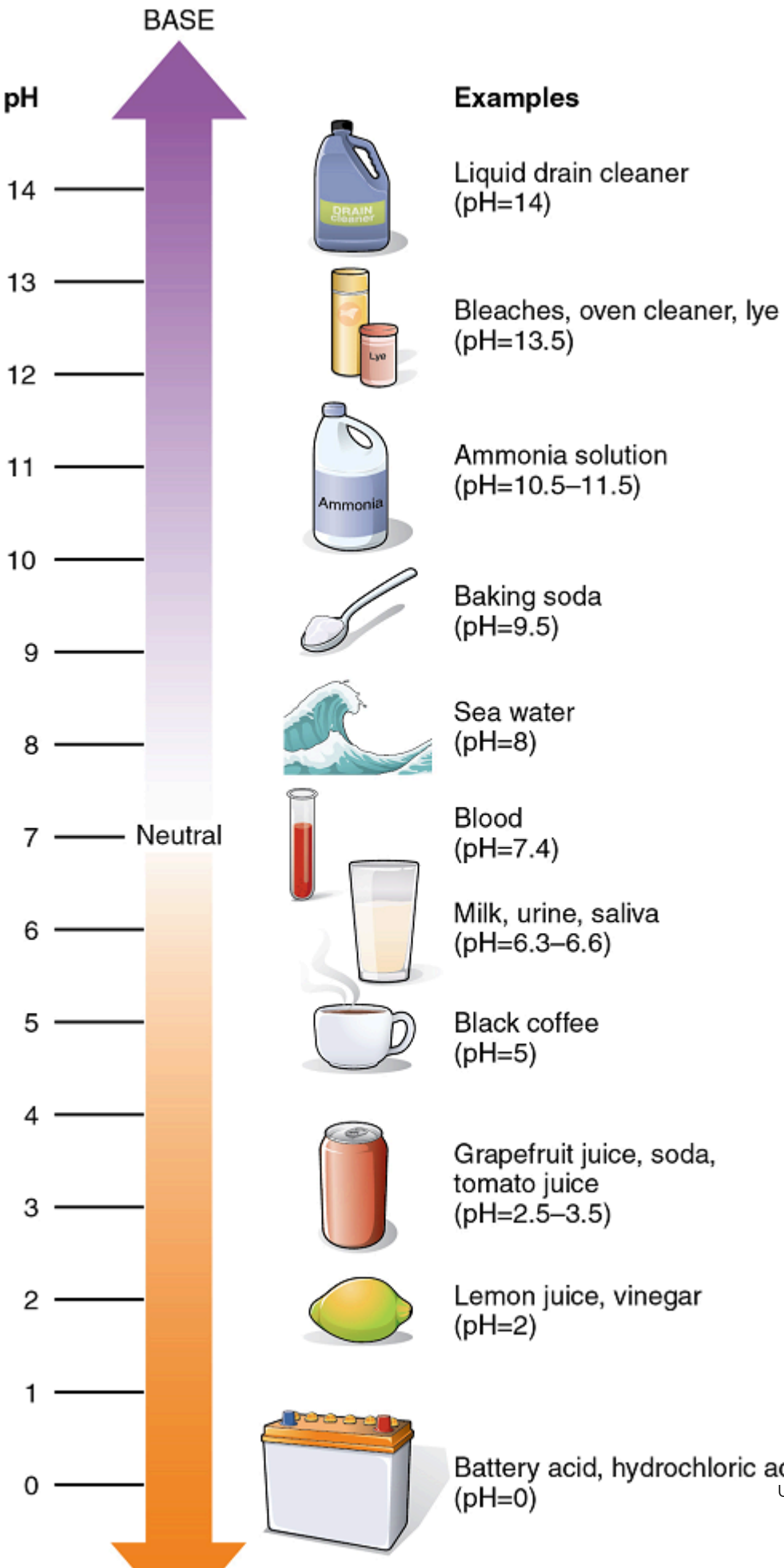


Figure 6. The pH Scale.

Acid-Base Balance: Proper physiological functioning depends on a very tight balance between the concentrations of acids and bases in the blood. Acid-base balance is measured using the pH scale (Figure 6). A variety of buffering systems permits blood and other bodily fluids to maintain a narrow pH range, even in the face of perturbations. A buffer is a chemical system that minimizes change in hydrogen ion concentration.

Buffer Systems in the Body: The buffer systems in the human body are extremely efficient, and different systems work at different rates. It takes only seconds for the chemical buffers in the blood to resist changes in the pH. The respiratory tract can adjust the blood pH upward in minutes by exhaling CO_2 from the body. The renal system can also adjust blood pH through the excretion of hydrogen ions (H^+) and the conservation of bicarbonate, but this process takes hours to days to have an effect.

The buffer systems functioning in blood plasma include plasma proteins, phosphate, and bicarbonate and carbonic acid buffers. The kidneys help control acid-base balance by excreting hydrogen ions and generating bicarbonate that helps maintain blood plasma pH within a normal range.

Protein Buffers in Blood Plasma and Cells: Protein buffer systems work predominantly inside cells. Nearly all proteins can function as buffers. Proteins are made up of amino acids, which contain positively charged amino groups and negatively charged carboxyl groups. The charged regions of free amino acids can bind hydrogen and hydroxyl ions, and thus function as buffers. Buffering by proteins accounts for two-thirds of the buffering power of the blood and most of the buffering within cells.

Phosphate Buffer: Phosphates are found in the blood in two forms: sodium dihydrogen phosphate (NaH_2PO_4), which is a weak acid, and sodium monohydrogen phosphate (Na_2HPO_4), which is a weak base. When Na_2HPO_4 comes into contact with a strong acid, such as HCl, the base reacts with the hydrogen ion released by the HCl to form the weak acid NaH_2PO_4 and sodium chloride (NaCl). When NaH_2PO_4 (the weak acid) comes into contact with a strong base, such as sodium hydroxide (NaOH), the weak acid releases H^+ ions which bind to the OH^- ions released by the base to produce water.

$$HCl + Na_2HPO_4 \rightarrow NaH_2PO_4 + NaCl$$

(strong acid) + (weak base) → (weak acid) + (a salt)

$$NaOH + NaH_2PO_4 \rightarrow Na_2HPO_4 + H_2O$$

(strong base) + (weak acid) → (weak base) + (water)

Bicarbonate-Carbonic Acid Buffer: The bicarbonate-carbonic acid buffer works in a fashion similar to phosphate buffers. When sodium bicarbonate ($NaHCO_3$), comes into contact with a strong acid, such as HCl, carbonic acid (H_2CO_3), which is a weak acid, and NaCl are formed. When carbonic acid comes into contact with a strong base, such as NaOH, bicarbonate and water are formed.

$$NaHCO_3 + HCl \rightarrow H_2CO_3 + NaCl$$

(sodium bicarbonate) + (strong acid) → (weak acid) + (a salt)

$$H_2CO_3 + NaOH \rightarrow NaHCO_3 + H_2O$$

(weak acid) + (strong base) → (sodium bicarbonate) + (water)

As with the phosphate buffer, a weak acid or weak base captures the free ions, and a significant change in pH is prevented. Bicarbonate ions and carbonic acid are present in the blood in a 20:1 ratio if the blood pH is within the normal range. With 20 times more bicarbonate than carbonic acid, this capture system is most efficient at buffering changes that would make the blood more acidic. This is useful because most of the body's metabolic wastes, such as lactic acid and ketone bodies, are acids. Carbonic acid levels in the blood are controlled by the expiration of CO_2 through the lungs. The level of bicarbonate in the blood is controlled through the renal system, where bicarbonate ions in the renal filtrate are conserved and passed back into the blood. However, the bicarbonate buffer is the primary buffering system of the interstitial fluid surrounding the cells in tissues throughout the body.

Respiratory Regulation of Acid-Base Balance: The respiratory system contributes to the balance of acids and

bases in the body by regulating the blood levels of carbonic acid (Figure 7). CO_2 in the blood readily reacts with water to form carbonic acid, and the levels of CO_2 and carbonic acid in the blood are in equilibrium. When the CO_2 level in the blood rises (as it does when you hold your breath), the excess CO_2 reacts with water to form additional carbonic acid, lowering blood pH. Increasing the rate and/or depth of respiration (which you might feel the "urge" to do after holding your breath) allows you to exhale more CO_2. The loss of CO_2 from the body reduces blood levels of carbonic acid and thereby adjusts the pH upward, toward normal levels. This process also works in the opposite direction, excessive deep and rapid breathing (as in hyperventilation) rids the blood of CO_2 and reduces the level of carbonic acid, making the blood too alkaline.

Renal Regulation of Acid-Base Balance: The renal regulation of the body's acid-base balance addresses the metabolic component of the buffering system. Whereas the respiratory system (together with breathing centers in the brain) controls the blood levels of carbonic acid by controlling the exhalation of CO_2, the renal system controls the blood levels of bicarbonate.

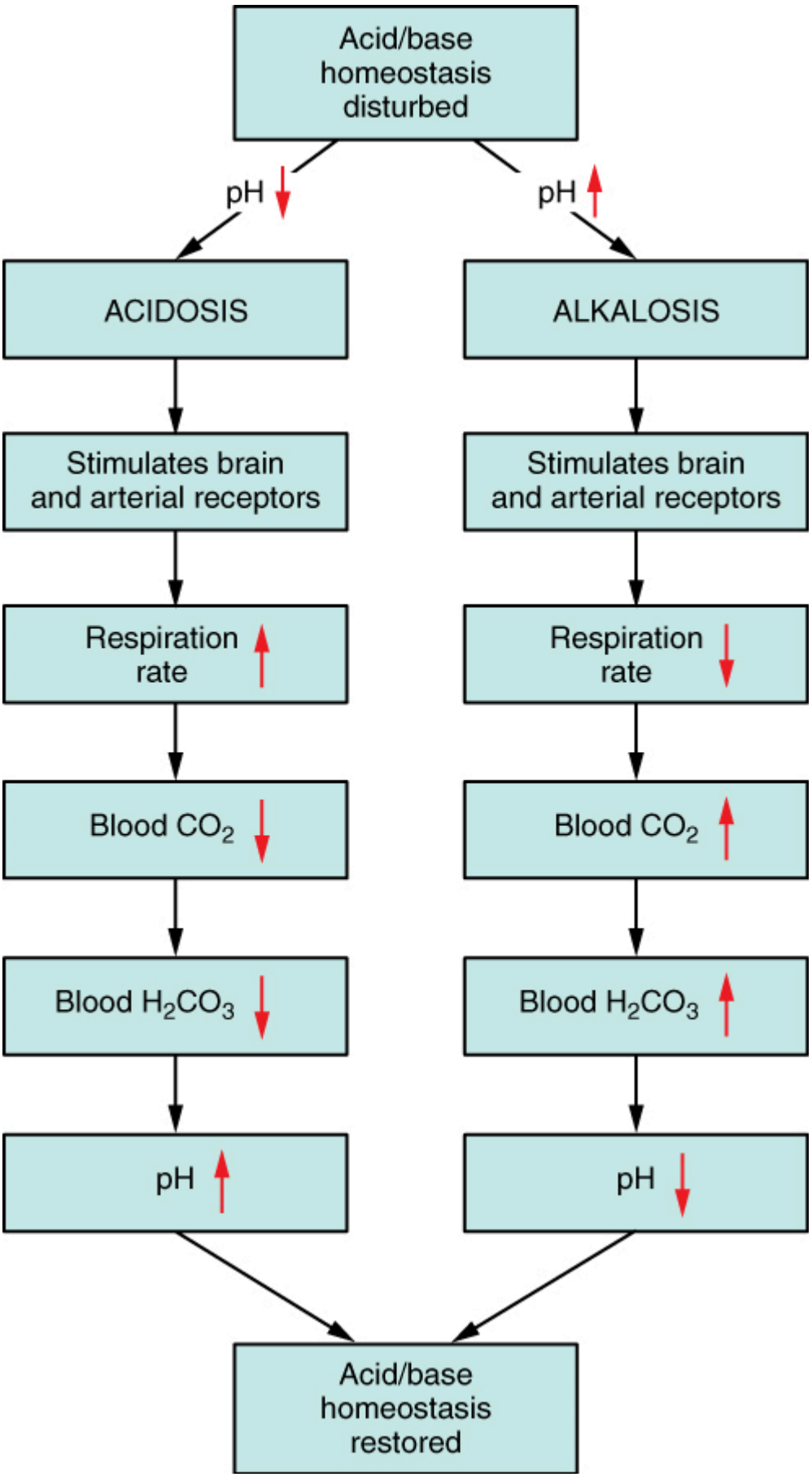


Figure 7. Respiratory Regulation of Blood pH. *The respiratory system can reduce blood pH by removing CO2 from the blood.*

A decrease of blood bicarbonate can result by certain diuretics or from excessive bicarbonate loss due to diarrhea. Low bicarbonate blood levels can also occur as a result of elevated levels of ketone bodies (common in unmanaged diabetes mellitus), which bind bicarbonate hence, lowering their concentration in the blood.

Disorders of Acid-Base Balance: Normal arterial blood pH is restricted to a very narrow range of 7.35 to 7.45. A person who has a blood pH below 7.35 is considered to be in acidosis, and a continuous blood pH below 7.0 can be fatal. Acidosis has several symptoms, including headache and confusion, and the individual can become lethargic and easily fatigued (Figure 8). A person who has a blood pH above 7.45 is considered to be in alkalosis, and a pH above 7.8 is fatal. Some symptoms of alkalosis include cognitive impairment (which can progress to unconsciousness), tingling or numbness in the extremities, muscle twitching and spasm, and nausea and vomiting. Both acidosis and alkalosis can be caused by either metabolic or respiratory disorders.

As discussed earlier in this chapter, the concentration of carbonic acid in the blood is dependent on the level of CO_2 in the body and the amount of CO_2 gas exhaled through the lungs. Thus, the respiratory contribution to acid-base balance is usually discussed in terms of CO_2 (rather than of carbonic acid). Remember that a molecule of carbonic acid is lost for every molecule of CO_2 exhaled, and a molecule of carbonic acid is formed for every molecule of CO_2 retained.

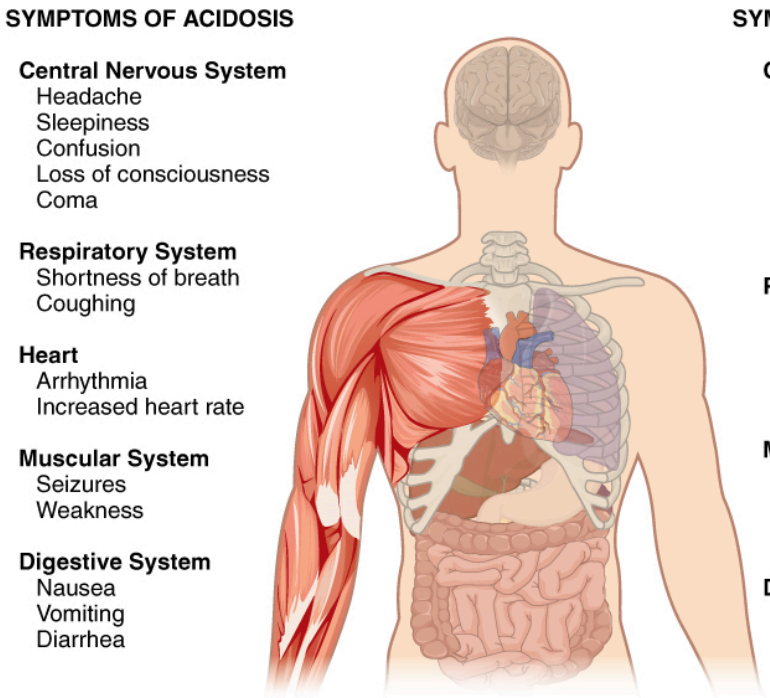


Figure 8. Symptoms of Acidosis and Alkalosis. *Symptoms of acidosis affect several organ systems. Both acidosis and alkalosis can be diagnosed using a blood test.*

Metabolic Acidosis: Primary Bicarbonate Deficiency: Metabolic acidosis occurs when the blood is too acidic (pH below 7.35) due to too little bicarbonate in the body fluids. At the normal pH of 7.40, the ratio of bicarbonate to carbonic acid buffer is 20:1. If a person's blood pH drops below 7.35, then he or she is in metabolic acidosis. The most common cause of metabolic acidosis is the presence of organic acids or excessive ketone bodies in the blood. Table 1 lists some other causes of metabolic acidosis.

The first three of the eight causes of metabolic acidosis listed are medical (or unusual physiological) conditions. Strenuous exercise can cause temporary metabolic acidosis due to the production of lactic acid. The last five causes result from the ingestion of specific substances. The active form of aspirin is its metabolite, acetylsalicylic acid. An overdose of aspirin causes acidosis due to the acidity of this drug. Metabolic acidosis can also result from uremia, which is the retention of urea and uric acid. Metabolic acidosis can also arise from diabetic ketoacidosis, wherein an excess of ketones is present in the blood. Other causes of metabolic

acidosis are a decrease in the excretion of hydrogen ions, which inhibits the conservation of bicarbonate ions, and excessive loss of bicarbonate ions through the gastrointestinal tract due to diarrhea.

Table 1: Common Causes of Metabolic Acidosis

Cause	Metabolite affected
Diarrhea	Bicarbonate
Uremia	Phosphoric, sulfuric, lactic acids
Diabetic ketoacidosis	Ketones
Strenuous exercise	Lactic acid
Methanol ingestion	Formic acid (metabolite of methanol)
Paraldehyde ingestion	β-Hydroxybutyric acid (metabolite of paraldehyde)
Isopropanol ingestion	Propionic acid (metabolite of isopropanol)
Ethylene glycol ingestion	Glycolic acid, oxalic & formic acids (metabolites of ethylene glycol)
Salicylate/ aspirin ingestion	Sulfasalicylic acid (metabolite of salicylate)

Metabolic Alkalosis: Primary Bicarbonate Excess: Metabolic alkalosis is the opposite of metabolic acidosis. It occurs when the blood is too alkaline (pH above 7.45) due to an excess of bicarbonate in the body fluids.

A transient excess of bicarbonate in the blood can follow ingestion of excessive amounts of bicarbonate, citrate, or antacids for conditions such as stomach acid reflux—known as heartburn. Other causes of metabolic alkalosis include the loss of hydrochloric acid from the stomach through vomiting, potassium depletion due to the use of diuretics for hypertension, and the excessive use of laxatives.

Respiratory Acidosis: Primary Carbonic Acid/CO_2 Excess: Respiratory acidosis occurs when the blood is overly acidic due to an excess of carbonic acid, resulting from too much CO_2 in the blood. Respiratory acidosis can result from anything that interferes with respiration, such as pneumonia, emphysema, or congestive heart failure.

Respiratory Alkalosis: Primary Carbonic Acid/CO_2 Deficiency: Respiratory alkalosis occurs when the blood is overly alkaline due to a deficiency in carbonic acid and CO_2 levels in the blood. This condition usually occurs when too much CO_2 is exhaled from the lungs, as occurs in hyperventilation, which is breathing that is deeper or more frequent than normal. An elevated respiratory rate leading to hyperventilation can be due to extreme emotional upset or fear, fever, infections, hypoxia, or abnormally high levels of catecholamines, such as epinephrine and norepinephrine.

Compensation Mechanisms: Various compensatory mechanisms exist to maintain blood pH within a narrow range, including buffers, respiration, and renal mechanisms. Although compensatory mechanisms usually work very well, when one of these mechanisms is not working properly (like kidney failure or respiratory disease), they have their limits. If the pH and bicarbonate to carbonic acid ratio are changed too drastically, the body may not

be able to compensate. Moreover, extreme changes in pH can denature proteins. Extensive damage to proteins in this way can result in disruption of normal metabolic processes, serious tissue damage, and ultimately death.

Respiratory Compensation: Respiratory compensation for metabolic acidosis increases the respiratory rate to drive off CO_2 and readjust the bicarbonate to carbonic acid ratio to the 20:1 level. This adjustment can occur within minutes. Respiratory compensation for metabolic alkalosis is not as adept as its compensation for acidosis. The normal response of the respiratory system to elevated pH is to increase the amount of CO_2 in the blood by decreasing the respiratory rate to conserve CO_2. The respiratory route is less efficient at compensating for metabolic alkalosis than for acidosis.

Metabolic Compensation: Metabolic and renal compensation for respiratory diseases that can create acidosis revolves around the conservation of bicarbonate ions. In cases of respiratory acidosis, the kidney increases the conservation of bicarbonate and secretion of H^+. These processes increase the concentration of bicarbonate in the blood, reestablishing the proper relative concentrations of bicarbonate and carbonic acid. In cases of respiratory alkalosis, the kidneys decrease the production of bicarbonate and conserve H^+ ions.

REPRODUCTION

Unit 10: Reproduction and Development

Unit outline

Part 1: Anatomy and Physiology of the Male Reproductive System

- Scrotum
- Testes
- Structure of formed sperm
- Sperm transport
- The penis
- Testosterone
- Function of testosterone

Part 2: Anatomy and Physiology of the Female Reproductive System

- External female genitalia
- Vagina
- Ovaries
- The ovarian cycle
- Hormonal control of the ovarian cycle
- The uterine tubes
- The uterus and cervix
- The menstrual cycle
- The breasts

Part 3: Fertilization

- Transit of sperm
- Contact between sperm and oocyte
- The zygote

Part 4: Embryonic Development

- Pre-implantation embryonic development
- Implantation
- Embryonic membranes
- Germ layers
- Development of the placenta

Part 5: Fetal Development

- The fetal circulatory system

Part 6: Maternal Changes During Pregnancy, Labor, and Birth

- Effects of hormones

- Physiology of labor
- Stages of childbirth

Part 7: Adjustments of the Infant at Birth and Postnatal Stages

- Reproductive adjustments
- Circulatory adjustments

Part 8: Lactation

- Structure of lactating breast
- The process of lactation
- Changes in composition of breast milk

Learning Objectives

At the end of this unit, you should be able to:

I. Describe the location, structure and function(s) of the components of the male reproductive system.

II. Describe the location, structure and function(s) of the components of the female reproductive system.

III. Describe the roles of FSH, LH, and testosterone or estrogens in the male and female reproductive systems

IV. Describe the mechanism, stages, and anatomical pathway of sperm release and ovum release

V. Describe the ovarian cycle as well as the role of the hypothalamus and anterior pituitary gland in this cycle.

VI. Describe the uterine (menstrual) cycle as well as the names and sources of any hormones involved.

VII. Correlate the ovarian and uterine (menstrual) cycles and explain their integrated hormonal regulation.

VIII. Describe the roles of the corpus luteum and placenta if pregnancy occurs.

IX. Distinguish between morula and blastocyst.

X. Identify the three primary germ layers and which body parts they develop into.

XI. Describe the formation and functions of each of the four extra-embryonic membranes.

XII. Describe the formation, structure, and functions of the placenta and umbilical cord.

XIII. Explain the mechanisms underlying the production of fraternal twins and identical twins.

XIV. Describe the endocrine regulation (estrogen, progesterone, prolactin, oxytocin) and maintenance of pregnancy, labor, and lactation.

XV. Name the three stages of labor and describe the events that occur during each stage.

XVI. Describe fetal circulation and the changes that occur in fetal circulation following delivery.

XVII. Describe the respiratory system of the neonate with respect to the importance of surfactant, fluid in the lungs, and initiation of ventilation.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Describe the location, structure and function(s) of each of the following components of the male reproductive system:

1. Testes
2. Epididymis
3. Vas deferens
4. Ejaculatory duct
5. Urethra
6. Seminal vesicles (seminal glands)
7. Prostate gland
8. Penis
9. Bulbourethral gland
10. Scrotum
11. Seminiferous tubules

II. Describe the location, structure and function(s) of each of the following components of the female reproductive system:

1. Ovaries
2. Uterine tubes
3. Uterus
4. Cervix
5. Clitoris
6. Vagina
7. Endometrium
8. Myometrium
9. Perimetrium
10. Bartholin's glands

III. Describe the roles of FSH, LH, and testosterone or estrogens in the male and female reproductive systems

IV. Describe the mechanism, stages, and anatomical pathway of sperm release and ovum release

V. Describe the ovarian cycle. Include in your description the name of the three stages of the cycle and the events that occur during each stage, as well as the role of the hypothalamus and anterior pituitary gland in this cycle.

VI. Describe the uterine (menstrual) cycle. Include in your description the name of the three stages of the cycle and the events that occur during each stage, as well as the names and sources of any hormones involved.

VII. Correlate the ovarian and uterine (menstrual) cycles and explain their integrated hormonal regulation by drawing a fully-annotated diagram showing the hormonal interactions between the ovarian and uterine cycles.

VIII. Describe the roles of the corpus luteum and placenta if pregnancy occurs.

IX. Distinguish between morula and blastocyst.

X. Identify the three primary germ layers and which body parts they develop into.

1. Name the primary germ layer from which each of the following structures develops:
    - Nervous system
    - Sensory organs
    - Epidermis
    - Hair
    - Nails
    - Skeleton
    - Muscles
    - Connective tissue
    - Heart
    - Blood vessels
    - Kidneys
    - Gastrointestinal tract lining
    - Liver
    - Pancreas
    - Lungs

XI. Describe the formation and functions of each of the four extra-embryonic membranes.

XII. Describe the formation, structure, and functions of the placenta and umbilical cord.

XIII. Explain the mechanisms underlying the production of fraternal twins and identical twins.

XIV. Describe the endocrine regulation (estrogen, progesterone, prolactin, oxytocin) and maintenance of pregnancy, labor, and lactation.

XV. Name the three stages of labor and describe the events that occur during each stage.

XVI. Describe fetal circulation and the changes that occur in fetal circulation following delivery.

1. Name all the vessels or structures found in the fetal circulation but not the adult, and for each one

name the structure they develop into after birth.

XVII. Describe the respiratory system of the neonate with respect to the importance of surfactant, fluid in the lungs, and initiation of ventilation.

Small, uncoordinated, and slick with amniotic fluid, a newborn encounters the world outside of her mother's womb. We do not often consider that a child's birth is proof of the healthy functioning of both her mother's and father's reproductive systems. Moreover, her parents' endocrine systems had to secrete the appropriate regulating hormones to induce the production and release of unique male and female gametes, reproductive cells containing the parents' genetic material (one set of 23 chromosomes). Her parent's reproductive behavior had to facilitate the transfer of male gametes—the sperm—to the female reproductive tract at just the right time to encounter the female gamete, an oocyte (egg). Finally, combination of the gametes (fertilization) had to occur, followed by implantation and development. In this unit, you will explore the male and female reproductive systems, whose healthy functioning can culminate in the powerful sound of a newborn's first cry.

Part 1: Anatomy and Physiology of the Male Reproductive System

Unique for its role in human reproduction, a **gamete** is a specialized sex cell carrying 23 chromosomes—one half the number in body cells. At fertilization, the chromosomes in one male gamete, called a **sperm** (or spermatozoon), combine with the chromosomes in one female gamete, called an oocyte. The function of the male reproductive system (Figure 1) is to produce sperm and transfer them to the female reproductive tract. The paired testes are a crucial component in this process, as they produce both sperm and androgens, the hormones that support male reproductive physiology. In humans, the most important male androgen is testosterone. Several accessory organs and ducts aid the process of sperm maturation and transport the sperm and other seminal components to the penis, which delivers sperm to the female reproductive tract. In this section, we examine each of these different structures, and discuss the process of sperm production and transport.

Watch this CrashCourse video for an overview of the male reproductive system! Direct link: https://youtu.be/-XQcnO4iX_U

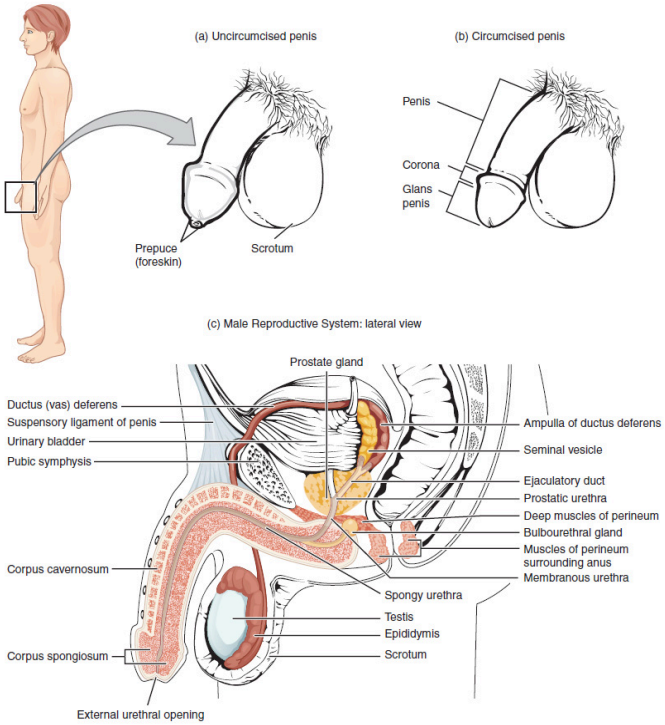
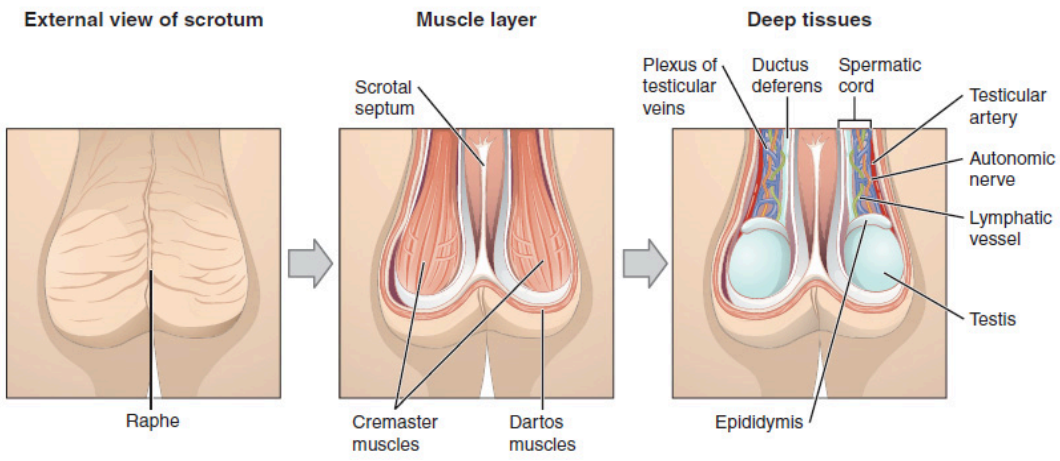


Figure 1. Male Reproductive System. *The structures of the male reproductive system include the testes, the epididymides, the penis, and the ducts and glands that produce and carry semen. Sperm exit the scrotum through the ductus deferens, which is bundled in the spermatic cord. The seminal vesicles and prostate gland add fluids to the sperm to create semen.*

Scrotum: The testes are located in a skin-covered, highly pigmented, muscular sac called the **scrotum** that extends from the body behind the penis (Figure 1). This location is important in sperm production, which occurs within the testes, and proceeds more efficiently when the testes are kept 2 to 4°C below core body temperature.

Figure 2. The Scrotum and Testes. *This anterior view shows the structures of the scrotum and testes.*



Testes: The **testes** (singular = testis) are the male **gonads**. They produce both sperm and androgens, such as testosterone, and are active throughout the reproductive lifespan of the male.

Paired ovals, the testes are each approximately 4 to 5 cm in length and are housed within the scrotum (Figure 2). Within the testes, sperm develop in structures called seminiferous tubules. During the seventh month of the developmental period of a male fetus, each testis moves through the abdominal musculature to descend into the scrotal cavity. This is called the "descent of the testis."

The tightly coiled **seminiferous tubules** form the bulk of each testis. They are composed of developing sperm

cells surrounding a lumen, the hollow center of the tubule, where formed sperm are released into the duct system of the testis (Figure 3). Specifically, from the lumens of the seminiferous tubules, sperm move into the straight tubules then into a fine meshwork of tubules and then leave the testis itself.

Inside the seminiferous tubules are supporting cells and developing sperm cells called germ cells. Germ cell development progresses from the basement membrane—at the perimeter of the tubule—toward the lumen.

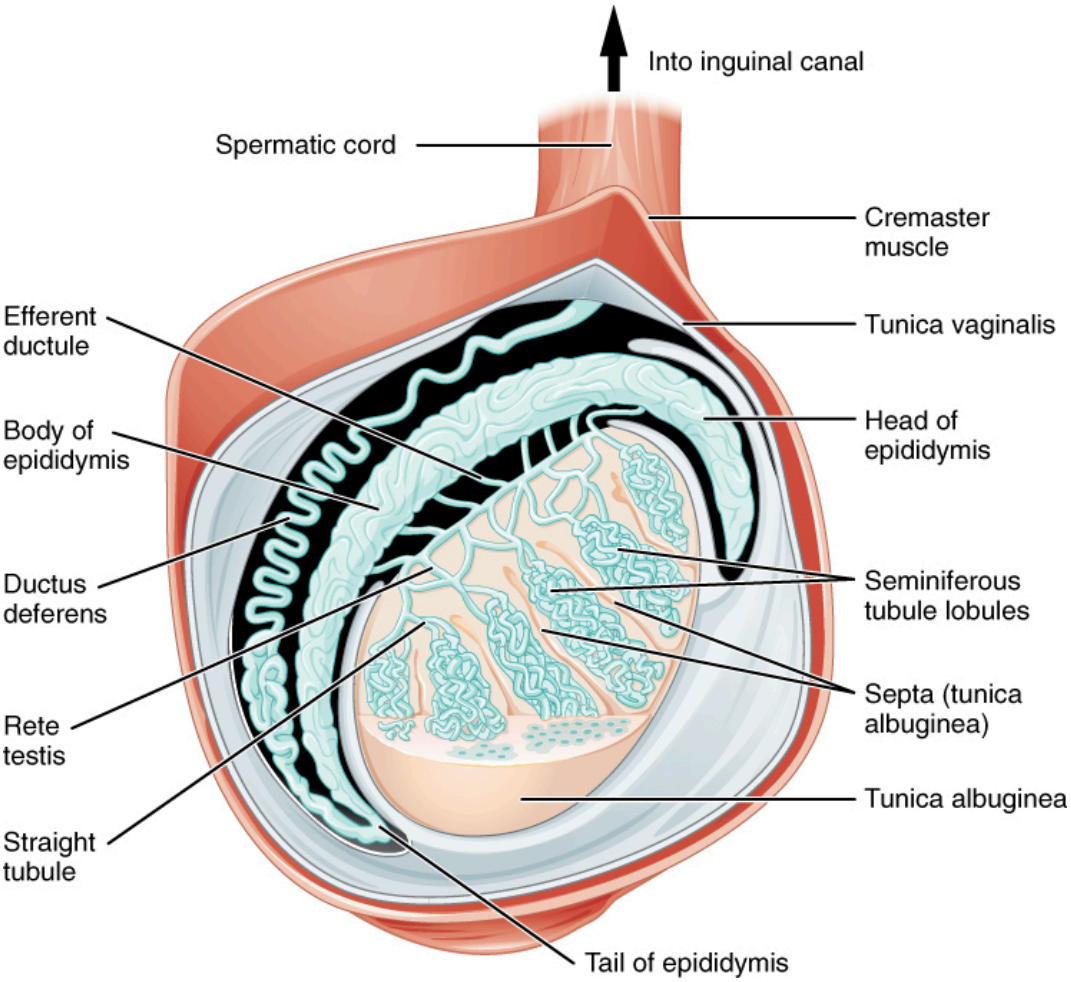


Figure 3. Anatomy of the Testis. *This sagittal view shows the seminiferous tubules, the site of sperm production. Formed sperm are transferred to the epididymis, where they mature. They leave the epididymis during an ejaculation via the ductus deferens.*

1. Sertoli Cells: Surrounding all stages of the developing sperm cells are elongate, branching Sertoli cells. Sertoli cells secrete signaling molecules that promote sperm production and can control whether germ cells live or die.

2. Germ Cells: The least mature cells line the basement membrane inside the tubule. These divide to produce primary and secondary spermatocytes, then spermatids, which finally produce formed sperm. This process is called **spermatogenesis**. The process begins at puberty, after which time sperm are produced constantly throughout a man's life. Eventually, the sperm are released into the lumen and are moved along a series of ducts in the testis toward a structure called the epididymis for the next step of sperm maturation.

Structure of Formed Sperm: Sperm are smaller than most cells in the body; in fact, the volume of a sperm cell is 85,000 times less than that of the female gamete. Approximately 100 to 300 million sperm are produced each day, whereas women typically ovulate only one oocyte per month as is true for most cells in the body, the structure of sperm cells speaks to their function. Sperm have a distinctive head, mid-piece, and tail region (Figure 4). The head of the sperm contains the extremely compact haploid (half the genetic content of a diploid somatic cell) nucleus with very little cytoplasm. These qualities contribute to the overall small size of the sperm

(the head is only 5 µm long). A structure called the acrosome covers most of the head of the sperm cell as a "cap" that is filled with lysosomal enzymes important for preparing sperm to participate in fertilization. Tightly packed mitochondria fill the midpiece of the sperm. ATP produced by these mitochondria will power the flagellum, which extends from the neck and the mid-piece through the tail of the sperm, enabling it to move the entire sperm cell. The central strand of the flagellum, the axial filament, is formed from one centriole inside the maturing sperm cell during the final stages of spermatogenesis.

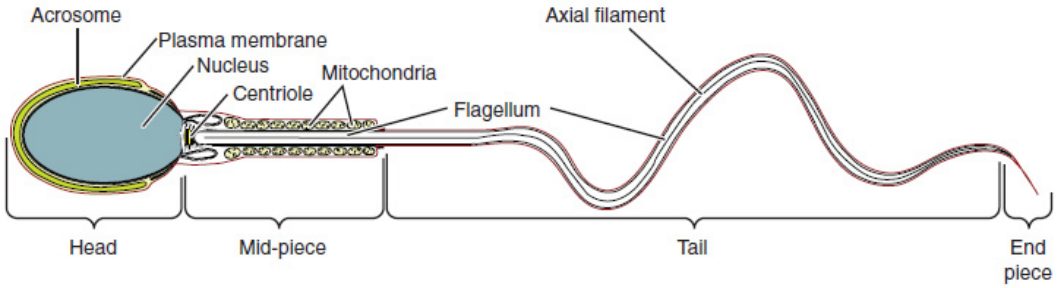


Figure 4. Structure of Sperm. *Sperm cells are divided into a head, containing DNA; a mid-piece, containing mitochondria; and a tail, providing motility. The acrosome is oval and somewhat flattened.*

Sperm Transport: To fertilize an egg, sperm must be moved from the seminiferous tubules in the testes, through the epididymis, and—later during ejaculation—along the length of the penis and out into the female reproductive tract.

1. Role of the Epididymis: From the lumen of the seminiferous tubules, the immotile sperm are surrounded by testicular fluid and moved to the **epididymis** (plural = epididymides), a coiled tube attached to the testis where newly formed sperm continue to mature (Figure 3). Though the epididymis does not take up much room in its tightly coiled state, it would be approximately 6 m (20 feet) long if straightened. It takes an average of 12 days for sperm to move through the coils of the epididymis, with the shortest recorded transit time in humans being one day. Sperm enter the head of the epididymis and are moved along predominantly by the contraction of smooth muscles lining the epididymal tubes. As they are moved along the length of the epididymis, the sperm further mature and acquire the ability to move under their own power. Once inside the female reproductive tract, they will use this ability to move independently toward the unfertilized egg. The more mature sperm are then stored in the tail of the epididymis (the final section) until ejaculation occurs.

2. Duct System: During ejaculation, sperm exit the tail of the epididymis and are pushed by smooth muscle contraction to the ductus deferens (also called the vas deferens). The ductus deferens is a thick, muscular tube that is bundled together inside the scrotum with connective tissue, blood vessels, and nerves into a structure called the **spermatic cord** (Figure 1 and Figure 2). Because the ductus deferens is physically accessible within the scrotum, surgical sterilization to interrupt sperm delivery can be performed by cutting and sealing a small section of the ductus (vas) deferens. This procedure is called a vasectomy, and it is an effective form of male birth control. Although it may be possible to reverse a vasectomy, clinicians consider the procedure permanent, and advise men to undergo it only if they are certain they no longer wish to father children.

From each epididymis, each ductus deferens extends superiorly into the abdominal cavity through the **inguinal canal** in the abdominal wall. From here, the ductus deferens continues posteriorly to the pelvic cavity, ending posterior to the bladder where it dilates in a region called the ampulla (meaning "flask").

Sperm make up only 5 percent of the final volume of **semen**, the viscous, whitish-gray fluid that the male ejaculates. The bulk of semen is produced by three critical accessory glands of the male reproductive system: the seminal vesicles, the prostate, and the bulbourethral glands.

3. Seminal Vesicles: As sperm pass through the ampulla of the ductus deferens at ejaculation, they mix with fluid from the associated **seminal vesicle** (Figure 1). The paired seminal vesicles are glands that contribute approximately 60 percent of the semen volume. Seminal vesicle fluid contains large amounts of fructose, which is used by the sperm mitochondria to generate ATP to allow movement through the female reproductive tract.

The fluid, now containing both sperm and seminal vesicle secretions, next moves into the associated **ejaculatory duct**, a short structure formed from the ampulla of the ductus deferens and the duct of the seminal vesicle. The paired ejaculatory ducts transport the seminal fluid into the next structure, the prostate gland.

4. Prostate Gland: The centrally located **prostate gland** sits anterior to the rectum at the base of the bladder surrounding the prostatic urethra, the portion of the urethra that runs within the prostate (Figure 1). About the size of a walnut, the prostate is formed of both muscular and glandular tissues. It excretes a fluid enriched with enzymes and citric acid to the passing seminal fluid—now called semen—that is critical to first coagulate and then decoagulate the semen following ejaculation. The temporary thickening of semen helps retain it within the female reproductive tract, providing time for sperm to utilize the fructose provided by seminal vesicle secretions. When the semen regains its fluid state, sperm can then pass farther into the female reproductive tract.

5. Bulbourethral Glands: The final addition to semen is made by two bulbourethral glands (or Cowper's glands) that release a thick, salty fluid that lubricates the end of the urethra and the vagina and helps to clean urine residues from the penile urethra. The fluid from these accessory glands is released after the male becomes sexually aroused, and shortly before the release of the semen. It is therefore sometimes called pre-ejaculate. It is important to note that, in addition to the lubricating proteins, it is possible for bulbourethral fluid to pick up sperm already present in the urethra, and therefore it may be able to cause pregnancy.

The Penis: The penis is the male organ of copulation (sexual intercourse). It is flaccid for non-sexual actions, such as urination, and turgid and rod-like with sexual arousal. When erect, the stiffness of the organ allows it to penetrate into the vagina and deposit semen into the female reproductive tract.

The shaft of the penis surrounds the urethra (Figure 1). The end of the penis, called the **glans penis**, has a high concentration of nerve endings, resulting in very sensitive skin that influences the likelihood of ejaculation (see Figure 1). The skin from the shaft extends down over the glans and forms a collar called the **prepuce** (or foreskin). The foreskin also contains a dense concentration of nerve endings, and both lubricate and protect the sensitive skin of the glans penis. A surgical procedure called circumcision, often performed for religious or social reasons, removes the prepuce, typically within days of birth.

Testosterone: Testosterone, an androgen, is a steroid hormone produced by **Leydig cells**. The alternate term for Leydig cells, interstitial cells, reflects their location between the seminiferous tubules in the testes. In male embryos, testosterone is secreted by Leydig cells by the seventh week of development, with peak concentrations reached in the second trimester. This early release of testosterone results in the anatomical differentiation of the male sexual organs. In childhood, testosterone concentrations are low. They increase during puberty, activating characteristic physical changes and initiating spermatogenesis.

Functions of Testosterone: The continued presence of testosterone is necessary to keep the male reproductive system working properly, and Leydig cells produce approximately 6 to 7 mg of testosterone per day. Testicular steroidogenesis (the manufacturing of androgens, including testosterone) results in testosterone concentrations that are 100 times higher in the testes than in the circulation. Maintaining these normal concentrations of testosterone promotes spermatogenesis, whereas low levels of testosterone can lead to infertility. In addition to intratesticular secretion, testosterone is also released into the systemic circulation and plays an important role in muscle development, bone growth, the development of secondary sex characteristics, and maintaining libido (sex drive) in both males and females. In females, the ovaries secrete small amounts of testosterone, although most is converted to estradiol. A small amount of testosterone is also secreted by the adrenal glands in both sexes.

Part 2: Anatomy and Physiology of the Female Reproductive System

The female reproductive system functions to produce gametes and reproductive hormones, just like the male reproductive system; however, it also has the additional task of supporting the developing fetus and delivering it to the outside world. Unlike its male counterpart, the female reproductive system is located primarily inside the pelvic cavity (Figure 5). Recall that the ovaries are the female gonads. The gamete they produce is called an

oocyte. We'll discuss the production of oocytes in detail shortly. First, let's look at some of the structures of the female reproductive system.

Watch this Crash Course video for an overview of the female reproductive system! Direct link: https://youtu.be/RFDatCchpus

External Female Genitals: The external female reproductive structures are referred to collectively as the vulva. The mons pubis is a pad of fat found in both males and females that is located at the anterior, over the pubic bone. After puberty, it becomes covered in pubic hair. In females, the **labia majora** (labia = "lips"; majora = "larger") are folds of hair-covered skin that begin just posterior to the mons pubis. The thinner and more pigmented **labia minora** (labia = "lips"; minora = "smaller") extend medial to the labia majora. Although they naturally vary in shape and size from woman to woman, the labia minora serve to protect the female urethra and the entrance to the female reproductive tract.

The superior, anterior portions of the labia minora come together to encircle the **clitoris** (or glans clitoris), an organ that originates from the same cells as the glans penis and has abundant nerves that make it important in sexual sensation and orgasm. The hymen is a thin membrane that sometimes partially covers the entrance to the vagina. The vaginal opening is located between the opening of the urethra and the anus. It is flanked by outlets to the **Bartholin's glands** (or greater vestibular glands).

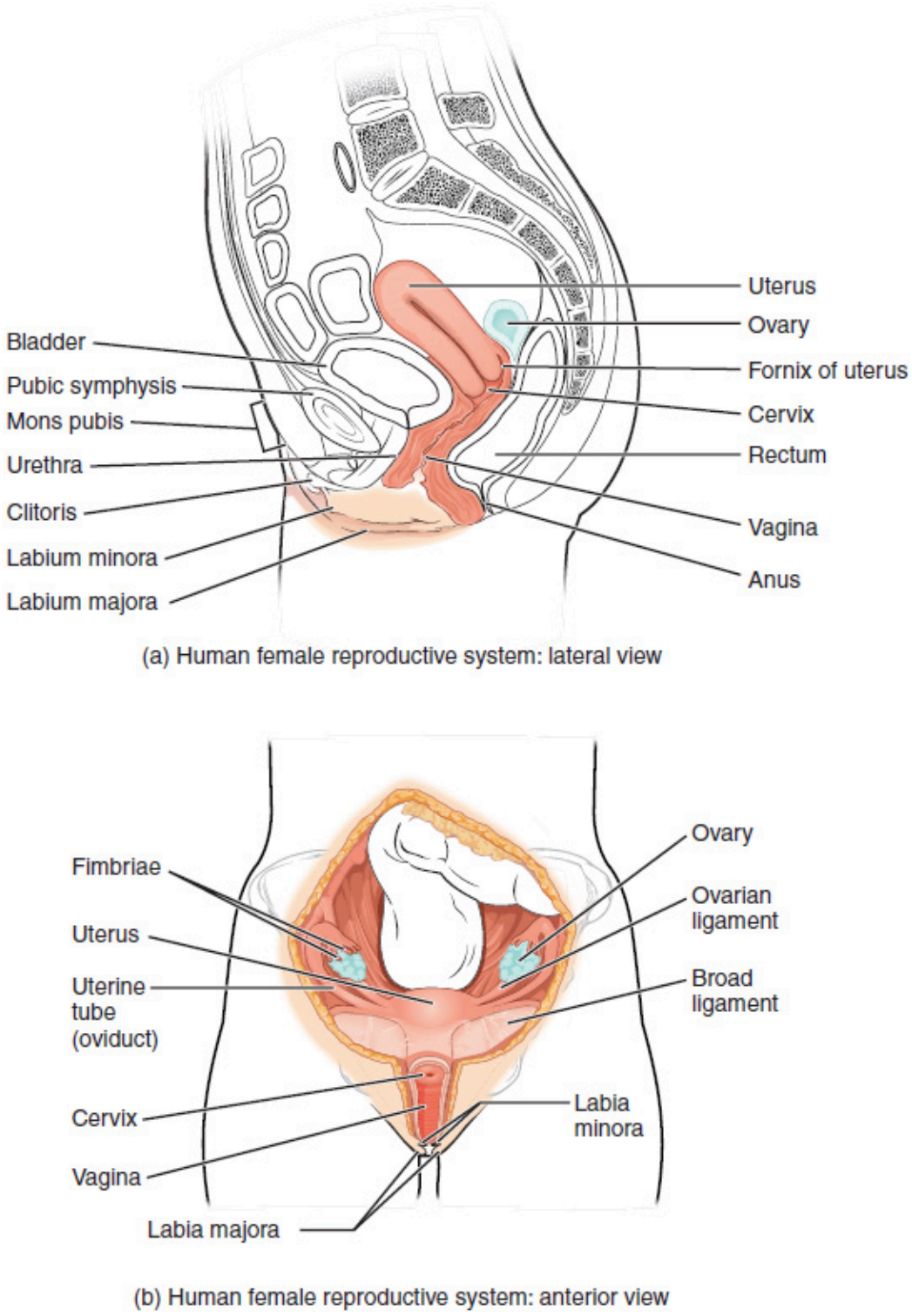


Figure 5. Female Reproductive System. *The major organs of the female reproductive system are located inside the pelvic cavity.*

Vagina: The **vagina** is a muscular canal (approximately 10 cm long) that serves as the entrance to the reproductive tract (Figure 5). It also serves as the exit from the uterus during menses and childbirth. The outer walls of the anterior and posterior vagina are formed into longitudinal columns, or ridges, and the superior portion of the vagina—called the fornix—meets the protruding uterine cervix. The walls of the vagina are

lined with an outer fibrous adventitia, a middle layer of smooth muscle, and an inner mucous membrane with transverse folds called rugae. Together, the middle and inner layers allow the expansion of the vagina to accommodate intercourse and childbirth. The Bartholin's glands and the lesser vestibular glands (located near the clitoris) secrete mucus, which keeps the vestibular area moist.

The vagina is home to a normal population of microorganisms that help to protect against infection by pathogenic bacteria, yeast, or other organisms that can enter the vagina. In a healthy woman, the most predominant type of vaginal bacteria is from the genus *Lactobacillus*. This family of beneficial bacterial flora secretes lactic acid, and thus protects the vagina by maintaining an acidic pH (below 4.5). Potential pathogens are less likely to survive in these acidic conditions. Lactic acid, in combination with other vaginal secretions, makes the vagina a self-cleansing organ.

Ovaries: The **ovaries** are the female gonads (Figure 5). Paired ovals, they are each about 2 to 3 cm in length, about the size of an almond. The ovaries are located within the pelvic cavity, and are supported by ligaments.

The bulk of the adult ovary is composed of a tissue framework. Oocytes develop within the outer layer of this stroma, each surrounded by supporting cells. This grouping of an oocyte and its supporting cells is called a **follicle**. The growth and development of ovarian follicles will be described shortly.

The Ovarian Cycle: The **ovarian cycle** is a set of predictable changes in a female's oocytes and ovarian follicles. During a woman's reproductive years, it is a roughly 28-day cycle that can be correlated with, but is not the same as, the menstrual cycle (discussed shortly). The cycle includes two interrelated processes: oogenesis (the production of female gametes) and folliculogenesis (the growth and development of ovarian follicles).

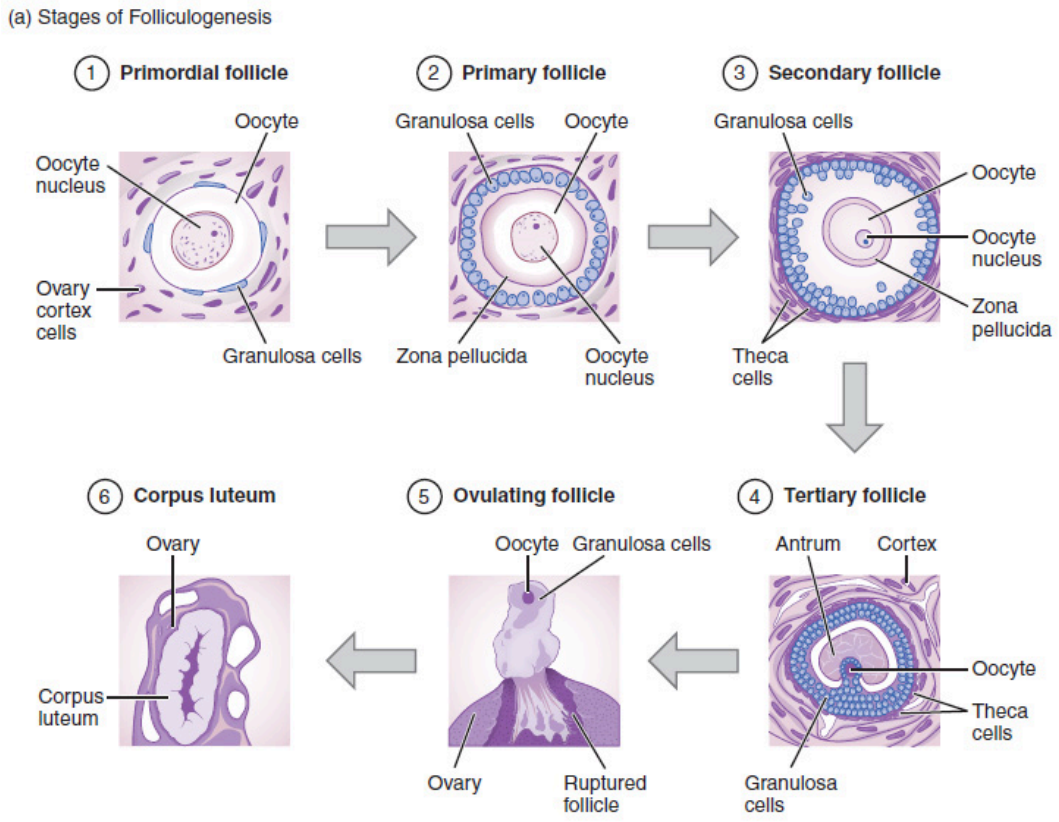
1. Oogenesis: Gametogenesis in females is called **oogenesis**. The process begins with the ovarian stem cells, or **oogonia**. Oogonia form primary oocytes in the fetal ovary prior to birth. These primary oocytes are then arrested in this stage, only to resume it years later, beginning at puberty and continuing until the woman is near menopause (the cessation of a woman's reproductive functions). The number of primary oocytes present in the ovaries declines from one to two million in an infant, to approximately 400,000 at puberty, to zero by the end of menopause.

The initiation of **ovulation**—the release of an oocyte from the ovary—marks the transition from puberty into reproductive maturity for women. From then on, throughout a woman's reproductive years, ovulation occurs approximately once every 28 days, triggered by a surge of LH just prior to ovulation.

The larger amount of cytoplasm contained in the female gamete is used to supply the developing zygote with nutrients during the period between fertilization and implantation into the uterus. Interestingly, sperm contribute only DNA at fertilization, not cytoplasm. Therefore, the cytoplasm and all the cytoplasmic organelles in the developing embryo are of maternal origin. This includes mitochondria, which contain their own DNA. Scientific research in the 1980s determined that mitochondrial DNA is maternally inherited, meaning that you can trace your mitochondrial DNA directly to your mother, her mother, and so on back through your female ancestors.

2. Folliculogenesis: Again, ovarian follicles are oocytes and their supporting cells. They grow and develop in a process called **folliculogenesis**, which typically leads to ovulation of one follicle approximately every 28 days, along with death to multiple other follicles. The death of ovarian follicles is called atresia and can occur at any point during follicular development. Recall that, a female infant at birth will have one to two million oocytes within her ovarian follicles, and that this number declines throughout life until menopause, when no follicles remain. As you will see next, follicles progress from primordial, to primary, to secondary and tertiary stages prior to ovulation—with the oocyte inside the follicle remaining as a primary oocyte until right before ovulation.

Folliculogenesis begins with follicles in a resting state. These small **primordial follicles** are present in newborn females and are the prevailing follicle type in the adult ovary (Figure 6). Primordial follicles have only a single flat layer of support cells, called **granulosa cells**, that surround the oocyte, and they can stay in this resting state for years—some until right before menopause.



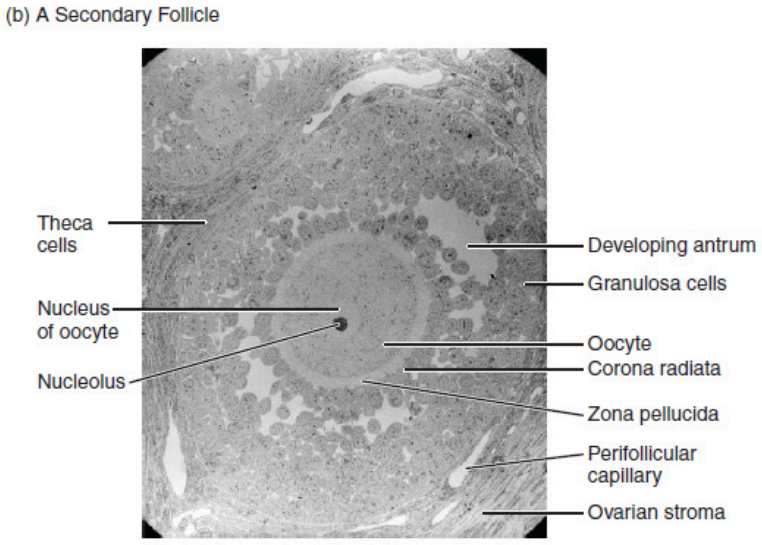


Figure 6. Folliculogenesis. *(a) The maturation of a follicle is shown in a clockwise direction proceeding from the primordial follicles. FSH stimulates the growth of a tertiary follicle, and LH stimulates the production of estrogen by granulosa and theca cells. Once the follicle is mature, it ruptures and releases the oocyte. Cells remaining in the follicle then develop into the corpus luteum. (b) In this electron micrograph of a secondary follicle, the oocyte, theca cells (thecae folliculi), and developing antrum are clearly visible. EM × 1100. (Micrograph provided by the Regents of University of Michigan Medical School © 2012)*

After puberty, a few primordial follicles will respond to a recruitment signal each day and will join a pool of immature growing follicles called primary follicles. Primary follicles start with a single layer of granulosa cells, but the granulosa cells then become active and transition from a flat or squamous shape to a rounded cuboidal shape as they increase in size and proliferate. As the granulosa cells divide, the follicles—now called **secondary follicles** (Figure 6)—increase in diameter, adding a new outer layer of connective tissue, blood vessels, and **theca cells**—cells that work with the granulosa cells to produce estrogens.

Within the growing secondary follicle, the primary oocyte now secretes a thin acellular membrane called the zona pellucida that will play a critical role in fertilization. A thick fluid, called follicular fluid, that has formed

between the granulosa cells also begins to collect into one large pool, or **antrum**. Follicles in which the antrum has become large and fully formed are considered **tertiary follicles** (or antral follicles). Several follicles reach the tertiary stage at the same time, and most of these will undergo atresia. The one that does not die will continue to grow and develop until ovulation, when it will expel its secondary oocyte surrounded by several layers of granulosa cells from the ovary. Keep in mind that most follicles do not make it to this point. In fact, roughly 99 percent of the follicles in the ovary will undergo atresia, which can occur at any stage of folliculogenesis.

Hormonal Control of the Ovarian Cycle: The process of development that we have just described, from primordial follicle to early tertiary follicle, takes approximately two months in humans. The final stages of development of a small cohort of tertiary follicles, ending with ovulation of a secondary oocyte, occur over a course of approximately 28 days. These changes are regulated by many of the same hormones that regulate the male reproductive system, including GnRH, LH, and FSH.

As in men, the hypothalamus produces GnRH, a hormone that signals the anterior pituitary gland to produce the gonadotropins FSH and LH (Figure 7). These gonadotropins leave the pituitary and travel through the bloodstream to the ovaries, where they bind to receptors on the granulosa and theca cells of the follicles. FSH stimulates the follicles to grow (hence its name of follicle-stimulating hormone), and the five or six tertiary follicles expand in diameter. The release of LH also stimulates the granulosa and theca cells of the follicles to produce the sex steroid hormone estradiol, a type of estrogen. This phase of the ovarian cycle, when the tertiary follicles are growing and secreting estrogen, is known as the follicular phase.

The more granulosa and theca cells a follicle has (that is, the larger and more developed it is), the more estrogen it will produce in response to LH stimulation. As a result of these large follicles producing large amounts of estrogen, systemic plasma estrogen concentrations increase. Following a classic negative feedback loop, the high concentrations of estrogen will inhibit the production of GnRH, LH, and FSH by the hypothalamus and pituitary. Because the large tertiary follicles require FSH to grow and survive at this point, this decline in FSH caused by negative feedback leads most of them to die (atresia). Typically, only one follicle, now called the dominant follicle, will survive this reduction in FSH, and this follicle will be the one that releases an oocyte.

When only the one dominant follicle remains in the ovary, it again begins to secrete estrogen. It produces more estrogen than all the developing follicles did together before the negative feedback occurred. It produces so much estrogen that the normal negative feedback does not occur. Instead, these extremely high concentrations of systemic plasma estrogen trigger a regulatory switch in the anterior pituitary that responds by secreting large amounts of LH and FSH into the bloodstream (Figure 7). The positive feedback loop by which more estrogen triggers release of more LH and FSH only occurs at this point in the cycle.

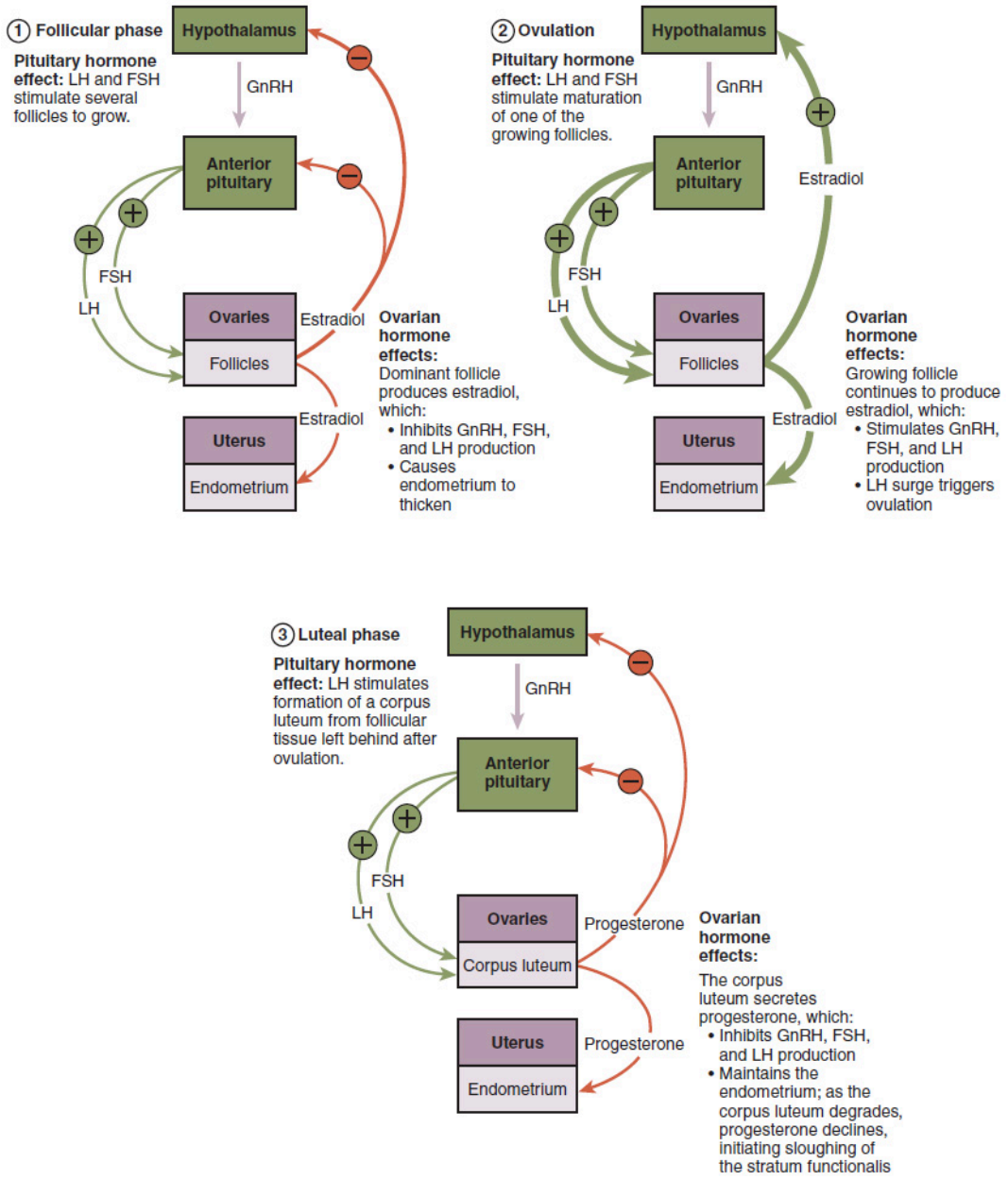


Figure 7. Hormonal Regulation of Ovulation. *The hypothalamus and pituitary gland regulate the ovarian cycle and ovulation. GnRH activates the anterior pituitary to produce LH and FSH, which stimulate the production of estrogen and progesterone by the ovaries.*

It is this large burst of LH (called the LH surge) that leads to ovulation of the dominant follicle. The LH surge induces many changes in the dominant follicle, including triggering proteases (enzymes that cleave proteins) to break down structural proteins in the ovary wall on the surface of the bulging dominant follicle. This degradation of the wall, combined with pressure from the large, fluid-filled antrum, results in the expulsion of the oocyte surrounded by granulosa cells into the peritoneal cavity. This release is **ovulation**.

In the next section, you will follow the ovulated oocyte as it travels toward the uterus, but there is one more important event that occurs in the ovarian cycle. The surge of LH also stimulates a change in the granulosa and theca cells that remain in the follicle after the oocyte has been ovulated. This change is called luteinization (recall that the full name of LH is luteinizing hormone), and it transforms the collapsed follicle into a new endocrine structure called the **corpus luteum**, a term meaning "yellowish body" (Figure 8). Instead of estrogen, the luteinized granulosa and theca cells of the corpus luteum begin to produce large amounts of the sex

steroid hormone progesterone, a hormone that is critical for the establishment and maintenance of pregnancy. Progesterone triggers negative feedback at the hypothalamus and pituitary, which keeps GnRH, LH, and FSH secretions low, so no new dominant follicles develop at this time.

The post-ovulatory phase of progesterone secretion is known as the luteal phase of the ovarian cycle. If pregnancy does not occur within 10 to 12 days, the corpus luteum will stop secreting progesterone and degrade into the **corpus albicans**, a nonfunctional "whitish body" that will disintegrate in the ovary over a period of several months. During this time of reduced progesterone secretion, FSH and LH are once again stimulated, and the follicular phase begins again with a new cohort of early tertiary follicles beginning to grow and secrete estrogen.

The Uterine Tubes: The **uterine tubes** (also called fallopian tubes or oviducts) serve as the conduit of the oocyte from the ovary to the uterus (Figure 8). Each of the two uterine tubes is close to, but not directly connected to, the ovary and divided into sections. The **isthmus** is the narrow medial end of each uterine tube that is connected to the uterus. The wide distal **infundibulum** flares out with slender, finger-like projections called **fimbriae**. The middle region of the tube, called the **ampulla**, is where fertilization often occurs. The uterine tubes also contain ciliated cells that beat in the direction of the uterus, producing a current that will be critical to move the oocyte.

Following ovulation, the secondary oocyte surrounded by a few granulosa cells is released into the peritoneal cavity. The nearby uterine tube, either left or right, receives the oocyte. Unlike sperm, oocytes lack flagella, and therefore cannot move on their own. So how do they travel into the uterine tube and toward the uterus? High concentrations of estrogen that occur around the time of ovulation induce contractions of the smooth muscle along the length of the uterine tube and the result is a coordinated movement that sweeps the surface of the ovary and the pelvic cavity. Current flowing toward the uterus is generated by coordinated beating of the cilia that line the lumen of the length of the uterine tube. These cilia beat more strongly in response to the high estrogen concentrations that occur around the time of ovulation. As a result of these mechanisms, the oocyte–granulosa cell complex is pulled into the interior of the tube. Once inside, the muscular contractions and beating cilia move the oocyte slowly toward the uterus. When fertilization does occur, sperm typically meet the egg while it is still moving through the ampulla.

If the oocyte is successfully fertilized, the resulting zygote will begin to divide into two cells, then four, and so on, as it makes its way through the uterine tube and into the uterus. There, it will implant and continue to grow. If the egg is not fertilized, it will simply degrade—either in the uterine tube or in the uterus, where it may be shed with the next menstrual period.

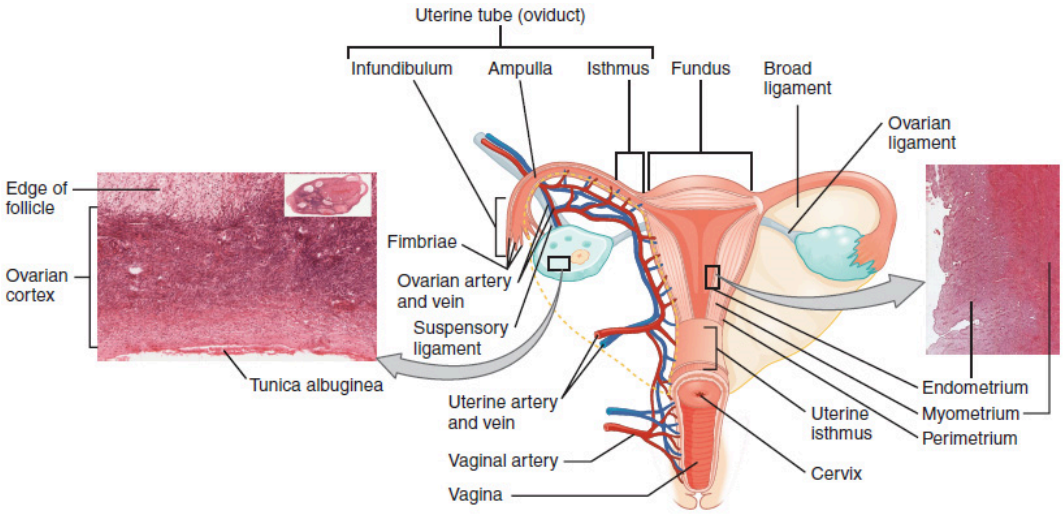


Figure 8. Ovaries, Uterine Tubes, and Uterus. *This anterior view shows the relationship of the ovaries, uterine tubes (oviducts), and uterus. Sperm enter through the vagina, and fertilization of an ovulated oocyte usually occurs in the distal uterine tube. From left to right, LM × 400, LM × 20. (Micrographs provided by the Regents of University of Michigan Medical School © 2012)*

The Uterus and Cervix: The uterus is the muscular organ that nourishes and supports the growing embryo (Figure 8). Its average size is approximately 5 cm wide by 7 cm long (approximately 2 in by 3 in) when a female is not pregnant. It has three sections. The portion of the uterus superior to the opening of the uterine tubes is called the **fundus**. The middle section of the uterus is called the **body of uterus** (or corpus). The **cervix** is the narrow inferior portion of the uterus that projects into the vagina. The cervix produces mucus secretions that become thin and stringy under the influence of high systemic plasma estrogen concentrations, and these secretions can facilitate sperm movement through the reproductive tract. Several ligaments maintain the position of the uterus within the abdominopelvic cavity.

The wall of the uterus is made up of three layers. The most superficial layer is the serous membrane, or **perimetrium**, which consists of epithelial tissue that covers the exterior portion of the uterus. The middle layer, or **myometrium**, is a thick layer of smooth muscle responsible for uterine contractions. Most of the uterus is myometrial tissue, and the muscle fibers run horizontally, vertically, and diagonally, allowing the powerful contractions that occur during labor and the less powerful contractions (or cramps) that help to expel menstrual blood during a woman's period.

The innermost layer of the uterus is called the endometrium. The **endometrium** contains a connective tissue lining, the lamina propria, which is covered by epithelial tissue that lines the lumen. Structurally, the endometrium consists of two layers: the stratum basalis and the stratum functionalis (the basal and functional layers). The stratum basalis layer is part of the lamina propria and is adjacent to the myometrium; this layer does not shed during menses. In contrast, the thicker stratum functionalis layer contains the glandular portion of the lamina propria and the endothelial tissue that lines the uterine lumen. It is the stratum functionalis that grows and thickens in response to increased levels of estrogen and progesterone. In the luteal phase of the menstrual cycle, special branches off of the uterine artery called spiral arteries supply the thickened stratum functionalis. This inner functional layer provides the proper site of implantation for the fertilized egg, and—should fertilization not occur—it is only the stratum functionalis layer of the endometrium that sheds during menstruation.

Recall that during the follicular phase of the ovarian cycle, the tertiary follicles are growing and secreting estrogen. At the same time, the stratum functionalis of the endometrium is thickening to prepare for a potential implantation. The post-ovulatory increase in progesterone, which characterizes the luteal phase, is key for maintaining a thick stratum functionalis. As long as a functional corpus luteum is present in the ovary, the endometrial lining is prepared for implantation. Indeed, if an embryo implants, signals are sent to the corpus luteum to continue secreting progesterone to maintain the endometrium, and thus maintain the pregnancy. If an embryo does not implant, no signal is sent to the corpus luteum and it degrades, ceasing progesterone production and ending the luteal phase. Without progesterone, the endometrium becomes thinner and, under the influence of prostaglandins, the spiral arteries of the endometrium constrict and rupture, preventing oxygenated blood from reaching the endometrial tissue. As a result, endometrial tissue dies and blood, pieces of the endometrial tissue, and white blood cells are shed through the vagina during menstruation, or the **menses**.

The Menstrual Cycle: Now that we have discussed the maturation of the cohort of tertiary follicles in the ovary, the build-up and then shedding of the endometrial lining in the uterus, and the function of the uterine tubes and vagina, we can put everything together to talk about the three phases of the **menstrual cycle**—the series of changes in which the uterine lining is shed, rebuilds, and prepares for implantation.

The timing of the menstrual cycle starts with the first day of menses, referred to as day one of a woman's period. Cycle length is determined by counting the days between the onset of bleeding in two subsequent cycles. Because the average length of a woman's menstrual cycle is 28 days, this is the time period used to identify the timing of events in the cycle. However, the length of the menstrual cycle varies among women, and even in the same woman from one cycle to the next, typically from 21 to 32 days.

Just as the hormones produced by the granulosa and theca cells of the ovary "drive" the follicular and luteal phases of the ovarian cycle, they also control the three distinct phases of the menstrual cycle. These are the menses phase, the proliferative phase, and the secretory phase.

1. Menses Phase: The **menses phase** of the menstrual cycle is the phase during which the lining is shed; that is, the days that the woman menstruates. Although it averages approximately five days, the menses phase can last from 2 to 7 days, or longer. The menses phase occurs during the early days of the follicular phase of the ovarian cycle, when progesterone, FSH, and LH levels are low (Figure 9). Recall that progesterone concentrations decline as a result of the degradation of the corpus luteum, marking the end of the luteal phase. This decline in progesterone triggers the shedding of the stratum functionalis of the endometrium.

2. Proliferative Phase: Once menstrual flow ceases, the endometrium begins to proliferate again, marking the beginning of the **proliferative phase** of the menstrual cycle (Figure 9). It occurs when the granulosa and theca cells of the tertiary follicles begin to produce increased amounts of estrogen. These rising estrogen concentrations stimulate the endometrial lining to rebuild.

Recall that the high estrogen concentrations will eventually lead to a decrease in FSH as a result of negative feedback, resulting in atresia of all but one of the developing tertiary follicles. The switch to positive feedback—which occurs with the elevated estrogen production from the dominant follicle—then stimulates the LH surge that will trigger ovulation. In a typical 28-day menstrual cycle, ovulation occurs on day 14. Ovulation marks the end of the proliferative phase as well as the end of the follicular phase.

3. Secretory Phase: In addition to prompting the LH surge, high estrogen levels increase the uterine tube contractions that facilitate the pick-up and transfer of the ovulated oocyte. High estrogen levels also slightly decrease the acidity of the vagina, making it more hospitable to sperm. In the ovary, the luteinization of the granulosa cells of the collapsed follicle forms the progesterone-producing corpus luteum, marking the beginning of the luteal phase of the ovarian cycle. In the uterus, progesterone from the corpus luteum begins the **secretory phase** of the menstrual cycle, in which the endometrial lining prepares for implantation (Figure 9). Over the next 10 to 12 days, the endometrial glands secrete a fluid rich in glycogen. If fertilization has occurred, this fluid will nourish the ball of cells now developing from the zygote. At the same time, the spiral arteries develop to provide blood to the thickened stratum functionalis.

If no pregnancy occurs within approximately 10 to 12 days, the corpus luteum will degrade into the corpus albicans. Levels of both estrogen and progesterone will fall, and the endometrium will grow thinner. Prostaglandins will be secreted that cause constriction of the spiral arteries, reducing oxygen supply. The endometrial tissue will die, resulting in menses—or the first day of the next cycle.

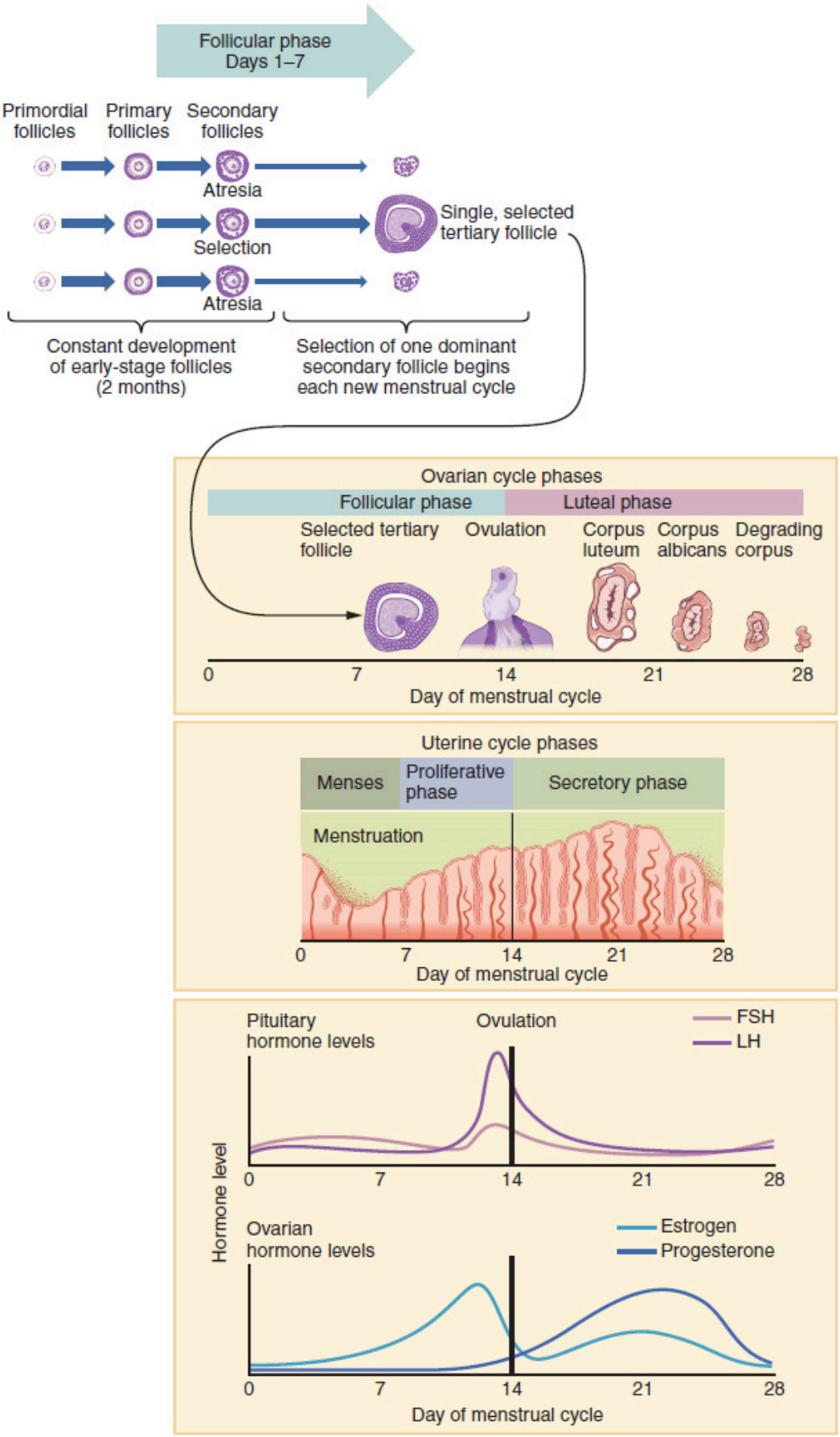


Figure 9. Hormone Levels in Ovarian and Menstrual Cycles. *The correlation of the hormone levels and their effects on the female reproductive system is shown in this timeline of the ovarian and menstrual cycles. The menstrual cycle begins at day one with the start of menses. Ovulation occurs around day 14 of a 28-day cycle, triggered by the LH surge.*

The Breasts: Although the breasts are located far from the other female reproductive organs, they are considered accessory organs of the female reproductive system. The function of the breasts is to supply milk to an infant in a process called lactation. The external features of the breast include a nipple surrounded by a pigmented **areola** (Figure 10), whose coloration may deepen during pregnancy. The areola is typically circular and can vary in size from 25 to 100 mm in diameter. The areolar region is characterized by small, raised areolar glands that secrete lubricating fluid during lactation to protect the nipple from chafing. When a baby nurses, or draws milk from the breast, the entire areolar region is taken into the mouth.

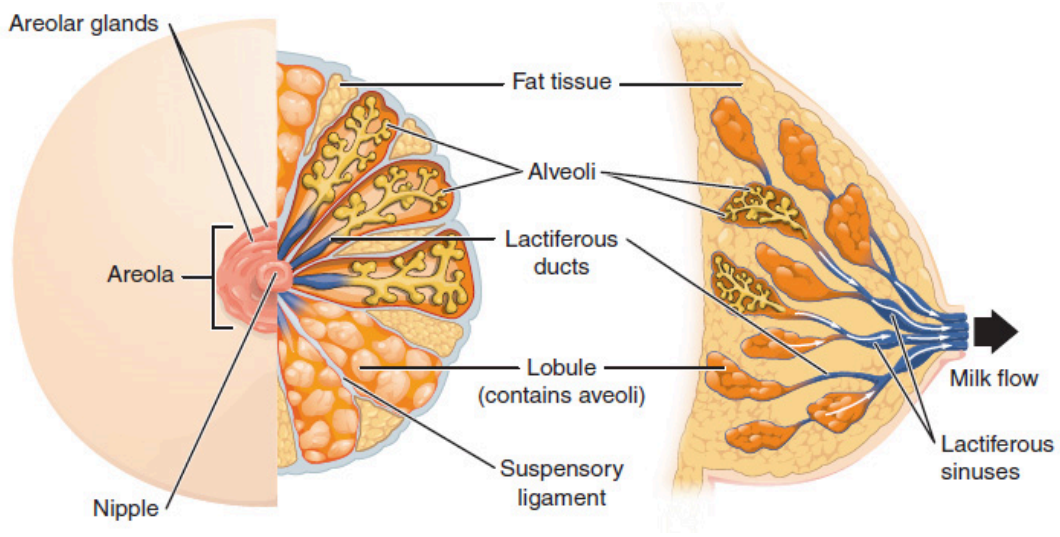


Figure 10. Anatomy of the Breast. *During lactation, milk moves from the alveoli through the lactiferous ducts to the nipple.*

Breast milk is produced by the **mammary glands**, which are modified sweat glands. The milk itself exits the breast through the nipple via 15 to 20 **lactiferous ducts** that open on the surface of the nipple. These lactiferous ducts each extend to a **lactiferous sinus** that connects to a glandular lobe within the breast itself that contains groups of milk-secreting cells in clusters called **alveoli** (Figure 10). The clusters can change in size depending on the amount of milk in the alveolar lumen. Once milk is made in the alveoli, stimulated myoepithelial cells that surround the alveoli contract to push the milk to the lactiferous sinuses. From here, the baby can draw milk through the lactiferous ducts by suckling. The lobes themselves are surrounded by fat tissue, which determines the size of the breast; breast size differs between individuals and does not affect the amount of milk produced. Supporting the breasts are multiple bands of connective tissue called suspensory ligaments that connect the breast tissue to the dermis of the overlying skin.

Part 3: Fertilization

Fertilization occurs when a sperm and an oocyte (egg) combine and their nuclei fuse. Because each of these reproductive cells is a haploid cell containing half of the genetic material needed to form a human being, their combination forms a diploid cell. This new single cell, called a **zygote**, contains all the genetic material needed to form a human—half from the mother and half from the father.

Transit of Sperm: Fertilization is a numbers game. During ejaculation, hundreds of millions of sperm (spermatozoa) are released into the vagina. Almost immediately, millions of these sperm are overcome by the acidity of the vagina (approximately pH 3.8), and millions more may be blocked from entering the uterus by thick cervical mucus. Of those that do enter, thousands are destroyed by phagocytic uterine leukocytes. Thus, the race into the uterine tubes, which is the most typical site for sperm to encounter the oocyte, is reduced to a few thousand contenders. Their journey—thought to be facilitated by uterine contractions—usually takes from 30 minutes to 2 hours. If the sperm do not encounter an oocyte immediately, they can survive in the uterine tubes for another 3–5 days. Thus, fertilization can still occur if intercourse takes place a few days

before ovulation. In comparison, an oocyte can survive independently for only approximately 24 hours following ovulation. Intercourse more than a day after ovulation will therefore usually not result in fertilization.

During the journey, fluids in the female reproductive tract prepare the sperm for fertilization through a process called **capacitation**, or priming. The fluids improve the motility of the spermatozoa. They also deplete cholesterol molecules embedded in the membrane of the head of the sperm, thinning the membrane in such a way that will help facilitate the release of the lysosomal (digestive) enzymes needed for the sperm to penetrate the oocyte's exterior once contact is made. Sperm must undergo the process of capacitation in order to have the "capacity" to fertilize an oocyte. If they reach the oocyte before capacitation is complete, they will be unable to penetrate the oocyte's thick outer layer of cells.

Contact Between Sperm and Oocyte: Upon ovulation, the oocyte released by the ovary is swept into—and along—the uterine tube. Fertilization must occur in the distal uterine tube because an unfertilized oocyte cannot survive the 72-hour journey to the uterus. As you will recall from your study of the oogenesis, this oocyte (specifically a secondary oocyte) is surrounded by two protective layers. **The corona radiata** is an outer layer of follicular (granulosa) cells that form around a developing oocyte in the ovary and remain with it upon ovulation. The underlying **zona pellucida** (pellucid = "transparent") is a transparent, but thick, glycoprotein membrane that surrounds the cell's plasma membrane.

As it is swept along the distal uterine tube, the oocyte encounters the surviving capacitated sperm, which stream toward it in response to chemical attractants released by the cells of the corona radiata. To reach the oocyte itself, the sperm must penetrate the two protective layers. The sperm first burrow through the cells of the corona radiata. Then, upon contact with the zona pellucida, the sperm bind to receptors in the zona pellucida. This initiates a process called the **acrosomal reaction** in which the enzyme-filled "cap" of the sperm, called the acrosome, releases its stored digestive enzymes. These enzymes clear a path through the zona pellucida that allows sperm to reach the oocyte. Finally, a single sperm makes contact with sperm-binding receptors on the oocyte's plasma membrane (Figure 11). The plasma membrane of that sperm then fuses with the oocyte's plasma membrane, and the head and mid-piece of the "winning" sperm enter the oocyte interior.

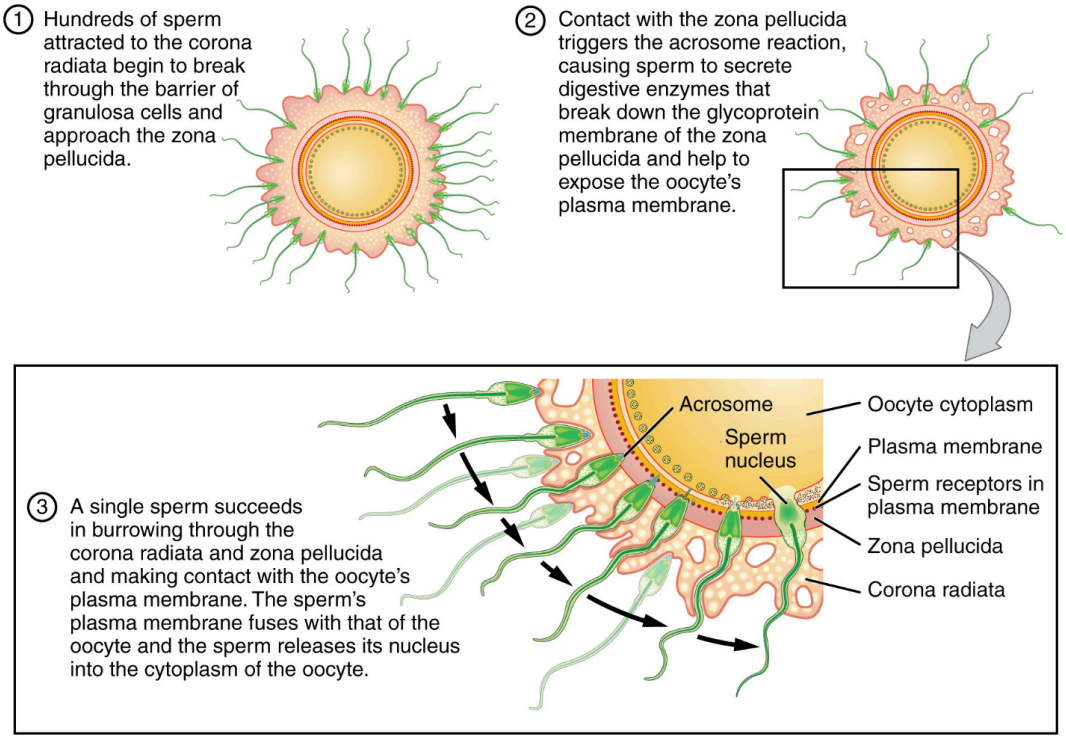


Figure 11. Sperm and the Process of Fertilization. *Before fertilization, hundreds of capacitated sperm must break through the surrounding corona radiata and zona pellucida so that one can contact and fuse with the oocyte plasma membrane.*

How do sperm penetrate the corona radiata? Some sperm undergo a spontaneous acrosomal reaction, which is an acrosomal reaction not triggered by contact with the zona pellucida. The digestive enzymes released by this reaction digest the extracellular matrix of the corona radiata. As you can see, the first sperm to reach the oocyte is never the one to fertilize it. Rather, hundreds of sperm cells must undergo the acrosomal reaction, each helping to degrade the corona radiata and zona pellucida until a path is created to allow one sperm to contact and fuse with the plasma membrane of the oocyte. If you consider the loss of millions of sperm between entry into the vagina and degradation of the zona pellucida, you can understand why a low sperm count can cause male infertility.

When the first sperm fuses with the oocyte, the oocyte deploys two mechanisms to prevent **polyspermy**, which is penetration by more than one sperm. This is critical because if more than one sperm were to fertilize the oocyte, the resulting zygote would be a triploid organism with three sets of chromosomes. This is incompatible with life.

The first mechanism is the fast block, which involves a near instantaneous change in sodium ion permeability upon binding of the first sperm, depolarizing the oocyte plasma membrane and preventing the fusion of additional sperm cells. The fast block sets in almost immediately and lasts for about a minute, during which time an influx of calcium ions following sperm penetration triggers the second mechanism, the slow block. In this process, referred to as the **cortical reaction**, cortical granules sitting immediately below the oocyte plasma membrane fuse with the membrane and release zonal inhibiting proteins and mucopolysaccharides into the space between the plasma membrane and the zona pellucida. Zonal inhibiting proteins cause the release of any other attached sperm and destroy the oocyte's sperm receptors, thus preventing any more sperm from binding. The mucopolysaccharides then coat the nascent zygote in an impenetrable barrier that, together with hardened zona pellucida, is called a **fertilization membrane**.

The Zygote: Upon fertilization the two haploid nuclei derived from the sperm and oocyte decondense, expand, and replicate their DNA. The pronuclei then migrate toward each other, their nuclear envelopes disintegrate, and the male- and female-derived genetic material intermingles. This step completes the process of fertilization and results in a single-celled diploid zygote with all the genetic instructions it needs to develop into a human.

Most of the time, a woman releases a single egg during an ovulation cycle. However, in approximately 1 percent of ovulation cycles, two eggs are released and both are fertilized. Two zygotes form, implant, and develop, resulting in the birth of dizygotic (or fraternal) twins. Because dizygotic twins develop from two eggs fertilized by two sperm, they are no more identical than siblings born at different times.

Much less commonly, a zygote can divide into two separate offspring during early development. This results in the birth of monozygotic (or identical) twins. Although the zygote can split as early as the two-cell stage, splitting occurs most commonly during the early blastocyst stage, with roughly 70–100 cells present. These two scenarios are distinct from each other, in that the twin embryos that separated at the two-cell stage will have individual placentas, whereas twin embryos that form from separation at the blastocyst stage will share a placenta and a chorionic cavity.

Watch this CrashCourse video to learn more about fertilization! Direct link: https://youtu.be/SUdAEGXLO-8

Part 4: Embryonic Development

Throughout this chapter, we will express embryonic and fetal ages in terms of weeks from fertilization, commonly called conception. The period of time required for full development of a fetus in utero is referred to as **gestation** (gestare = "to carry" or "to bear"). It can be subdivided into distinct gestational periods. A developing human is referred to as an **embryo** during weeks 3–8, and a **fetus** from the ninth week of gestation until birth. In this section, we'll cover the embryonic stages of development, which are characterized by cell division, migration, and differentiation. By the end of the embryonic period, all the organ systems are structured in rudimentary form, although the organs themselves are either nonfunctional or only semi-functional.

Pre-implantation Embryonic Development: Following fertilization, the zygote and its associated membranes, together referred to as the **conceptus**, continue to be projected toward the uterus by peristalsis and beating cilia. During its journey to the uterus, the zygote undergoes five or six rapid mitotic cell divisions. Although each **cleavage** results in more cells, it does not increase the total volume of the conceptus (Figure 12). Each daughter cell produced by cleavage is called a **blastomere** (blastos = "germ," in the sense of a seed or sprout).

Approximately 3 days after fertilization, a 16-cell conceptus reaches the uterus. The cells that had been loosely grouped are now compacted and look more like a solid mass. The name given to this structure is the **morula** (morula = "little mulberry"). Once inside the uterus, the conceptus floats freely for several more days. It continues to divide, creating a ball of approximately 100 cells, and consuming nutritive endometrial secretions called uterine milk while the uterine lining thickens. The ball of now tightly bound cells starts to secrete fluid and organize themselves around a fluid-filled cavity, the **blastocoel**. At this developmental stage, the conceptus is referred to as a **blastocyst**. Within this structure, a group of cells forms into an **inner cell mass**, which is fated to become the embryo. The cells that form the outer shell are called **trophoblasts** (trophe = "to feed" or "to nourish"). These cells will develop into the chorionic sac and the fetal portion of the **placenta** (the organ of nutrient, waste, and gas exchange between mother and the developing offspring).

The inner mass of embryonic cells is totipotent during this stage, meaning that each cell has the potential to differentiate into any cell type in the human body. Totipotency lasts for only a few days before the cells' fates are set as being the precursors to a specific lineage of cells.

As the blastocyst forms, the trophoblast excretes enzymes that begin to degrade the zona pellucida. In a process called "hatching," the conceptus breaks free of the zona pellucida in preparation for implantation.

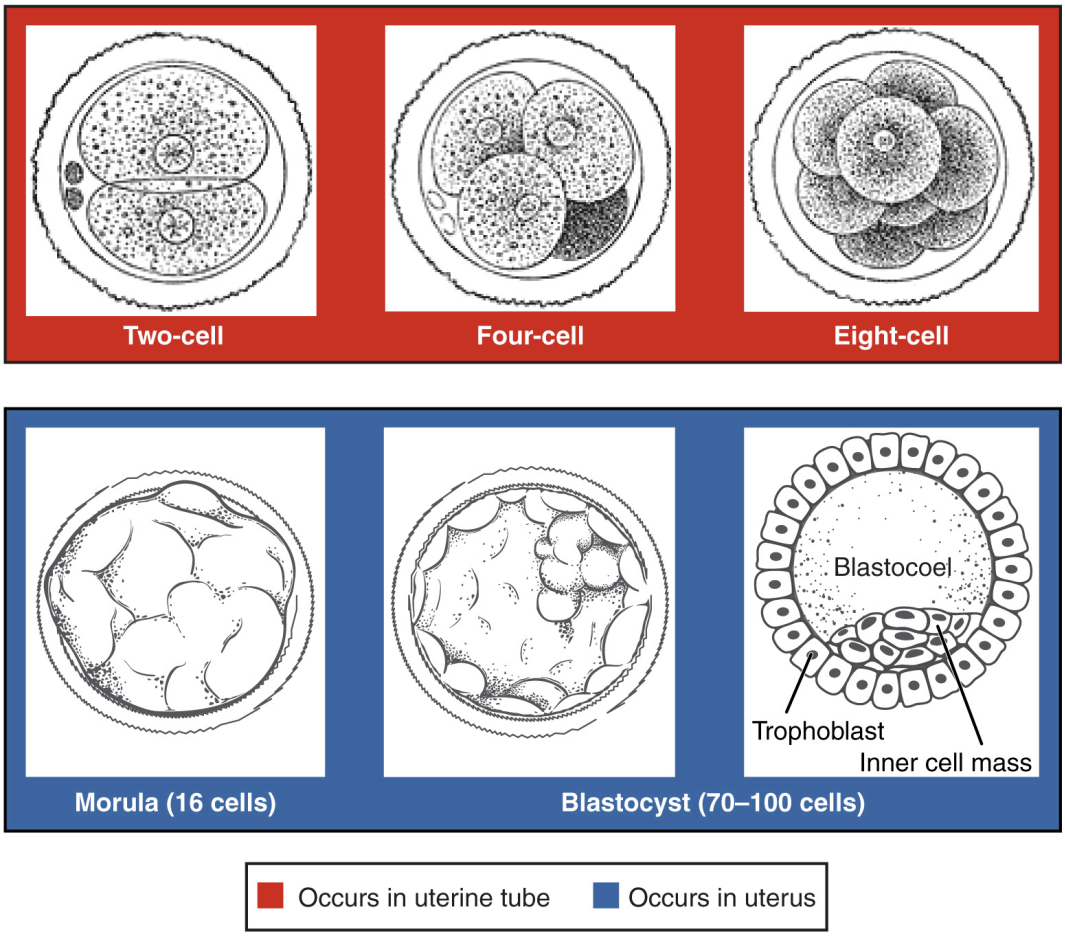


Figure 12. Early Embryonic Development. *Cleavages make use of the abundant cytoplasm of the conceptus as the cells rapidly divide without changing the total volume.*

Implantation: At the end of the first week, the blastocyst comes in contact with the uterine wall and adheres to it, embedding itself in the uterine lining via the trophoblast cells. Thus begins the process of implantation, which signals the end of the embryonic stage of development (Figure 13). Implantation can be accompanied by minor bleeding. The blastocyst typically implants in the fundus of the uterus or on the posterior wall. However, if the endometrium is not fully developed and ready to receive the blastocyst, the blastocyst will detach and find a better spot. A significant percentage (50–75 percent) of blastocysts fail to implant; when this occurs, the blastocyst is shed with the endometrium during menses. The high rate of implantation failure is one reason why pregnancy typically requires several ovulation cycles to be achieved.

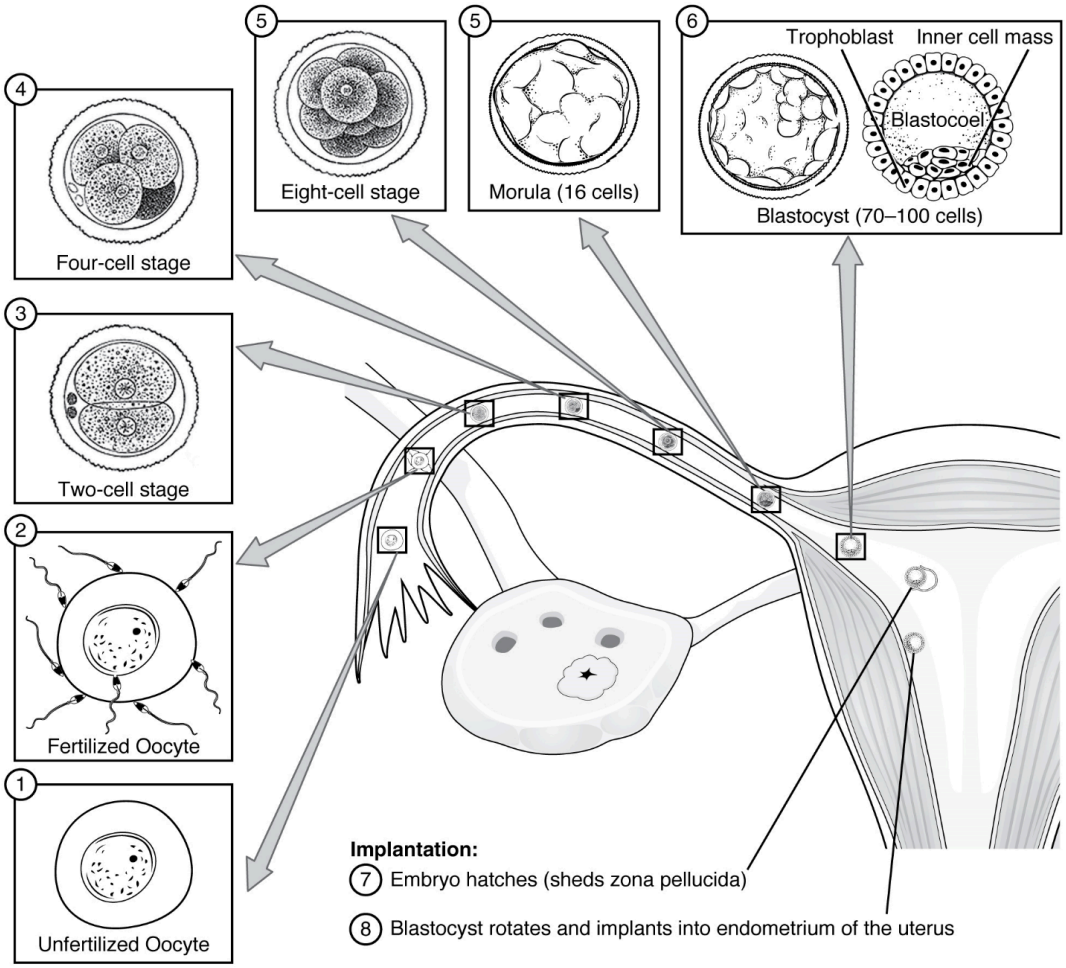


Figure 13. Early Embryonic Development. *Ovulation, fertilization, cleavage embryonic development, and implantation occur at specific locations within the female reproductive system in a time span of approximately 1 week.*

When implantation succeeds and the blastocyst adheres to the endometrium, the superficial cells of the trophoblast fuse with each other, forming the **syncytiotrophoblast**, a multinucleated body that digests endometrial cells to firmly secure the blastocyst to the uterine wall. In response, the uterine mucosa rebuilds itself and envelops the blastocyst (Figure 14). The trophoblast secretes **human chorionic gonadotropin (hCG)**, a hormone that directs the corpus luteum to survive, enlarge, and continue producing progesterone and estrogen to suppress menses. These functions of hCG are necessary for creating an environment suitable for the developing embryo. As a result of this increased production, hCG accumulates in the maternal bloodstream and is excreted in the urine. Implantation is complete by the middle of the second week. Just a few days after implantation, the trophoblast has secreted enough hCG for an at-home urine pregnancy test to give a positive result.

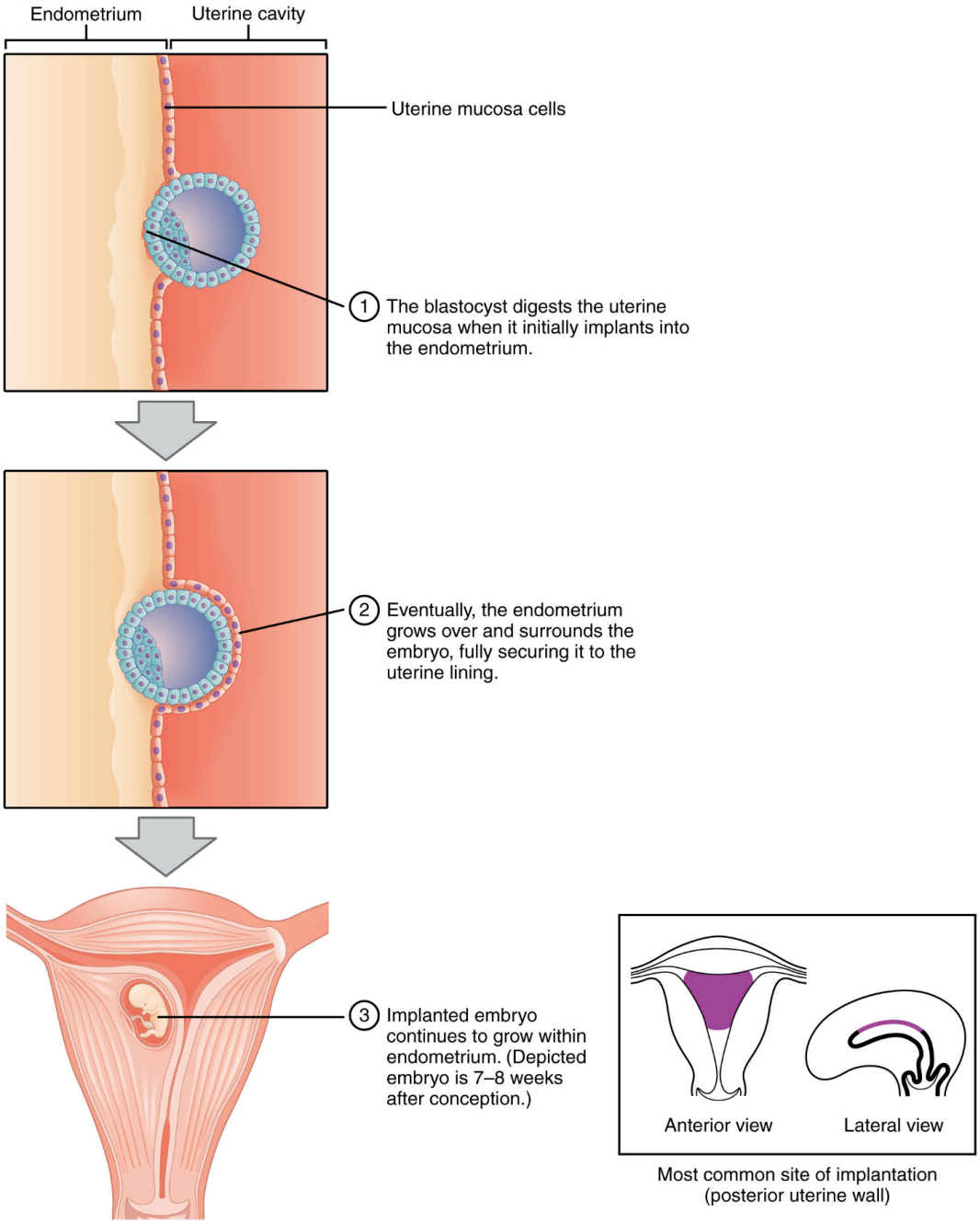


Figure 14. Implantation. *During implantation, the trophoblast cells of the blastocyst adhere to the endometrium and digest endometrial cells until it is attached securely.*

Embryonic Membranes: During the second week of development, with the embryo implanted in the uterus, cells within the blastocyst start to organize into layers. Some grow to form the extra-embryonic membranes needed to support and protect the growing embryo: the amnion, the yolk sac, the allantois, and the chorion (Figure 16). The **amnion** fills with amniotic fluid and eventually grows to surround the embryo. The **yolk sac** supplies some nutrients absorbed from the trophoblast and also provides primitive blood circulation to the developing embryo for the second and third week of development. During week 3, a finger-like outpocketing of the yolk sac develops into the **allantois**, a primitive excretory duct of the embryo that will become part of the urinary bladder. Together, the stalks of the yolk sac and allantois establish the outer structure of the umbilical cord. The last of the extra-embryonic membranes is the **chorion**, which is the one membrane that surrounds all

others. The development of the chorion will be discussed in more detail shortly, as it relates to the growth and development of the placenta.

Germ Layers: As the third week of development begins, the two-layered disc of cells becomes a three-layered disc through the process of gastrulation, during which the cells transition from an undifferentiated to a more differentiated state. Three groups of cells, called germ layers, are formed: the **endoderm**, **mesoderm** (middle), and **ectoderm**.

Each of these germ layers will develop into specific structures in the embryo. Whereas the ectoderm and endoderm form tightly connected epithelial sheets, the mesodermal cells are less organized and exist as a loosely connected cell community. The ectoderm gives rise to cell lineages that differentiate to become the central and peripheral nervous systems, sensory organs, epidermis, hair, and nails. Mesodermal cells ultimately become the skeleton, muscles, connective tissue, heart, blood vessels, and kidneys. The endoderm goes on to form the epithelial lining of the gastrointestinal tract, liver, and pancreas, as well as the lungs (Figure 15).

Development of the Placenta: During the first several weeks of development, the cells of the endometrium—referred to as decidual cells—nourish the nascent embryo. During prenatal weeks 4–12, the developing placenta gradually takes over the role of feeding the embryo, and the decidual cells are no longer needed. The mature placenta is composed of tissues derived from the embryo, as well as maternal tissues of the endometrium. The placenta connects to the conceptus via the **umbilical cord**, which carries deoxygenated blood and wastes from the fetus through two umbilical arteries; nutrients and oxygen are carried from the mother to the fetus through the single umbilical vein.

The maternal portion of the placenta develops from the deepest layer of the endometrium, the decidua basalis. To form the embryonic portion of the placenta, the syncytiotrophoblast and the underlying cells of the trophoblast (cytotrophoblast cells) begin to proliferate along with a layer of extraembryonic mesoderm cells. These form the **chorionic membrane**, which envelops the entire conceptus as the chorion. The chorionic membrane forms finger-like structures called **chorionic villi** that burrow into the endometrium like tree roots, making up the fetal portion of the placenta. The cytotrophoblast cells perforate the chorionic villi, burrow farther into the endometrium, and remodel maternal blood vessels to augment maternal blood flow surrounding the villi. Meanwhile, fetal mesenchymal cells derived from the mesoderm fill the villi and differentiate into blood vessels, including the three umbilical blood vessels that connect the embryo to the developing placenta (Figure 16).

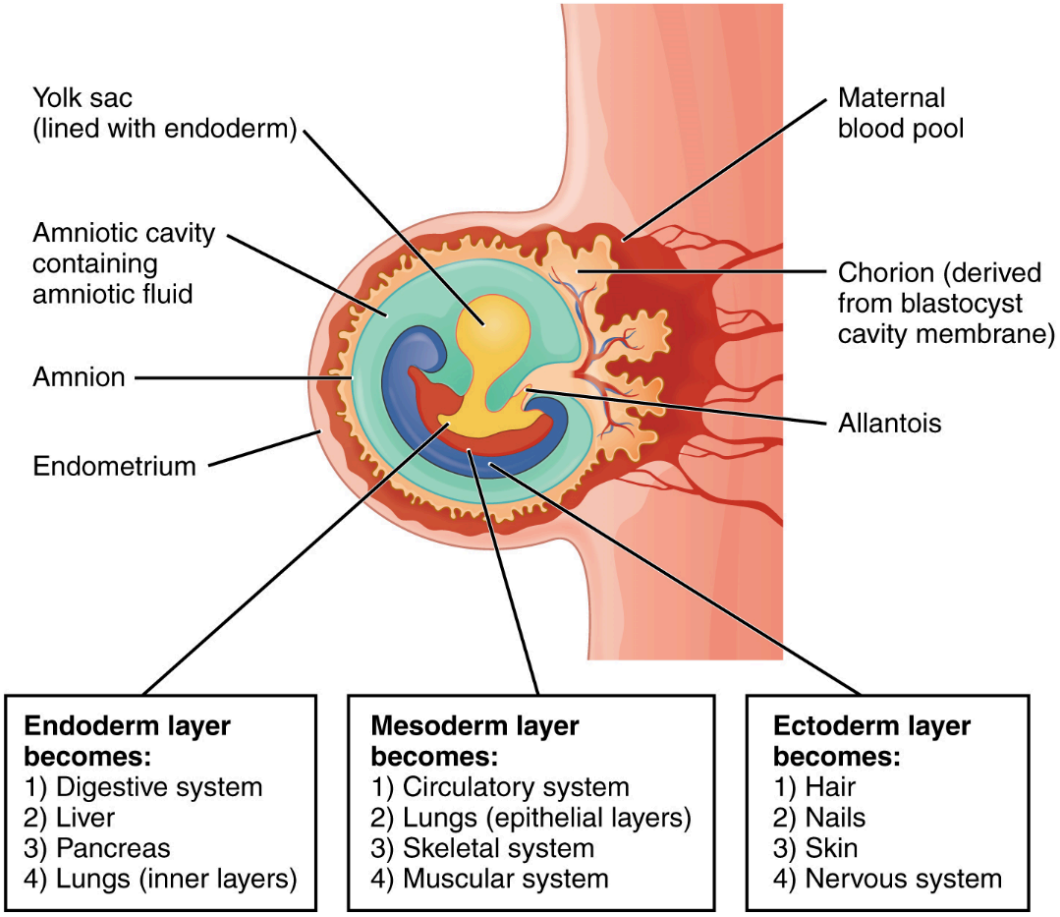


Figure 15. Fates of Germ Layers in Embryo. *Following gastrulation of the embryo in the third week, embryonic cells of the ectoderm, mesoderm, and endoderm begin to migrate and differentiate into the cell lineages that will give rise to mature organs and organ systems in the infant.*

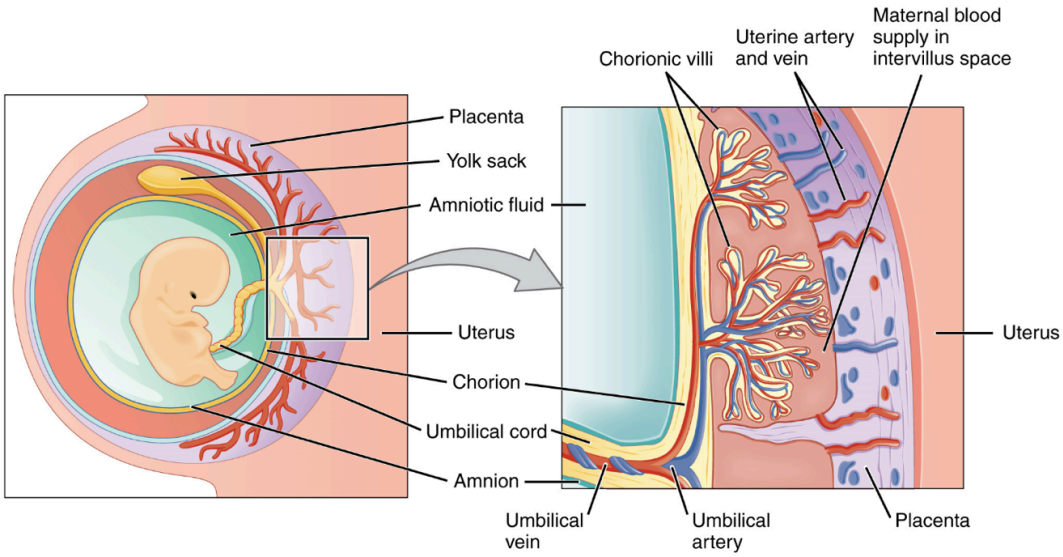


Figure 16. Cross-Section of the Placenta. *In the placenta, maternal and fetal blood components are conducted through the surface of the chorionic villi, but maternal and fetal bloodstreams never mix directly.*

The placenta develops throughout the embryonic period and during the first several weeks of the fetal period; **placentation** is complete by weeks 14–16. As a fully developed organ, the placenta provides nutrition and excretion, respiration, and endocrine function (Table 1 and Figure 17). It receives blood from the fetus through

the umbilical arteries. Capillaries in the chorionic villi filter fetal wastes out of the blood and return clean, oxygenated blood to the fetus through the umbilical vein. Nutrients and oxygen are transferred from maternal blood surrounding the villi through the capillaries and into the fetal bloodstream.

Maternal and fetal blood do not mingle because blood cells cannot move across the placenta. This separation prevents the mother's cytotoxic T cells from reaching and subsequently destroying the fetus, which bears "non-self" antigens.

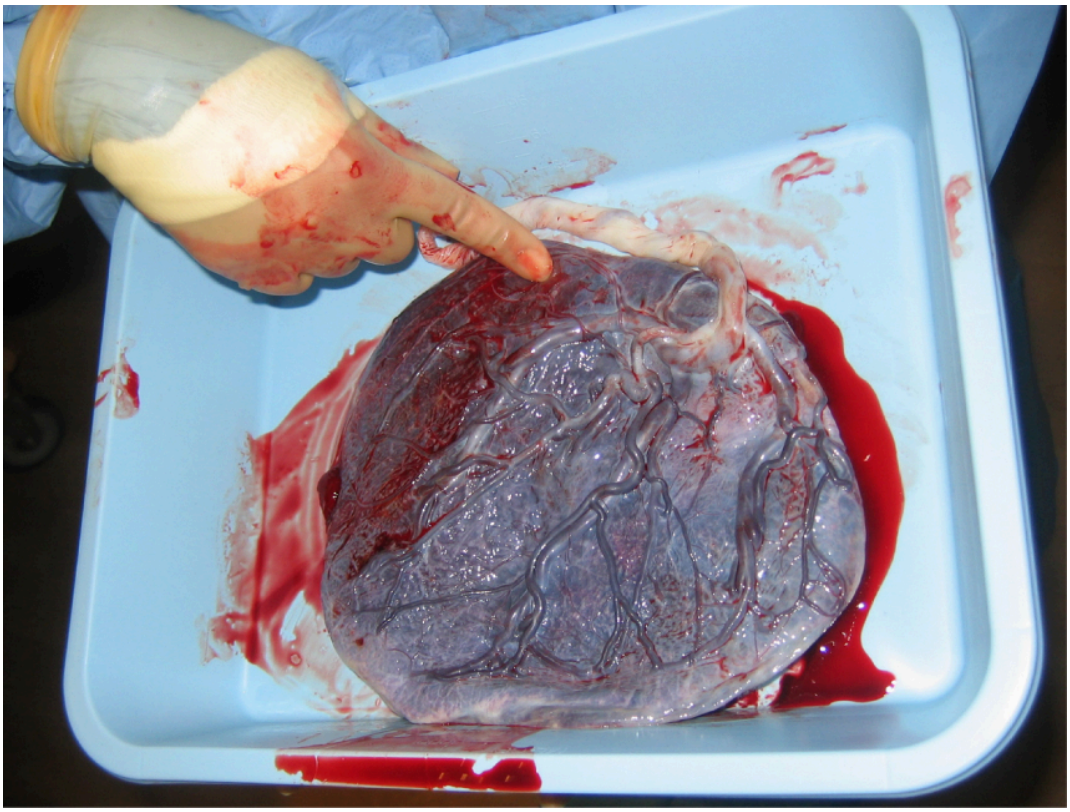


Figure 17. Placenta. *This post-expulsion placenta and umbilical cord (white) are viewed from the fetal side.*

Table 1: Functions of the Placenta

Nutrition & digestion	Respiration	Endocrine function
Mediates diffusion of maternal glucose, amino acids, fatty acids, vitamins, minerals	Mediates maternal-to-fetal oxygen transport	Secretes hCG, estrogens, and progesterone to maintain the pregnancy and stimulate maternal & fetal development
Stores nutrients during early pregnancy to accommodate increased fetal demand later in pregnancy	Mediates fetal-to-maternal carbon dioxide transport	Mediates transmission of maternal hormones into fetal blood
Excretes & filters fetal nitrogenous wastes into maternal blood		Mediates transmission of fetal hormones into maternal blood

During the sixth week, uncontrolled fetal limb movements begin to occur. The gastrointestinal system develops too rapidly for the embryonic abdomen to accommodate it, and the intestines temporarily loop into the umbilical cord. Paddle-shaped hands and feet develop fingers and toes by the process of apoptosis (programmed cell death), which causes the tissues between the fingers to disintegrate. By week 7, the facial

structure is more complex and includes nostrils, outer ears, and lenses (Figure 18). By the eighth week, the head is nearly as large as the rest of the embryo's body, and all major brain structures are in place. The external genitalia are apparent, but at this point, male and female embryos are indistinguishable. Bone begins to replace cartilage in the embryonic skeleton through the process of ossification. By the end of the embryonic period, the embryo is approximately 3 cm (1.2 in) from crown to rump and weighs approximately 8 g (0.25 oz).

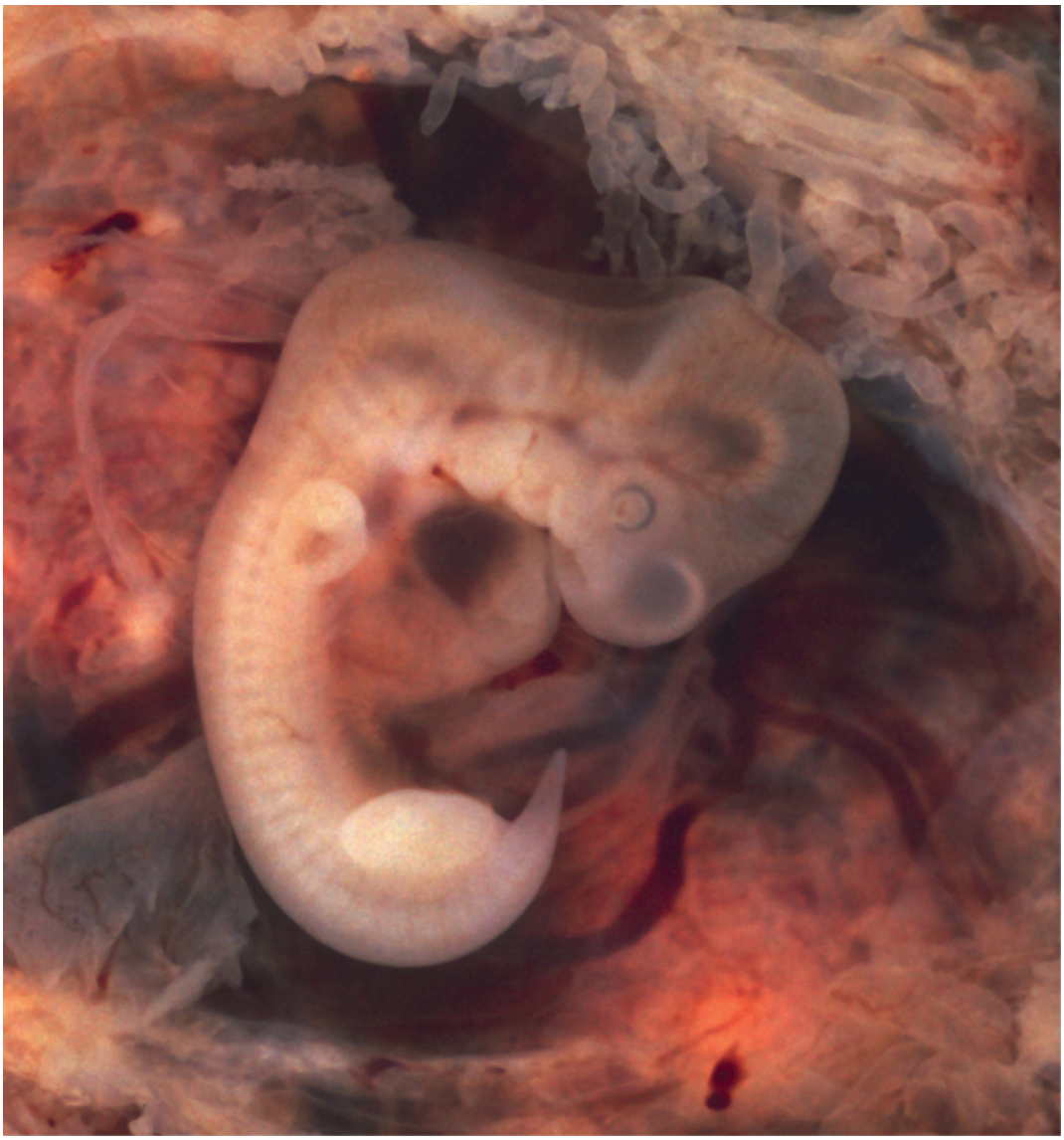


Figure 18. Embryo at 7 Weeks. *An embryo at the end of 7 weeks of development is only 10 mm in length, but its developing eyes, limb buds, and tail are already visible. (This embryo was derived from an ectopic pregnancy.) (credit: Ed Uthman)*

Watch this Crash Course video to learn about the stages of embryonic development! Direct link: https://youtu.be/BtsSbZ85yiQ

Part 5: Fetal Development

A developing human is called a fetus from the ninth week of gestation until birth. This 30-week period of development is marked by continued cell growth and differentiation, which fully develop the structures and functions of the immature organ systems formed during the embryonic period. The completion of fetal development results in a newborn who, although still immature in many ways, is capable of survival outside the womb.

The Fetal Circulatory System: During prenatal development, the fetal circulatory system is integrated with the placenta via the umbilical cord so that the fetus receives both oxygen and nutrients from the placenta. However, after childbirth, the umbilical cord is severed, and the newborn's circulatory system must be reconfigured. When the heart first forms in the embryo, it exists as two parallel tubes derived from mesoderm and lined with endothelium, which then fuse together. As the embryo develops into a fetus, the tube-shaped heart folds and further differentiates into the four chambers present in a mature heart. Unlike a mature cardiovascular system, however, the fetal cardiovascular system also includes circulatory shortcuts, or shunts. A **shunt** is an anatomical (or sometimes surgical) diversion that allows blood flow to bypass immature organs such as the lungs and liver until childbirth.

The placenta provides the fetus with necessary oxygen and nutrients via the umbilical vein. (Remember that veins carry blood toward the heart. In this case, the blood flowing to the fetal heart is oxygenated because it comes from the placenta. The respiratory system is immature and cannot yet oxygenate blood on its own.)

From the umbilical vein, the oxygenated blood flows toward the inferior vena cava, all but bypassing the immature liver, via the **ductus venosus** shunt (Figure 19). The liver receives just a trickle of blood, which is all that it needs in its immature, semifunctional state. Blood flows from the inferior vena cava to the right atrium, mixing with fetal venous blood along the way.

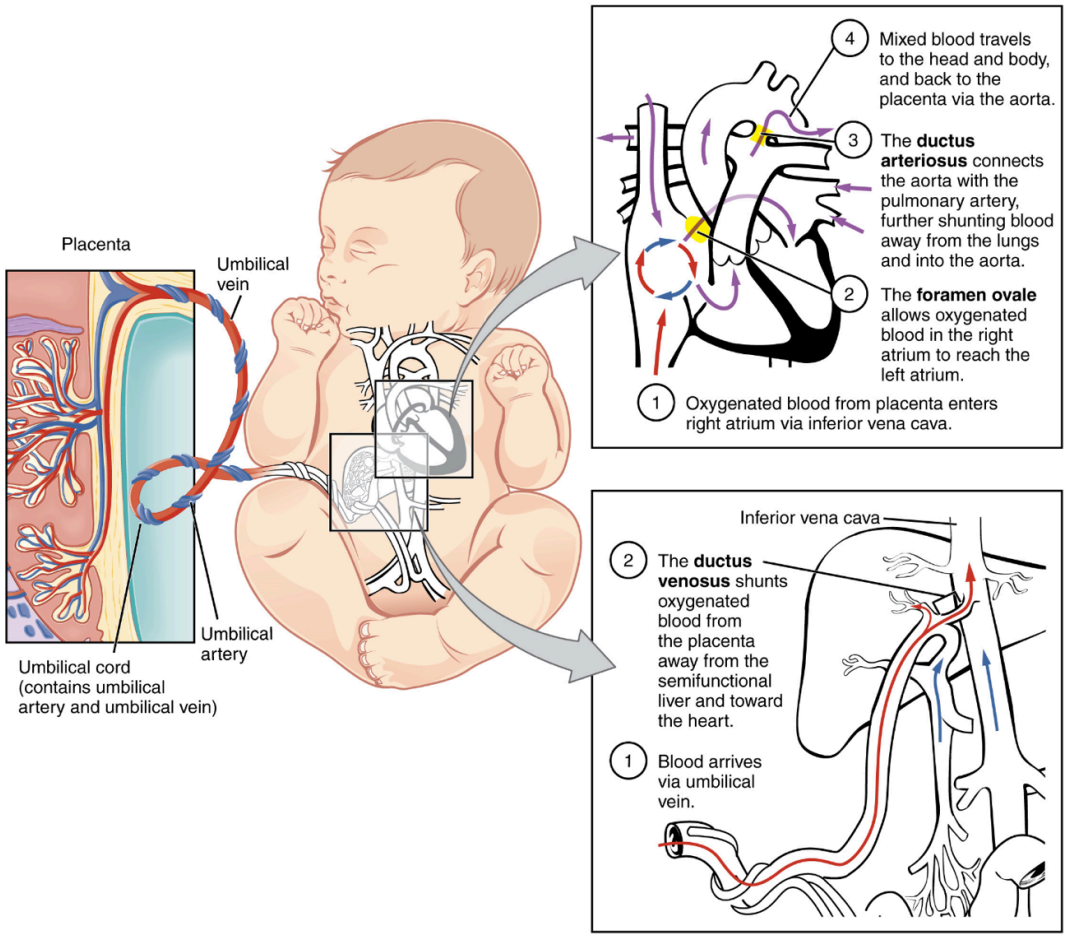


Figure 19. Fetal Circulatory System. *The fetal circulatory system includes three shunts to divert blood from undeveloped and partially functioning organs, as well as blood supply to and from the placenta.*

Although the fetal liver is semifunctional, the fetal lungs are nonfunctional. The fetal circulation therefore bypasses the lungs by shifting some of the blood through the **foramen ovale**, a shunt that directly connects the right and left atria and avoids the pulmonary trunk altogether. Most of the rest of the blood is pumped to the right ventricle, and from there, into the pulmonary trunk, which splits into pulmonary arteries. However, a shunt within the pulmonary artery, the **ductus arteriosus**, diverts a portion of this blood into the aorta. This ensures that only a small volume of oxygenated blood passes through the immature pulmonary circuit, which has only minor metabolic requirements. Blood vessels of uninflated lungs have high resistance to flow, a condition that encourages blood to flow to the aorta, which presents much lower resistance. The oxygenated blood moves through the foramen ovale into the left atrium, where it mixes with the now deoxygenated blood returning from the pulmonary circuit. This blood then moves into the left ventricle, where it is pumped into the aorta. Some of this blood moves through the coronary arteries into the myocardium, and some moves through the carotid arteries to the brain.

The descending aorta carries partially oxygenated and partially deoxygenated blood into the lower regions of the body. It eventually passes into the umbilical arteries through branches of the internal iliac arteries. The deoxygenated blood collects waste as it circulates through the fetal body and returns to the umbilical cord. Thus, the two umbilical arteries carry blood low in oxygen and high in carbon dioxide and fetal wastes. This blood is filtered through the placenta, where wastes diffuse into the maternal circulation. Oxygen and nutrients from the mother diffuse into the placenta and from there into the fetal blood, and the process repeats.

Prior to birth, the lungs are filled with amniotic fluid, mucus, and surfactant. As the fetus is squeezed through the birth canal, the fetal thoracic cavity is compressed, expelling much of this fluid. Some fluid remains, however, but is rapidly absorbed by the body shortly after birth. The first inhalation occurs within 10 seconds after birth and not only serves as the first inspiration, but also acts to inflate the lungs. **Pulmonary surfactant** is critical for inflation to occur, as it reduces the surface tension of the alveoli. Preterm birth around 26 weeks frequently results in severe respiratory distress, although with current medical advancements, some babies may survive. Prior to 26 weeks, sufficient pulmonary surfactant is not produced, and the surfaces for gas exchange have not formed adequately; therefore, survival is low.

During approximately weeks 16–20, as the fetus grows and limb movements become more powerful, the mother may begin to feel **quickening**, or fetal movements. However, space restrictions limit these movements and typically force the growing fetus into the "fetal position," with the arms crossed and the legs bent at the knees. Sebaceous glands coat the skin with a waxy, protective substance called **vernix caseosa** that protects and moisturizes the skin and may provide lubrication during childbirth. A silky hair called lanugo also covers the skin during weeks 17–20, but it is shed as the fetus continues to grow.

Developmental weeks 21–30 are characterized by rapid weight gain, which is important for maintaining a stable body temperature after birth. The bone marrow completely takes over erythrocyte synthesis, and the axons of the spinal cord begin to be myelinated, or coated in the electrically insulating glial cell sheaths that are necessary for efficient nervous system functioning. (The process of myelination is not completed until adolescence.) During this period, the fetus grows eyelashes. The eyelids are no longer fused and can be opened and closed. The lungs begin producing surfactant, a substance that reduces surface tension in the lungs and assists proper lung expansion after birth. Inadequate surfactant production in premature newborns may result in respiratory distress syndrome, and as a result, the newborn may require surfactant replacement therapy, supplemental oxygen, or maintenance in a continuous positive airway pressure (CPAP) chamber during their first days or weeks of life. In male fetuses, the testes descend into the scrotum near the end of this period. The fetus at 30 weeks measures 28 cm (11 in) from crown to rump and exhibits the approximate body proportions of a full-term newborn, but still is much leaner.

The fetus continues to lay down subcutaneous fat from week 31 until birth. The added fat fills out the hypodermis, and the skin transitions from red and wrinkled to soft and pink. Lanugo is shed, and the nails grow to the tips of the fingers and toes. Immediately before birth, the average crown-to-rump length is 35.5–40.5 cm (14–16 in), and the fetus weighs approximately 2.5–4 kg (5.5–8.8 lbs). Once born, the newborn is no longer

confined to the fetal position, so subsequent measurements are made from head-to-toe instead of from crown-to-rump. At birth, the average length is approximately 51 cm (20 in).

Part 6: Maternal Changes During Pregnancy, Labor, and Birth

A full-term pregnancy lasts approximately 270 days (approximately 38.5 weeks) from conception to birth. Because it is easier to remember the first day of the last menstrual period (LMP) than to estimate the date of conception, obstetricians set the due date as 284 days (approximately 40.5 weeks) from the LMP. This assumes that conception occurred on day 14 of the woman's cycle, which is usually a good approximation. The 40 weeks of an average pregnancy are usually discussed in terms of three trimesters, each approximately 13 weeks. During the second and third trimesters, the pre-pregnancy uterus—about the size of a fist—grows dramatically to contain the fetus, causing a number of anatomical changes in the mother (Figure 20).

Effects of Hormones: Virtually all the effects of pregnancy can be attributed in some way to the influence of hormones—particularly estrogens, progesterone, and hCG. During weeks 7–12 from the LMP, the pregnancy hormones are primarily generated by the corpus luteum. Progesterone secreted by the corpus luteum stimulates the production of decidual cells of the endometrium that nourish the blastocyst before placentation. As the placenta develops and the corpus luteum degenerates during weeks 12–17, the placenta gradually takes over as the endocrine organ of pregnancy.

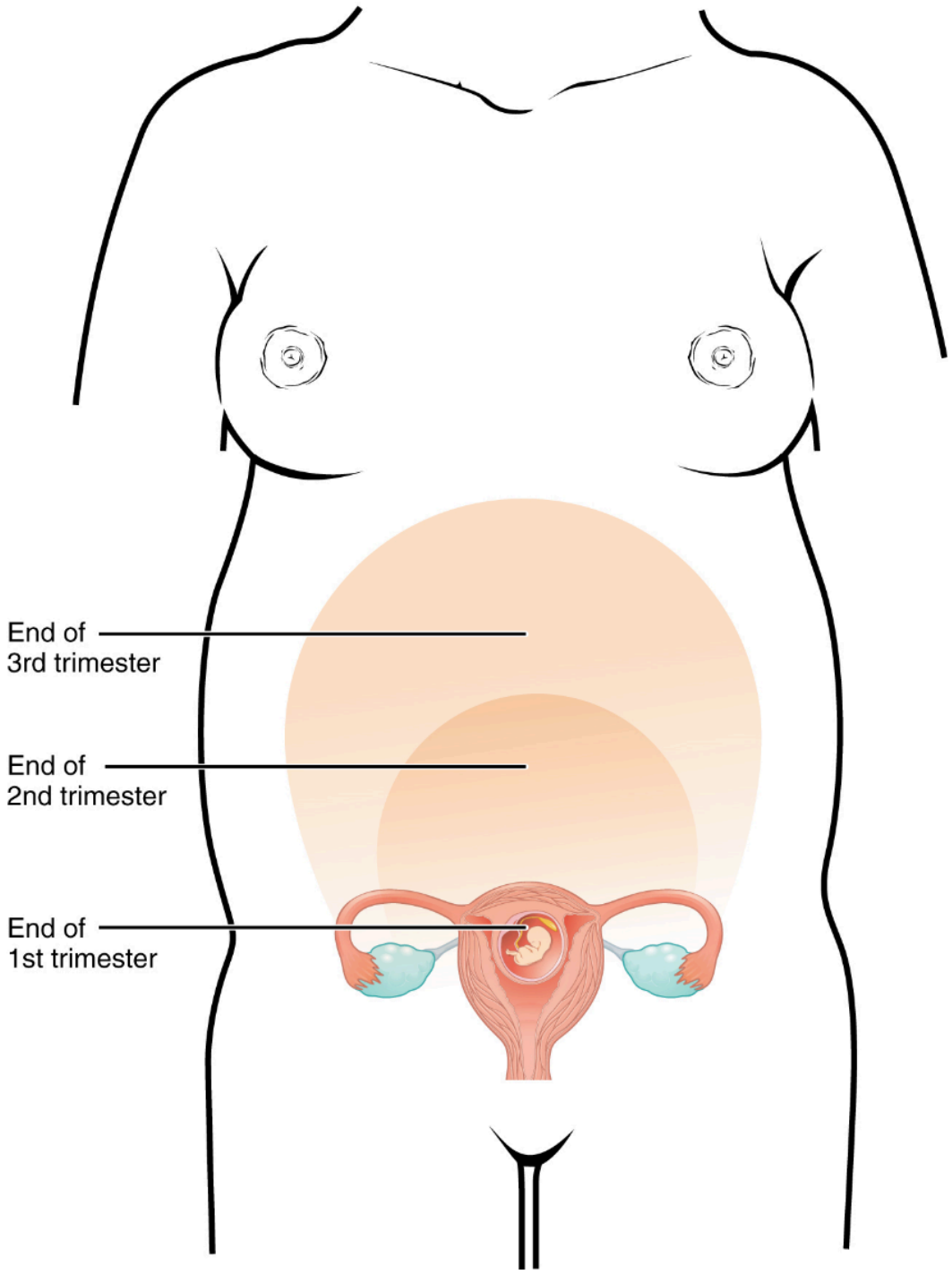


Figure 20. Size of Uterus Throughout Pregnancy. *The uterus grows throughout pregnancy to accommodate the fetus.*

The placenta converts weak androgens secreted by the maternal and fetal adrenal glands to estrogens, which are necessary for pregnancy to progress. Estrogen levels climb throughout the pregnancy, increasing 30-fold by childbirth. Estrogens have the following actions:

- They suppress FSH and LH production, effectively preventing ovulation. (This function is the biological basis of hormonal birth control pills.)
- They induce the growth of fetal tissues and are necessary for the maturation of the fetal lungs and liver.

- They promote fetal viability by regulating progesterone production and triggering fetal synthesis of cortisol, which helps with the maturation of the lungs, liver, and endocrine organs such as the thyroid gland and adrenal gland.
- They stimulate maternal tissue growth, leading to uterine enlargement and mammary duct expansion and branching.

Relaxin, another hormone secreted by the corpus luteum and then by the placenta, helps prepare the mother's body for childbirth. It increases the elasticity of the symphysis pubis joint and pelvic ligaments, making room for the growing fetus and allowing expansion of the pelvic outlet for childbirth. Relaxin also helps dilate the cervix during labor.

The placenta takes over the synthesis and secretion of progesterone throughout pregnancy as the corpus luteum degenerates. Like estrogen, progesterone suppresses FSH and LH. It also inhibits uterine contractions, protecting the fetus from preterm birth. This hormone decreases in late gestation, allowing uterine contractions to intensify and eventually progress to true labor. The placenta also produces hCG. In addition to promoting survival of the corpus luteum, hCG stimulates the male fetal gonads to secrete testosterone, which is essential for the development of the male reproductive system.

Physiology of Labor: Childbirth, or parturition, typically occurs within a week of a woman's due date, unless the woman is pregnant with more than one fetus, which usually causes her to go into labor early. As a pregnancy progresses into its final weeks, several physiological changes occur in response to hormones that trigger labor.

First, recall that progesterone inhibits uterine contractions throughout the first several months of pregnancy. As the pregnancy enters its seventh month, progesterone levels plateau and then drop. Estrogen levels, however, continue to rise in the maternal circulation (Figure 21). The increasing ratio of estrogen to progesterone makes the myometrium (the uterine smooth muscle) more sensitive to stimuli that promote contractions (because progesterone no longer inhibits them). Moreover, in the eighth month of pregnancy, fetal cortisol rises, which boosts estrogen secretion by the placenta and further overpowers the uterine-calming effects of progesterone.

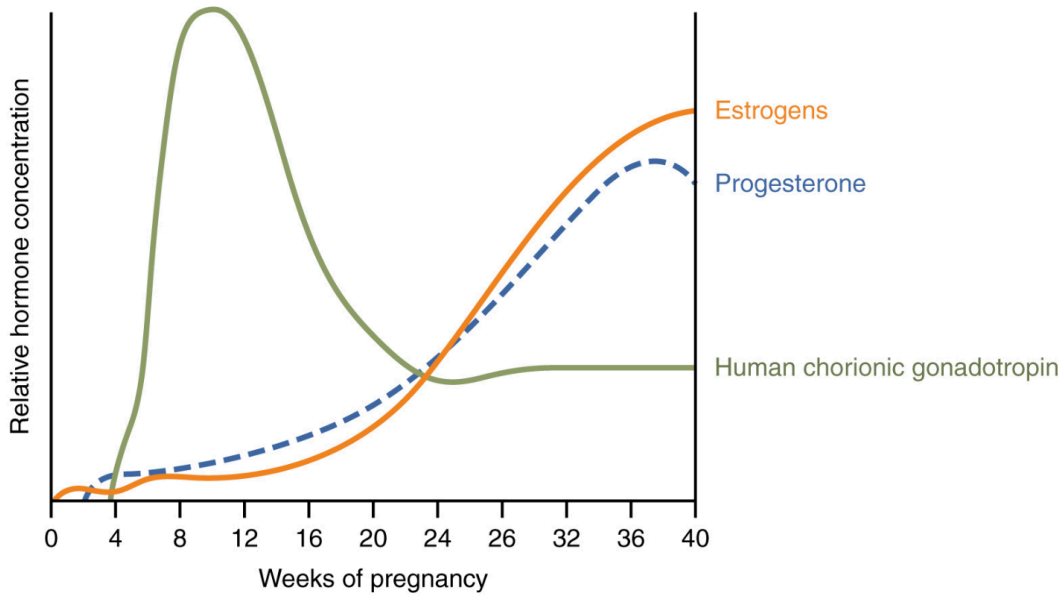


Figure 21. Hormones Initiating Labor. *A positive feedback loop of hormones works to initiate labor.*

Meanwhile, the posterior pituitary has been boosting its secretion of oxytocin, a hormone that stimulates the contractions of labor. At the same time, the myometrium increases its sensitivity to oxytocin by expressing

more receptors for this hormone. As labor nears, oxytocin begins to stimulate stronger, more painful uterine contractions, which—in a positive feedback loop—stimulate the secretion of prostaglandins from fetal membranes. Like oxytocin, prostaglandins also enhance uterine contractile strength.

Finally, stretching of the myometrium and cervix by a full-term fetus in the vertex (head-down) position is regarded as a stimulant to uterine contractions. The sum of these changes initiates the regular contractions known as true labor, which become more powerful and more frequent with time. The pain of labor is attributed to myometrial hypoxia during uterine contractions.

Stages of Childbirth: The process of childbirth can be divided into three stages: cervical dilation, expulsion of the newborn ending with birth, and afterbirth (Figure 22).

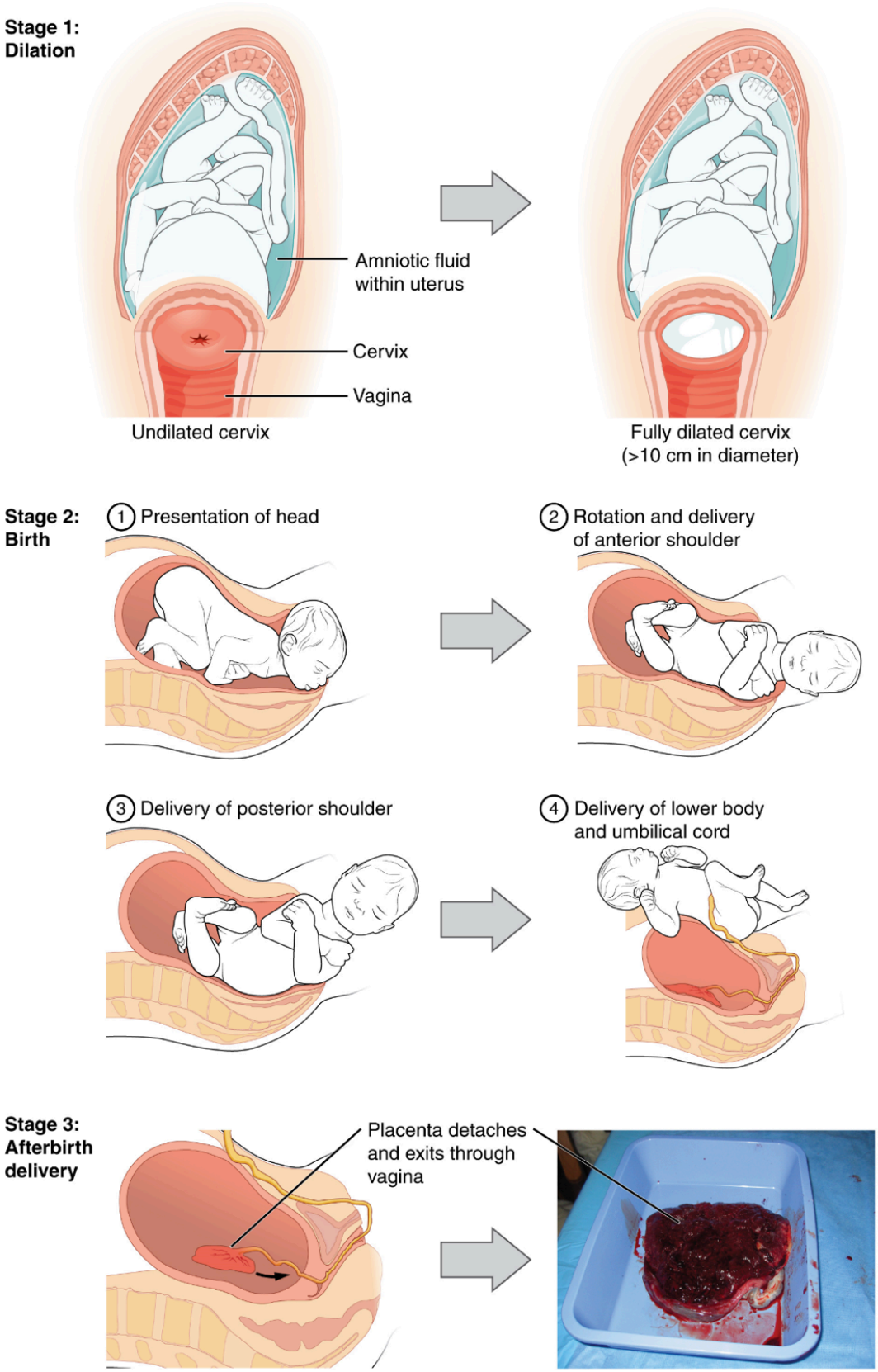


Figure 22. Stages of Childbirth. *The stages of childbirth begin with early cervical dilation, continue through full dilation and birth (expulsion of the newborn); and finally delivery of the placenta and associated fetal membranes. The position of the newborn's shoulder is described relative to the mother.*

1. Cervical Dilation: For vaginal birth to occur, the cervix must dilate fully to 10 cm in diameter—wide enough

to deliver the newborn's head. The dilation stage is the longest stage of labor and typically takes 6–12 hours. However, it varies widely and may take minutes, hours, or days, depending in part on whether the mother has given birth before; in each subsequent labor, this stage tends to be shorter.

True labor progresses in a positive feedback loop in which uterine contractions stretch the cervix, causing it to dilate and efface, or become thinner. Cervical stretching induces reflexive uterine contractions that dilate and efface the cervix further. In addition, cervical dilation boosts oxytocin secretion from the pituitary, which in turn triggers more powerful uterine contractions. When labor begins, uterine contractions may occur only every 3–30 minutes and last only 20–40 seconds; however, by the end of this stage, contractions may occur as frequently as every 1.5–2 minutes and last for a full minute.

Each contraction sharply reduces oxygenated blood flow to the fetus. For this reason, it is critical that a period of relaxation occur after each contraction. Fetal distress, measured as a sustained decrease or increase in the fetal heart rate, can result from severe contractions that are too powerful or lengthy for oxygenated blood to be restored to the fetus. Such a situation can be cause for an emergency birth with vacuum, forceps, or surgically by Caesarian section.

2. Expulsion Stage: The expulsion stage begins when the fetal head enters the birth canal and ends with birth of the newborn. It typically takes up to 2 hours, but it can last longer or be completed in minutes, depending in part on the orientation of the fetus. The most common presentation and associated with the greatest ease of vaginal birth is when the fetus faces the maternal spinal cord and the smallest part of the head (the posterior aspect called the occiput) exits the birth canal first.

Upon birth of the newborn's head, an obstetrician will aspirate mucus from the mouth and nose before the newborn's first breath. Once the head is birthed, the rest of the body usually follows quickly. The umbilical cord is then double-clamped, and a cut is made between the clamps. This completes the second stage of childbirth.

3. Afterbirth: The delivery of the placenta and associated membranes, commonly referred to as the afterbirth, marks the final stage of childbirth. After expulsion of the newborn, the myometrium continues to contract. This movement shears the placenta from the back of the uterine wall. It is then easily delivered through the vagina. Continued uterine contractions then reduce blood loss from the site of the placenta. Delivery of the placenta marks the beginning of the postpartum period—the period of approximately 6 weeks immediately following childbirth during which the mother's body gradually returns to a non-pregnant state. If the placenta does not birth spontaneously within approximately 30 minutes, it is considered retained, and the obstetrician may attempt manual removal.

Part 7: Adjustments of the Infant at Birth and Postnatal Stages

From a fetal perspective, the process of birth is a crisis. In the womb, the fetus was snuggled in a soft, warm, dark, and quiet world. The placenta provided nutrition and oxygen continuously. Suddenly, the contractions of labor and vaginal childbirth forcibly squeeze the fetus through the birth canal, limiting oxygenated blood flow during contractions and shifting the skull bones to accommodate the small space. After birth, the newborn's system must make drastic adjustments to a world that is colder, brighter, and louder, and where he or she will experience hunger and thirst. The neonatal period (neo- = "new"; -natal = "birth") spans the first to the thirtieth day of life outside of the uterus.

Respiratory Adjustments: Although the fetus "practices" breathing by inhaling amniotic fluid in utero, there is no air in the uterus and thus no true opportunity to breathe. (There is also no need to breathe because the placenta supplies the fetus with all the oxygenated blood it needs.) During gestation, the partially collapsed lungs are filled with amniotic fluid and exhibit very little metabolic activity. Several factors stimulate newborns to take their first breath at birth. First, labor contractions temporarily constrict umbilical blood vessels, reducing oxygenated blood flow to the fetus and elevating carbon dioxide levels in the blood. High carbon dioxide levels cause acidosis and stimulate the respiratory center in the brain, triggering the newborn to take a breath.

The first breath typically is taken within 10 seconds of birth, after mucus is aspirated from the infant's mouth and nose. The first breaths inflate the lungs to nearly full capacity and dramatically decrease lung pressure

and resistance to blood flow, causing a major circulatory reconfiguration. Pulmonary alveoli open, and alveolar capillaries fill with blood. Amniotic fluid in the lungs drains or is absorbed, and the lungs immediately take over the task of the placenta, exchanging carbon dioxide for oxygen by the process of respiration.

Circulatory Adjustments

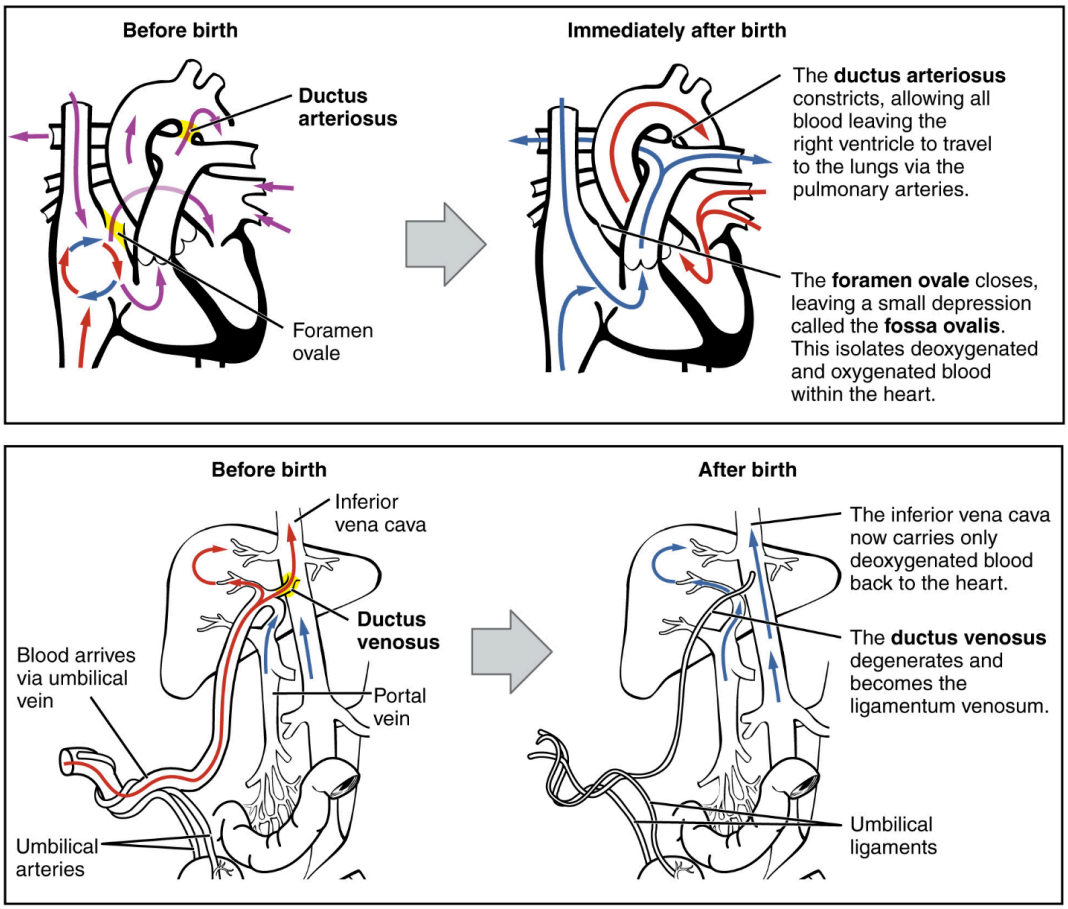


Figure 23. Neonatal Circulatory System. *A newborn's circulatory system reconfigures immediately after birth. The three fetal shunts have been closed permanently, facilitating blood flow to the liver and lungs.*

The newborn's first breath is vital to initiate the transition from the fetal to the neonatal circulatory pattern. Inflation of the lungs decreases blood pressure throughout the pulmonary system, as well as in the right atrium and ventricle. In response to this pressure change, the flow of blood temporarily reverses direction through the foramen ovale, moving from the left to the right atrium, and blocking the shunt with two flaps of tissue. Within 1 year, the tissue flaps usually fuse over the shunt, turning the foramen ovale into the fossa ovalis (Figure 23). The ductus arteriosus constricts as a result of increased oxygen concentration and becomes the ligamentum arteriosum (Figure 23). Closing of the ductus arteriosus ensures that all blood pumped to the pulmonary circuit will be oxygenated by the newly functional neonatal lungs. The ductus venosus degenerates to become the ligamentum venosum beneath the liver.

Part 8: Lactation

Lactation is the process by which milk is synthesized and secreted from the mammary glands of the postpartum female breast in response to an infant sucking at the nipple. Breast milk provides ideal nutrition and passive immunity for the infant, encourages mild uterine contractions to return the uterus to its pre-pregnancy size (i.e., involution), and induces a substantial metabolic increase in the mother, consuming the fat reserves stored during pregnancy.

Structure of the Lactating Breast: Mammary glands are modified sweat glands. The non-pregnant and

non-lactating female breast is composed primarily of adipose and collagenous tissue, with mammary glands making up a very minor proportion of breast volume. The mammary gland is composed of milk-transporting lactiferous ducts, which expand and branch extensively during pregnancy in response to estrogen, growth hormone, cortisol, and prolactin. Moreover, in response to progesterone, clusters of breast alveoli bud from the ducts and expand outward toward the chest wall. Breast alveoli are balloon-like structures lined with milk-secreting cuboidal cells, or lactocytes, that are surrounded by a net of contractile myoepithelial cells. Milk is secreted from the lactocytes, fills the alveoli, and is squeezed into the ducts. Clusters of alveoli that drain to a common duct are called lobules; the lactating female has 12–20 lobules organized radially around the nipple. Milk drains from lactiferous ducts into lactiferous sinuses that meet at 4 to 18 perforations in the nipple, called nipple pores.

The Process of Lactation: The pituitary hormone **prolactin** is instrumental in the establishment and maintenance of breast milk supply. It also is important for the mobilization of maternal micronutrients for breast milk.

Near the fifth week of pregnancy, the level of circulating prolactin begins to increase, eventually rising to approximately 10–20 times the pre-pregnancy concentration. During pregnancy, prolactin and other hormones prepare the breasts anatomically for the secretion of milk. The level of prolactin plateaus in late pregnancy, at a level high enough to initiate milk production. However, estrogen, progesterone, and other placental hormones inhibit prolactin-mediated milk synthesis during pregnancy. It is not until the placenta is expelled that this inhibition is lifted and milk production commences.

After childbirth, the baseline prolactin level drops sharply, but it is restored for a 1-hour spike during each feeding to stimulate the production of milk for the next feeding. With each prolactin spike, estrogen and progesterone also increase slightly.

When the infant suckles, sensory nerve fibers in the areola trigger a neuroendocrine reflex that results in milk secretion from lactocytes into the alveoli. The posterior pituitary releases oxytocin, which stimulates myoepithelial cells to squeeze milk from the alveoli so it can drain into the lactiferous ducts, collect in the lactiferous sinuses, and discharge through the nipple pores. It takes less than 1 minute from the time when an infant begins suckling (the latent period) until milk is secreted via the **let-down reflex** (Figure 24).

Changes in the Composition of Breast Milk: In the final weeks of pregnancy, the alveoli swell with colostrum, a thick, yellowish substance that is high in protein but contains less fat and glucose than mature breast milk. Colostrum is secreted during the first 48–72 hours postpartum. Only a small volume of colostrum is produced—approximately 3 ounces in a 24-hour period—but it is sufficient for the newborn in the first few days of life. Colostrum is rich with immunoglobulins, which confer gastrointestinal, and also likely systemic, immunity as the newborn adjusts to a nonsterile environment.

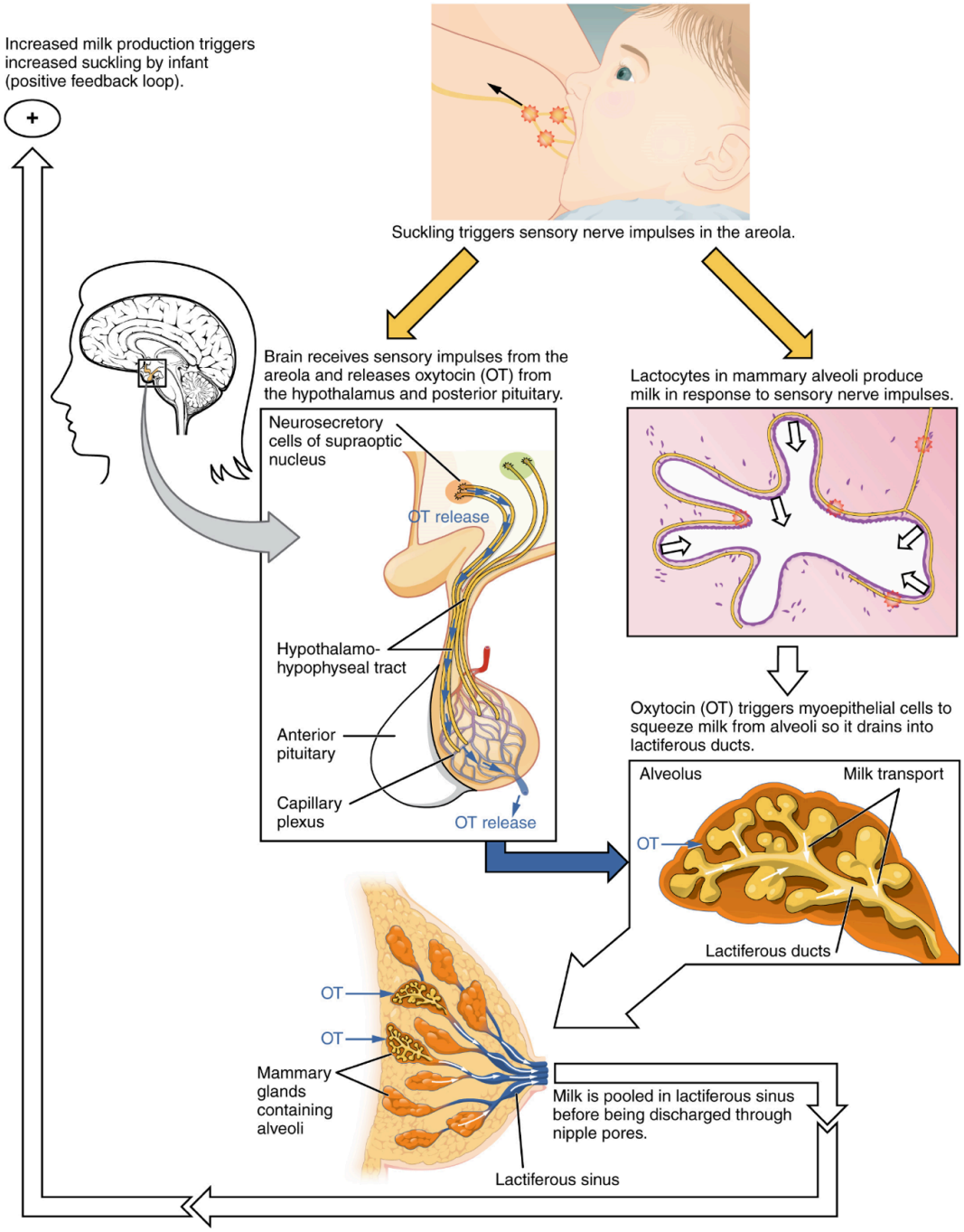


Figure 24. Let-Down Reflex. *A positive feedback loop ensures continued milk production as long as the infant continues to breastfeed.*

Unit 11: Human Genetics

Unit Outline

Part 1: Cell Growth and Division

- The cell cycle
- Interphase
- The structure of chromosomes
- Mitosis and cytokinesis
- Meiosis

Part 2: Patterns of Inheritance

- From genotype to phenotype
- Mendel's theory of inheritance
- Autosomal dominant inheritance
- Autosomal recessive inheritance
- X-linked dominant or recessive inheritance
- Mutations
- Chromosomal disorders
- Detecting genetic disorders

Learning Objectives

At the end of this unit, you should be able to:

I. Distinguish between:

1. Genes and alleles
2. Characters and traits
3. Genotype and phenotype
4. Chromatin, chromosome, and chromatid
5. Haploid and diploid cells

II. Describe the difference between the karyotypes of a normal human female and a normal human male.

III. Describe the characteristics and appearance of cells during each stage of mitosis and meiosis.

IV. Compare and contrast the processes of mitosis and meiosis.

V. Define the term nondisjunction.

VI. Describe the two processes that occur during meiosis that generate genetic variability between gametes (crossing over and independent assortment).

VII. Distinguish between:

1. Homozygous and heterozygous genotypes
2. Dominant and recessive alleles
3. Autosomal and sex-linked inheritance

VIII. Provide the genotypes and phenotypes, of both the parents and offspring, in the following cases:

1. An autosomal gene when:
    - Both parents are heterozygous
    - One parent is heterozygous, and the other is homozygous for the dominant allele
    - One parent is heterozygous, and the other is homozygous for the recessive allele
    - Both parents are homozygous for the recessive allele
    - Both parents are homozygous for the dominant allele
2. A sex-linked gene when:
    - The mother is heterozygous, and the father has the dominant allele
    - The mother is heterozygous, and the father has the recessive allele
    - The mother is homozygous for the dominant allele, and the father has the recessive allele
    - The mother is homozygous for the dominant allele, and the father has the dominant allele
    - The mother is homozygous for the recessive allele, and the father has the recessive allele
    - The mother is homozygous for the recessive allele, and the father has the dominant allele

IX. What is a mutation? Which type of cells would need to mutate for that mutation to be expressed in any offspring?

X. What is a wild-type allele?

XI. Describe the genetic determinants of genetic disorders.

Learning Objectives and Guiding Questions

At the end of this unit, you should be able to complete all the following tasks, including answering the guiding questions associated with each task.

I. Distinguish between:

1. Genes and alleles
2. Characters and traits

3. Genotype and phenotype
4. Chromatin, chromosome, and chromatid
5. Haploid and diploid cells

II. Describe the difference between the karyotypes of a normal human female and a normal human male.

III. Describe the characteristics and appearance of cells during each stage of mitosis and meiosis.

1. Draw a series of diagrams showing the appearance of chromosomes in a cell during each stage of mitosis.
2. Draw a series of diagrams showing the appearance of chromosomes in a cell during each stage of meiosis I.
3. Draw a series of diagrams showing the appearance of chromosomes in a cell during each stage of meiosis II.

IV. Compare and contrast the processes of mitosis and meiosis.

V. Define the term nondisjunction.

1. Describe when it can occur and what the result of it is.

VI. Describe the two processes that occur during meiosis that generate genetic variability between gametes (crossing over and independent assortment).

VII. Distinguish between:

1. Homozygous and heterozygous genotypes
2. Dominant and recessive alleles
3. Autosomal and sex-linked inheritance

VIII. Provide the genotypes and phenotypes, of both the parents and offspring, in the following cases:

1. An autosomal gene when:
    - Both parents are heterozygous
    - One parent is heterozygous, and the other is homozygous for the dominant allele
    - One parent is heterozygous, and the other is homozygous for the recessive allele
    - Both parents are homozygous for the recessive allele
    - Both parents are homozygous for the dominant allele
2. A sex-linked gene when:
    - The mother is heterozygous, and the father has the dominant allele
    - The mother is heterozygous, and the father has the recessive allele
    - The mother is homozygous for the dominant allele, and the father has the recessive allele
    - The mother is homozygous for the dominant allele, and the father has the dominant allele
    - The mother is homozygous for the recessive allele, and the father has the recessive allele
    - The mother is homozygous for the recessive allele, and the father has the dominant allele

IX. What is a mutation? Which type of cells would need to mutate for that mutation to be expressed in any offspring?

X. What is a wild-type allele?

XI. Describe the genetic determinants of genetic disorders.

1. Describe the chromosomal abnormality present in, and list the expected karyotype of, an individual with each of the following disorders:
    - Klinefelter syndrome
    - Turner syndrome
    - Down syndrome
2. Describe the type of inheritance demonstrated by each of the following conditions:
    - Tay-Sachs disease
    - Huntington's disease

Part 1: Cell Growth and Division

While there are a few cells in the body that do not undergo cell division (such as gametes, red blood cells, most neurons, and some muscle cells), most somatic cells divide regularly. A **somatic cell** is a general term for a body cell, and all human cells, except for the cells that produce eggs and sperm (which are referred to as germ cells), are somatic cells. Somatic cells contain two copies of each of their chromosomes (one copy received from each parent). A homologous pair of chromosomes is the two versions of a single chromosome found in each somatic cell. The human is a **diploid** organism, having 23 homologous pairs of chromosomes in each of the somatic cells. The condition of having pairs of chromosomes is known as diploidy.

Cells in the body replace themselves over the lifetime of a person. For example, the cells lining the gastrointestinal tract must be frequently replaced when constantly "worn off" by the movement of food through the gut. But what triggers a cell to divide, and how does it prepare for and complete cell division? The **cell cycle** is the sequence of events in the life of the cell from the moment it is created at the end of a previous cycle of cell division until it then divides itself, generating two new cells.

The Cell Cycle: One "turn" or cycle of the cell cycle consists of two general phases: interphase, followed by cell division (mitosis and cytokinesis). **Interphase** is the period of the cell cycle during which the cell is not dividing. The majority of cells are in interphase most of the time. **Mitosis** is the division of genetic material, during which the cell nucleus breaks down and two new, fully functional, nuclei are formed. **Cytokinesis** divides the cytoplasm into two distinctive cells.

Interphase: A cell grows and carries out all normal metabolic functions and processes in a period called G_1 (Figure 1). G_1 phase (gap 1 phase) is the first gap, or growth phase in the cell cycle. For cells that will divide again, G_1 is followed by replication of the DNA, during the S phase. The S phase (synthesis phase) is period during which a cell replicates its DNA.

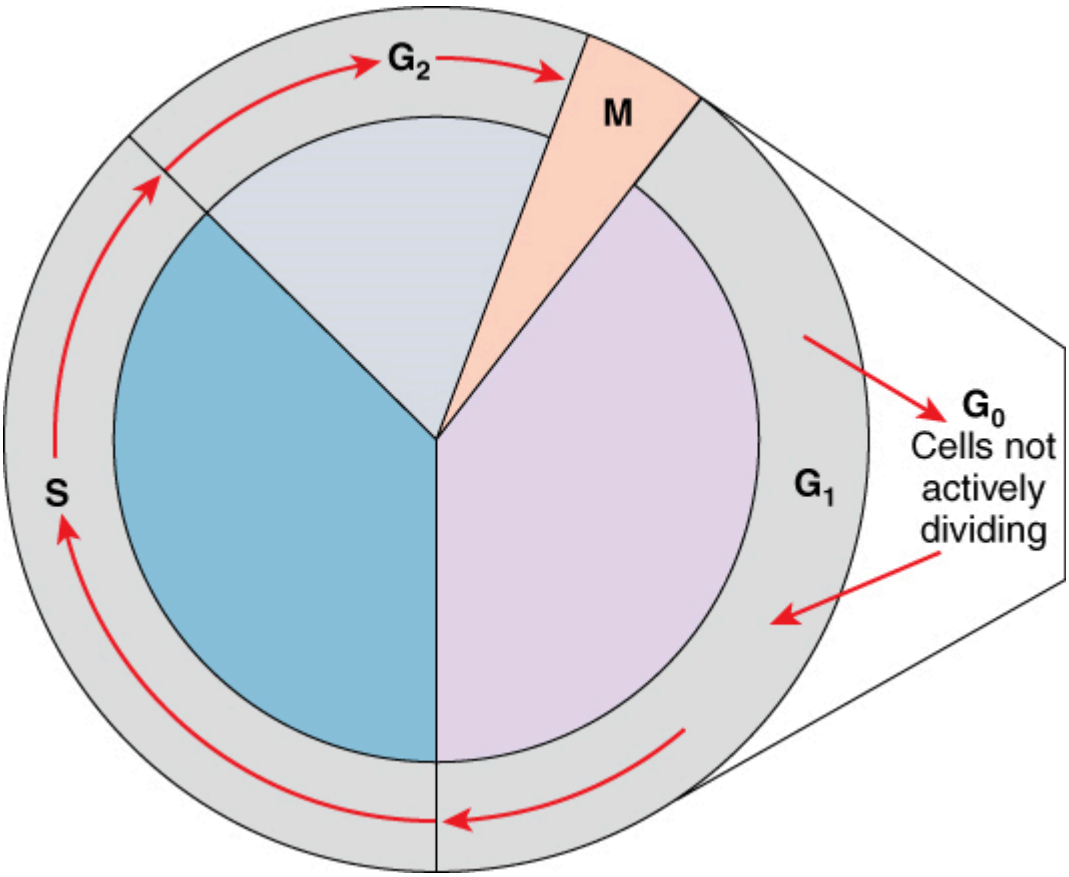


Figure 1. Cell Cycle. *The two major phases of the cell cycle include mitosis (cell division), and interphase, when the cell grows and performs all of its normal functions. Interphase is further subdivided into G1, S, and G2 phases.*

After the synthesis phase, the cell proceeds through the G_2 phase. The **G_2 phase** is a second gap phase, during which the cell continues to grow and makes the necessary preparations for mitosis. Between G_1, S, and G_2 phases, cells will vary the most in their duration of the G_1 phase. It is here that a cell might spend a couple of hours, or many days. The S phase typically lasts between 8-10 hours and the G_2 phase approximately 5 hours. In contrast to these phases, the **G_0 phase** is a resting phase of the cell cycle. Cells that have temporarily stopped dividing and are resting (a common condition) and cells that have permanently ceased dividing (like nerve cells) are said to be in G_0.

The Structure of Chromosomes: Billions of cells in the human body divide every day. During the synthesis phase (S, for DNA synthesis) of interphase, the amount of DNA within the cell precisely doubles. Therefore, after DNA replication but before cell division, each cell actually contains two copies of each chromosome. Each copy of the chromosome is referred to as a **sister chromatid** and is physically bound to the other copy. The **centromere** is the structure that attaches one sister chromatid to another. Because a human cell has 46 chromosomes, during this phase, there are 92 chromatids (46 × 2) in the cell. Make sure not to confuse the concept of a pair of chromatids (one chromosome and its exact copy attached during mitosis) and a homologous pair of chromosomes (two paired chromosomes which were inherited separately, one from each parent) (Figure 2).

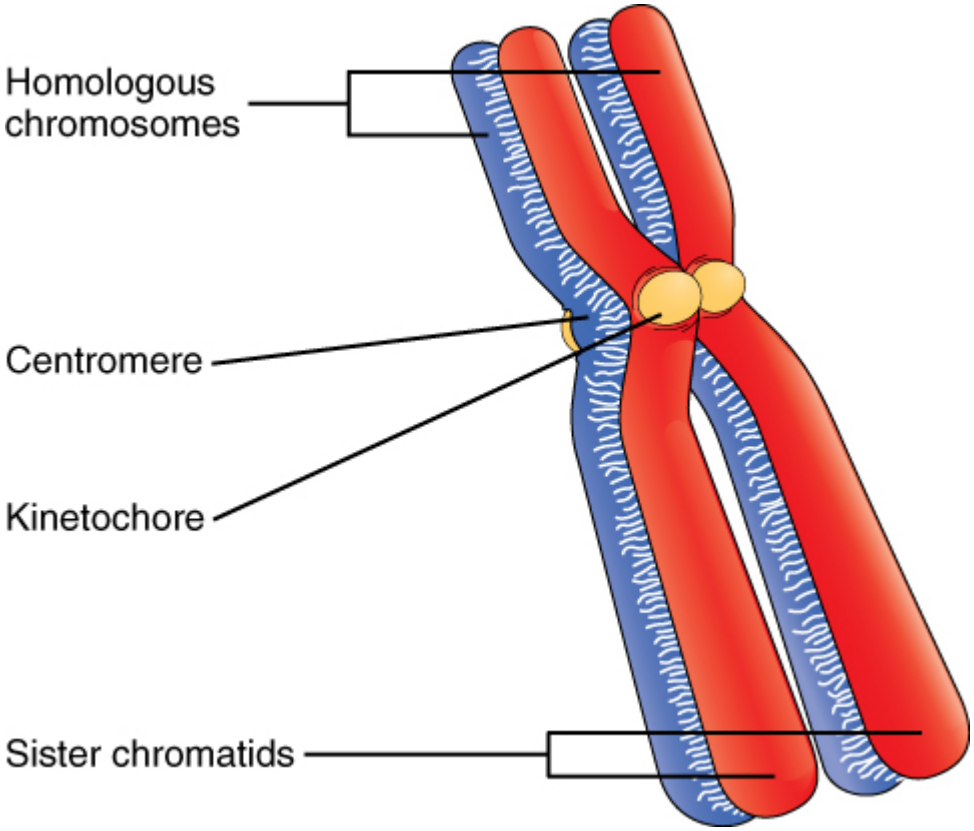


Figure 2. A Homologous Pair of Chromosomes with their Attached Sister Chromatids *(as they appear during meiosis). The red and blue colors correspond to a homologous pair of chromosomes. Each member of the pair was separately inherited from one parent. Each chromosome in the homologous pair is also bound to an identical sister chromatid, which is produced by DNA replication, and results in the familiar "X" shape.*

Mitosis and Cytokinesis: The mitotic phase of the cell typically takes between 1 and 2 hours. During this phase, a cell undergoes two major processes. First, it completes mitosis, during which the contents of the nucleus are equitably pulled apart and distributed between its two halves. Cytokinesis then occurs, dividing the cytoplasm and cell body into two new cells. Mitosis is divided into four major stages that take place after interphase (Figure 3) and in the following order: prophase, metaphase, anaphase, and telophase. The process is then followed by cytokinesis.

Prophase is the first phase of mitosis, during which the loosely packed chromatin coils and condenses into visible chromosomes. During prophase, each chromosome becomes visible forming the familiar X-shape of sister chromatids. The nucleolus disappears early during this phase, and the nuclear envelope also disintegrates. A major occurrence during prophase concerns a very important structure that contains the origin site for microtubule growth. Recall the cellular structures called centrioles that serve as origin points from which microtubules extend. These tiny structures also play a very important role during mitosis. A **centrosome** is a pair of centrioles together. The cell contains two centrosomes side-by-side, which begin to move apart during prophase. As the centrosomes migrate to two different sides of the cell, microtubules begin to extend from each like long fingers from two hands extending toward each other. The **mitotic spindle** is the structure composed of the centrosomes and their emerging microtubules.

Near the end of prophase there is an invasion of the nuclear area by microtubules from the mitotic spindle. The nuclear membrane has disintegrated, and the microtubules attach themselves to the centromeres that adjoin pairs of sister chromatids. The **kinetochore** is a protein structure on the centromere that is the point of attachment between the mitotic spindle and the sister chromatids. This stage is referred to as late prophase or **prometaphase** to indicate the transition between prophase and metaphase.

Metaphase is the second stage of mitosis. During this stage, the sister chromatids, with their attached microtubules, line up along an imaginary linear plane in the middle of the cell, called the **metaphase plate.** The

microtubules are now poised to pull apart the sister chromatids and bring one from each pair to each side of the cell.

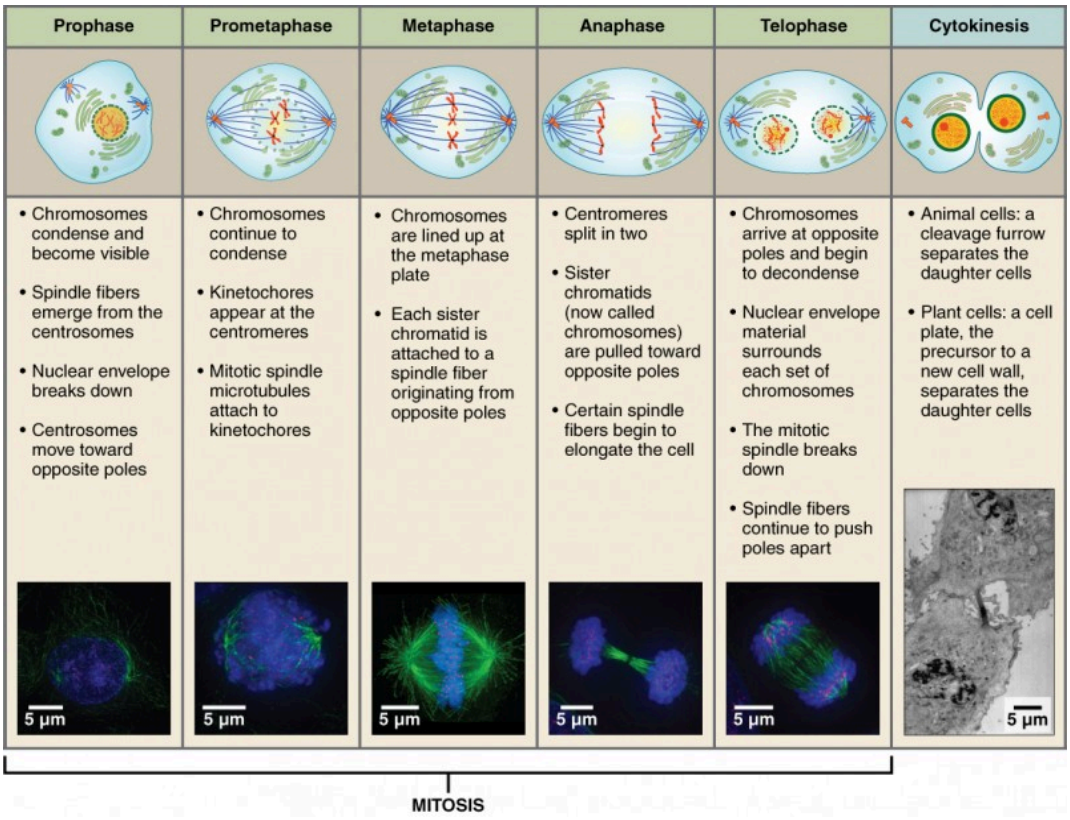


Figure 3. Cell Division: Mitosis Followed by Cytokinesis. *The stages of cell division oversee the separation of identical genetic material into two new nuclei, followed by the division of the cytoplasm.*

Anaphase is the third stage of mitosis. Anaphase takes place over a few minutes, when the pairs of sister chromatids are separated from one another, forming individual chromosomes once again. These chromosomes are pulled to opposite ends of the cell by their kinetochores, as the microtubules shorten. Each end of the cell receives one partner from each pair of sister chromatids, ensuring that the two new daughter cells will contain identical genetic material.

Telophase is the final stage of mitosis. Telophase is characterized by the formation of two new daughter nuclei at either end of the dividing cell. These newly formed nuclei surround the genetic material, which uncoils such that the chromosomes return to loosely packed chromatin. Nucleoli also reappear within the new nuclei, and the mitotic spindle breaks apart, each new cell receiving its own complement of DNA, organelles, membranes, and centrioles. At this point, the cell is already beginning to split in half as cytokinesis begins.

The **cleavage furrow** is a contractile band made up of microfilaments that forms around the midline of the cell during cytokinesis. (Recall that microfilaments consist of actin.) This contractile band squeezes the two cells apart until they finally separate. Two new cells are now formed. One of these cells (the "stem cell") enters its own cell cycle; able to grow and divide again at some future time. The other cell transforms into the functional cell of the tissue, typically replacing an "old" cell there.

Meiosis: Meiosis, unlike mitosis, is not part of the cell cycle of most cells, but only of the germ cells. The daughter cells generated by meiosis are four **haploid** cells are and not genetically identical to the parent cell.

Meiosis is divided into two major stages, meiosis I and meiosis II, that is each further divided into four main stages that are similar to those of mitosis: prophase, metaphase, anaphase, and telophase (Figure 4).

Prophase I is the first phase of meiosis, during which the loosely packed chromatin coils and condenses into visible chromosomes, in a manner similar to prophase of mitosis. In prophase I, however, homologous

chromosomes – chromosomes that contain the same genes – pair together and exchange genetic information with each other. Although pairs of chromosomes contain the same genes, they may contain different variants of those genes known as **alleles**. This process, known as **crossing over**, can occur at many points along a chromosome's length, contributing to genetic variation and resulting in chromosomes that may contain chromatids that are no longer identical to each other.

Metaphase I is the second stage of meiosis. During this stage, the pairs of homologous chromosomes line up along a linear plane in the middle of the cell. Similar to mitosis, the central location where the chromosomes line up is called a metaphase plate. However, unlike mitosis the chromosomes are lined up in pairs. These pairs are arranged in somewhat random orientations relative to each other, in that although they are all lined up at the metaphase plate, the material and paternal chromosomes are not necessarily all on the same side of the plate. This lack of regard to the orientation of other chromosomes results in the independent assortment of maternal and paternal genetic information into separate daughter cells.

Anaphase I is the third stage of meiosis. Microtubules pull entire chromosomes to opposite sides of the cell, while leaving the individual chromatids paired. This results in half as many chromosomes being delivered to either side of the cell as were found in the original parent cell.

Telophase I is the final stage of meiosis I; much like telophase of mitosis, telophase I results in the formation of two new daughter nuclei at either end of the dividing cell, surrounding the genetic material. However, in this case each daughter cell has only half of the number of chromosomes of the parent cell, and may have a different complement of alleles than the parent cell. Each chromosome at this stage still consists of two chromatids that then need to be separated.

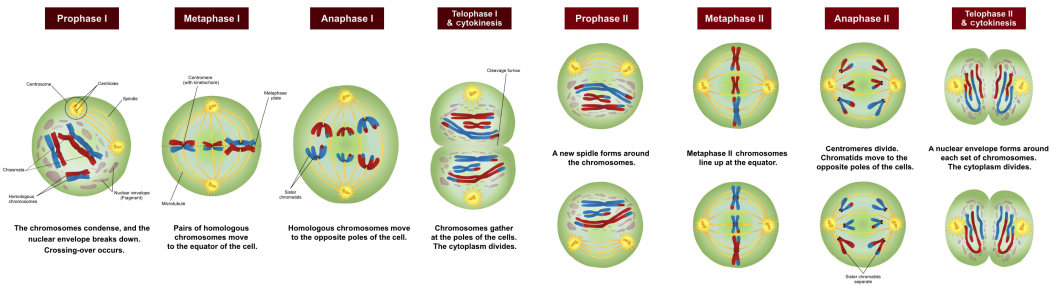


Figure 4. Cell Division: Meiosis. *The stages of cell division oversee the separation of a cell's genetic material into two new nuclei that each contain half of the genetic material of the parent cell. Image credit: Ali Zifan, Wikimedia Commons*

Each of the two cells resulting from meiosis I will therefore need to go through a second round of division known as **meiosis II**. The behaviour of chromosomes in meiosis II is remarkably similar to that of chromosomes during mitosis. However, cells that enter meiosis II have half as many chromosomes as a cell entering mitosis.

Prophase II is the fifth phase of meiosis and the first phase of meiosis II. Again, chromatin is condensed into visible chromosomes and spindle fibers form.

Metaphase II is the second stage of meiosis II. During this stage, the chromosomes line up along the metaphase plate. As in mitosis, the chromosomes are unpaired and simply line up along the central region of the cell.

Anaphase II is the third stage of meiosis II. Microtubules pull the two chromatids of each chromosome to opposite ends of the cell.

Telophase II is the final stage of meiosis. Since the original parent cell produced two cells that then went on to divide a second time, there are now a total of four daughter cells, each having half the genetic material of the original parental cell. Due to crossing-over between chromosomes and the independent assortment of chromosomes that occurred during meiosis, the four resulting daughter cells are likely to be genetically different from each other.

Watch this Crash Course video to learn more about the process of mitosis! Direct link: https://youtu.be/L0k-enzoeOM

Watchthis Amoeba Sisters video to learn more about mitosis! Direct link: https://youtu.be/f-ldPgEfAHI

Watch this Amoeba Sisters video to learn about the process of meiosis! Direct link: a href=https://youtu.be/VzDMG7ke69g>https://youtu.be/VzDMG7ke69g

Part 2: Patterns of Inheritance

We have discussed the events that lead to the development of a newborn. But what makes each newborn unique? The answer lies, of course, in the DNA in the sperm and oocyte that combined to produce that first diploid cell, the human zygote.

From Genotype to Phenotype: Each human body cell has a full complement of DNA stored in 23 pairs of chromosomes that can be organized in a systematic way in an arrangement called a **karyotype** (Figure 5). Among these is one pair of chromosomes, called the sex chromosomes, that determines the sex of the individual (XX in females, XY in males). The remaining 22 chromosome pairs are called autosomal chromosomes. Each of these chromosomes carries hundreds or even thousands of genes, each of which codes for the assembly of a particular protein—that is, genes are "expressed" as proteins. An individual's complete genetic makeup is referred to as his or her genotype. The characteristics that the genes express, whether they are physical, behavioral, or biochemical, are a person's phenotype.

You inherit one chromosome in each pair—a full complement of 23—from each parent. This occurs when the sperm and oocyte combine at the moment of your conception. Homologous chromosomes—those that make up a complementary pair—have genes for the same characteristics in the same location on the chromosome.

Because one copy of a gene, an **allele**, is inherited from each parent, the alleles in these complementary pairs may vary. Take for example an allele that encodes for dimples. A child may inherit the allele encoding for dimples on the chromosome from the father and the allele that encodes for smooth skin (no dimples) on the chromosome from the mother.

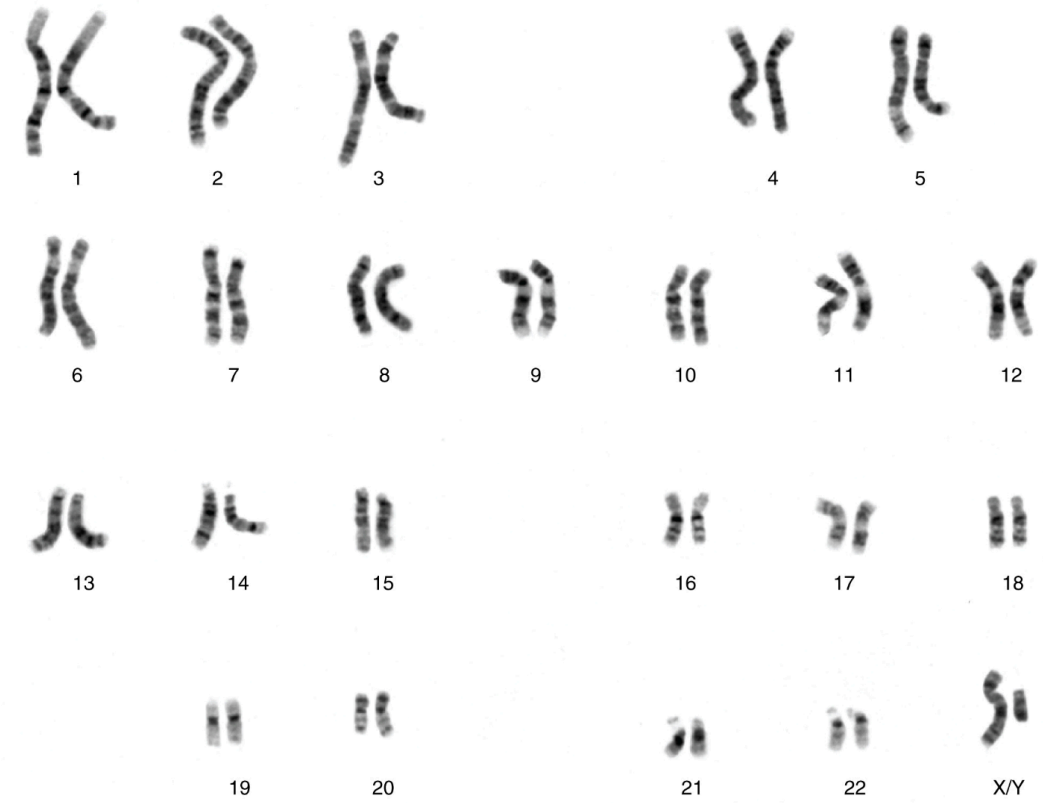


Figure 5. Chromosomal Complement of a Male. *Each pair of chromosomes contains hundreds to thousands of genes. The banding patterns are nearly identical for the two chromosomes within each pair, indicating the same organization of genes. As is visible in this karyotype, the only exception to this is the XY sex chromosome pair in males. (credit: National Human Genome Research Institute)*

Although a person can have two identical alleles for a single gene (a **homozygous** state), it is also possible for a person to have two different alleles (a **heterozygous** state). The two alleles can interact in several different ways. The expression of an allele can be dominant, for which the activity of this gene will mask the expression of a nondominant, or recessive, allele. Sometimes dominance is complete; at other times, it is incomplete. In some cases, both alleles are expressed at the same time in a form of expression known as codominance.

In the simplest scenario, a single pair of genes will determine a single heritable characteristic. However, it is quite common for multiple genes to interact to confer a feature. For instance, eight or more genes—each with their own alleles—determine eye color in humans. Moreover, although any one person can only have two alleles corresponding to a given gene, more than two alleles commonly exist in a population. This phenomenon is called multiple alleles. For example, there are three different alleles that encode ABO blood type; these are designated I^A, I^B, and i.

Over 100 years of theoretical and experimental genetics studies, and the more recent sequencing and annotation of the human genome, have helped scientists to develop a better understanding of how an individual's genotype is expressed as their phenotype. This body of knowledge can help scientists and medical professionals to predict, or at least estimate, some of the features that an offspring will inherit by examining the genotypes or phenotypes of the parents. One important application of this knowledge is to identify an individual's risk for certain heritable genetic disorders. However, most diseases have a multigenic pattern of inheritance and can also be affected by the environment, so examining the genotypes or phenotypes of a person's parents will provide only limited information about the risk of inheriting a disease. Only for a handful of

single-gene disorders can genetic testing allow clinicians to calculate the probability with which a child born to the two parents tested may inherit a specific disease.

Mendel's Theory of Inheritance: Our contemporary understanding of genetics rests on the work of a nineteenth-century monk. Working in the mid-1800s, long before anyone knew about genes or chromosomes, Gregor Mendel discovered that garden peas transmit their physical characteristics to subsequent generations in a discrete and predictable fashion. When he mated, or crossed, two true-breeding (pure-breeding) pea plants that differed by a certain characteristic, the first-generation offspring all looked like one of the parents. For instance, when he crossed tall and dwarf true-breeding pea plants, all of the offspring were tall. Mendel called tallness **dominant** because it was expressed in offspring when it was present in a purebred parent. He called dwarfism **recessive** because it was masked in the offspring if one of the purebred parents possessed the dominant characteristic. Note that tallness and dwarfism are variations on the characteristic of height. Mendel called such a variation a **trait**. We now know that these traits are the expression of different alleles of the gene encoding height.

Mendel performed thousands of crosses in pea plants with differing traits for a variety of characteristics. And he repeatedly came up with the same results—among the traits he studied, one was always dominant, and the other was always recessive. (Remember, however, that this dominant–recessive relationship between alleles is not always the case; some alleles are codominant, and sometimes dominance is incomplete.)

Using his understanding of dominant and recessive traits, Mendel tested whether a recessive trait could be lost altogether in a pea lineage or whether it would resurface in a later generation. By crossing the second-generation offspring of purebred parents with each other, he showed that the latter was true: recessive traits reappeared in third-generation plants in a ratio of 3:1 (three offspring having the dominant trait and one having the recessive trait). Mendel then proposed that characteristics such as height were determined by heritable "factors" that were transmitted, one from each parent, and inherited in pairs by offspring.

In the language of genetics, Mendel's theory applied to humans says that if an individual receives two dominant alleles, one from each parent, the individual's phenotype will express the dominant trait. If an individual receives two recessive alleles, then the recessive trait will be expressed in the phenotype. Individuals who have two identical alleles for a given gene, whether dominant or recessive, are said to be homozygous for that gene (homo- = "same"). Conversely, an individual who has one dominant allele and one recessive allele is said to be heterozygous for that gene (hetero- = "different" or "other"). In this case, the dominant trait will be expressed, and the individual will be phenotypically identical to an individual who possesses two dominant alleles for the trait.

It is common practice in genetics to use capital and lowercase letters to represent dominant and recessive alleles. Using Mendel's pea plants as an example, if a tall pea plant is homozygous, it will possess two tall alleles (*TT*). A dwarf pea plant must be homozygous because its dwarfism can only be expressed when two recessive alleles are present (*tt*). A heterozygous pea plant (*Tt*) would be tall and phenotypically indistinguishable from a tall homozygous pea plant because of the dominant tall allele. Mendel deduced that a 3:1 ratio of dominant to recessive would be produced by the random segregation of heritable factors (genes) when crossing two heterozygous pea plants. In other words, for any given gene, parents are equally likely to pass down either one of their alleles to their offspring in a haploid gamete, and the result will be expressed in a dominant–recessive pattern if both parents are heterozygous for the trait.

Because of the random segregation of gametes, the laws of chance and probability come into play when predicting the likelihood of a given phenotype. Consider a cross between an individual with two dominant alleles for a trait (*AA*) and an individual with two recessive alleles for the same trait (*aa*). All of the parental gametes from the dominant individual would be A, and all of the parental gametes from the recessive individual would be a (Figure 6). All of the offspring of that second generation, inheriting one allele from each parent, would have the genotype *Aa*, and the probability of expressing the phenotype of the dominant allele would be 4 out of 4, or 100 percent.

This seems simple enough, but the inheritance pattern gets interesting when the second-generation *Aa*

individuals are crossed. In this generation, 50 percent of each parent's gametes are *A* and the other 50 percent are a. By Mendel's principle of random segregation, the possible combinations of gametes that the offspring can receive are *AA, Aa, aA* (which is the same as *Aa*), and *aa*. Because segregation and fertilization are random, each offspring has a 25 percent chance of receiving any of these combinations. Therefore, if an *Aa* × *Aa* cross were performed 1000 times, approximately 250 (25 percent) of the offspring would be *AA*; 500 (50 percent) would be *Aa* (that is, *Aa* plus *aA*); and 250 (25 percent) would be *aa*. The genotypic ratio for this inheritance pattern is 1:2:1. However, we have already established that *AA* and *Aa* (and *aA*) individuals all express the dominant trait (i.e., share the same phenotype), and can therefore be combined into one group. The result is Mendel's third-generation phenotype ratio of 3:1.

Mendel's observation of pea plants also included many crosses that involved multiple traits, which prompted him to formulate the principle of independent assortment. The law states that the members of one pair of genes (alleles) from a parent will sort independently from other pairs of genes during the formation of gametes. Applied to pea plants, that means that the alleles associated with the different traits of the plant, such as color, height, or seed type, will sort independently of one another. This holds true except when two alleles happen to be located close to one other on the same chromosome. Independent assortment provides for a great degree of diversity in offspring.

Mendelian genetics represent the fundamentals of inheritance, but there are two important qualifiers to consider when applying Mendel's findings to inheritance studies in humans. First, as we've already noted, not all genes are inherited in a dominant–recessive pattern. Although all diploid individuals have two alleles for every gene, allele pairs may interact to create several types of inheritance patterns, including incomplete dominance and codominance.

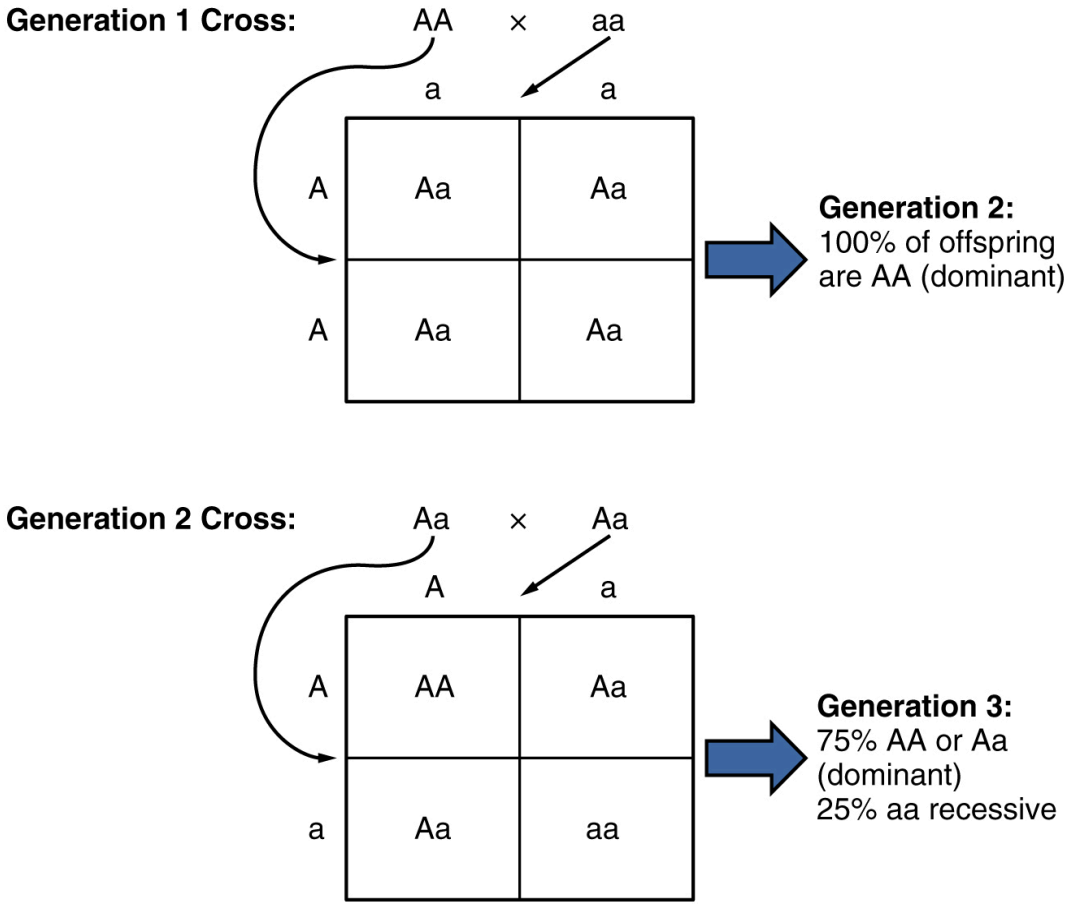


Figure 6. Random Segregation. *In the formation of gametes, it is equally likely that either one of a pair alleles from one parent will be passed on to the offspring. This figure follows the possible combinations of alleles through two generations following a first-generation cross of homozygous dominant and homozygous recessive parents. The recessive phenotype, which is masked in the second generation, has a 1 in 4, or 25 percent, chance of reappearing in the third generation.*

Secondly, Mendel performed his studies using thousands of pea plants. He was able to identify a 3:1 phenotypic ratio in second-generation offspring because his large sample size overcame the influence of variability resulting from chance. In contrast, no human couple has ever had thousands of children. If we know that a man and woman are both heterozygous for a recessive genetic disorder, we would predict that one in every four of their children would be affected by the disease. In real life, however, the influence of chance could change that ratio significantly. For example, if a man and a woman are both heterozygous for cystic fibrosis, a recessive genetic disorder that is expressed only when the individual has two defective alleles, we would expect one in four of their children to have cystic fibrosis. However, it is entirely possible for them to have seven children, none of whom is affected, or for them to have two children, both of whom are affected. For each individual child, the presence or absence of a single gene disorder depends on which alleles that child inherits from his or her parents.

Autosomal Dominant Inheritance: In the case of cystic fibrosis, the disorder is recessive to the normal phenotype. However, a genetic abnormality may be dominant to the normal phenotype. When the dominant allele is located on one of the 22 pairs of autosomes (non-sex chromosomes), we refer to its inheritance pattern as autosomal dominant. An example of an autosomal dominant disorder is neurofibromatosis type I, a disease that induces tumor formation within the nervous system that leads to skin and skeletal deformities. Consider a couple in which one parent is heterozygous for this disorder (and who therefore has neurofibromatosis), *Nn*, and one parent is homozygous for the normal gene, *nn*. The heterozygous parent would have a 50 percent chance of passing the dominant allele for this disorder to his or her offspring, and the homozygous parent would always pass the normal allele. Therefore, four possible offspring genotypes are equally likely to occur: *Nn, Nn, nn*, and *nn*. That is, every child of this couple would have a 50 percent chance of inheriting neurofibromatosis. This inheritance pattern is shown in Figure 7, in a form called a **Punnett square**, named after its creator, the British geneticist Reginald Punnett.

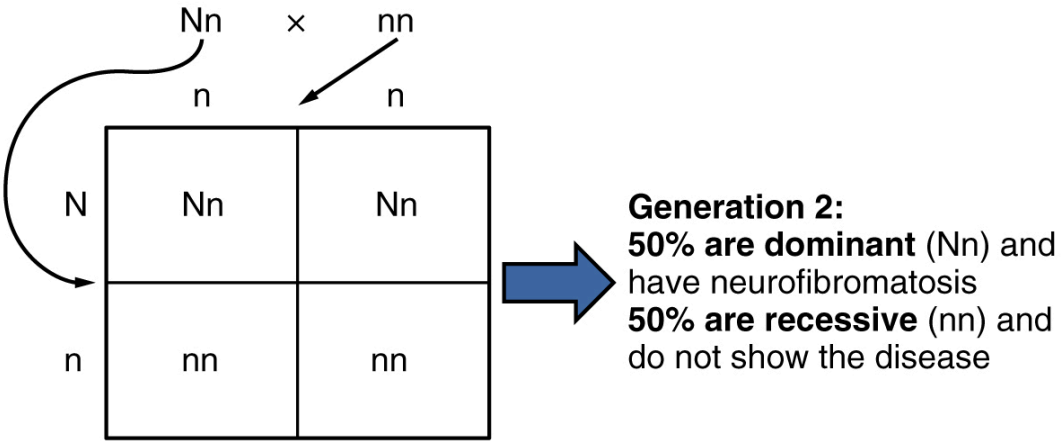


Figure 7. Autosomal Dominant Inheritance. *Inheritance pattern of an autosomal dominant disorder, such as neurofibromatosis, is shown in a Punnett square.*

Other genetic diseases that are inherited in this pattern are achondroplastic dwarfism, Marfan syndrome, and Huntington's disease. Because autosomal dominant disorders are expressed by the presence of just one gene, an individual with the disorder will know that he or she has at least one faulty gene. The expression of the disease may manifest later in life, after the childbearing years, which is the case in Huntington's disease (discussed in more detail later in this section).

Autosomal Recessive Inheritance: When a genetic disorder is inherited in an autosomal recessive pattern, the disorder corresponds to the recessive phenotype. Heterozygous individuals will not display symptoms of this disorder, because their unaffected gene will compensate. Such an individual is called a carrier. Carriers for an autosomal recessive disorder may never know their genotype unless they have a child with the disorder.

An example of an autosomal recessive disorder is cystic fibrosis (CF). CF is characterized by the chronic

accumulation of a thick, tenacious mucus in the lungs and digestive tract. Decades ago, children with CF rarely lived to adulthood. With advances in medical technology, the average lifespan in developed countries has increased into middle adulthood. CF is a relatively common disorder that occurs in approximately 1 in 2000 Caucasians. A child born to two CF carriers would have a 25 percent chance of inheriting the disease. This is the same 3:1 dominant:recessive ratio that Mendel observed in his pea plants would apply here. The pattern is shown in Figure 8, using a diagram that tracks the likely incidence of an autosomal recessive disorder on the basis of parental genotypes.

On the other hand, a child born to a CF carrier and someone with two unaffected alleles would have a 0 percent probability of inheriting CF, but would have a 50 percent chance of being a carrier. Other examples of autosome recessive genetic illnesses include the blood disorder sickle-cell anemia, the fatal neurological disorder Tay–Sachs disease, and the metabolic disorder phenylketonuria.

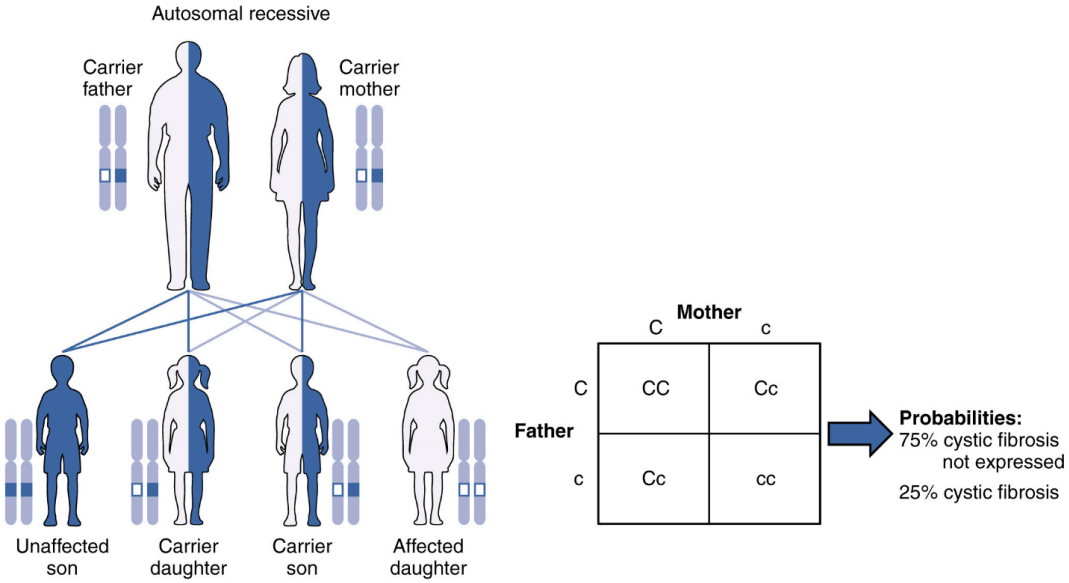


Figure 8. Autosomal Recessive Inheritance. *The inheritance pattern of an autosomal recessive disorder with two carrier parents reflects a 3:1 probability of expression among offspring. (credit: U.S. National Library of Medicine)*

X-linked Dominant or Recessive Inheritance: An X-linked transmission pattern involves genes located on the X chromosome of the 23rd pair (Figure 9). Recall that a male typically has one X and one Y chromosome. When a father transmits a Y chromosome, the child is genetically male, and when he transmits an X chromosome, the child is genetically female. A mother can transmit only an X chromosome, as both her sex chromosomes are X chromosomes.

For genes on either sex chromosome, when examining inheritance it is important to keep track of which chromosome they are on as well as the allele present. When an abnormal allele for a gene that occurs on the X chromosome is dominant over the normal allele, the pattern is described as **X-linked dominant**. Such an allele would be symbolized using a capital letter superscript on a capital X, e.g.: X^A. Thus an otherwise normal individual carrying a dominant allele of the X-linked gene "A" could have any of the following genotypes: X^AX^A (female, homozygous dominant, abnormal phenotype), X^AX^a (female, heterozygous, abnormal phenotype), X^AY (male, abnormal phenotype). Note that for any X-linked gene, males are expected to only have a single allele because they normally only have a single X chromosome, whereas females will be expected to have two alleles – that may be the same or different – because they normally have two X chromosomes.

An example of an X-linked dominent trait is vitamin D–resistant rickets: an affected father would pass the disease allele to all of his daughters, but none of his sons, because he donates only the Y chromosome to his sons (Figure 9a). If it is the mother who is affected, all of her children—male or female—would have a 50 percent chance of inheriting the disorder because she can only pass an X chromosome on to her children (Figure 9b).

For an affected female, the inheritance pattern would be identical to that of an autosomal dominant inheritance pattern in which one parent is heterozygous and the other is homozygous for the normal gene.

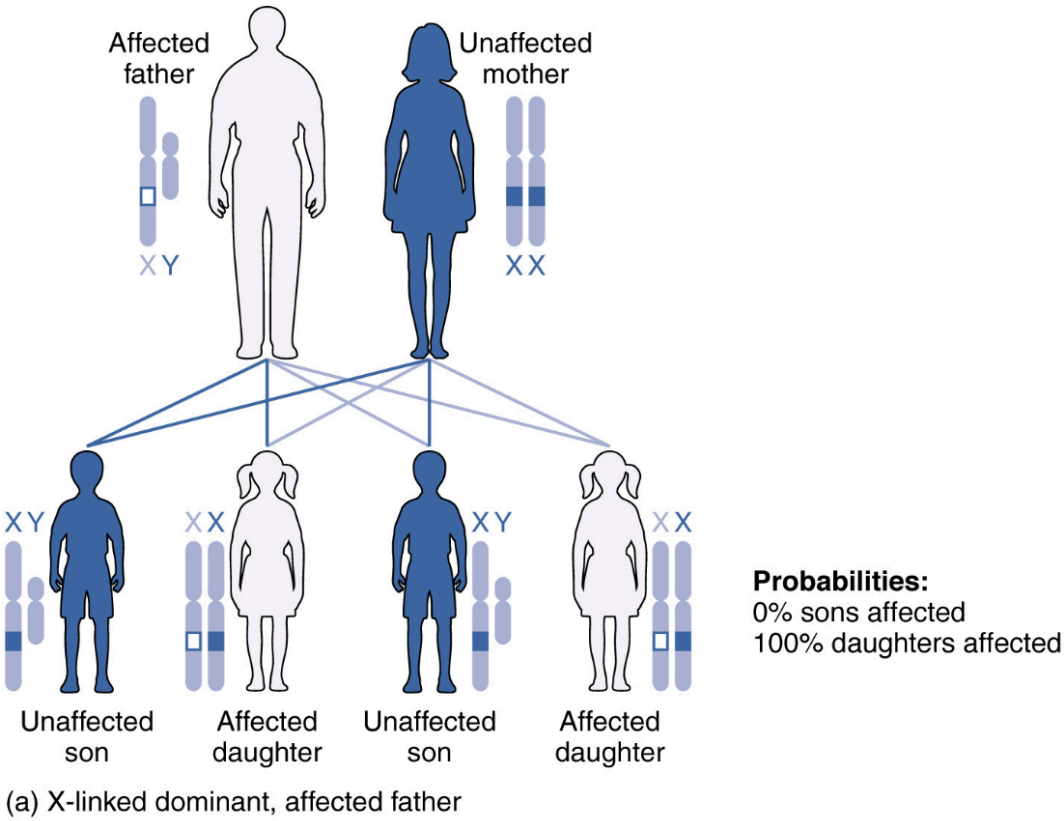


Figure 9. X-Linked Patterns of Inheritance. *A chart of X-linked dominant inheritance patterns differs depending on whether (a) the father or (b) the mother is affected with the disease. (credit: U.S. National Library of Medicine)*

X-linked recessive inheritance is much more common because females can be carriers of the disease yet still have a normal phenotype. This inheritance pattern occurs when an abnormal allele for a gene that occurs on the X chromosome is recessive to the normal allele. Such an allele would be symbolized using a lower-case letter superscript on a capital X, e.g.: X^b. Thus an otherwise normal individual carrying the abnormal allele of the X-linked gene "B" could have any of the following genotypes: X^BX^b (female, heterozygous, normal phenotype), X^bX^b (female, homozygous recessive, abnormal phenotype), X^bY (male, abnormal phenotype). Again, for an X-linked gene males are expected to have only a single allele (on their one X chromosome), whereas females are expected to have two alleles that may be the same or different (one on each of their two X chromosomes).

Diseases transmitted by X-linked recessive inheritance include color blindness, the blood-clotting disorder hemophilia, and some forms of muscular dystrophy. For an example of X-linked recessive inheritance, consider parents in which the mother is an unaffected carrier and the father is normal. None of the daughters would have the disease because they receive a normal gene from their father. However, they have a 50 percent chance of receiving the disease gene from their mother and becoming a carrier. In contrast, 50 percent of the sons would be affected (Figure 10).

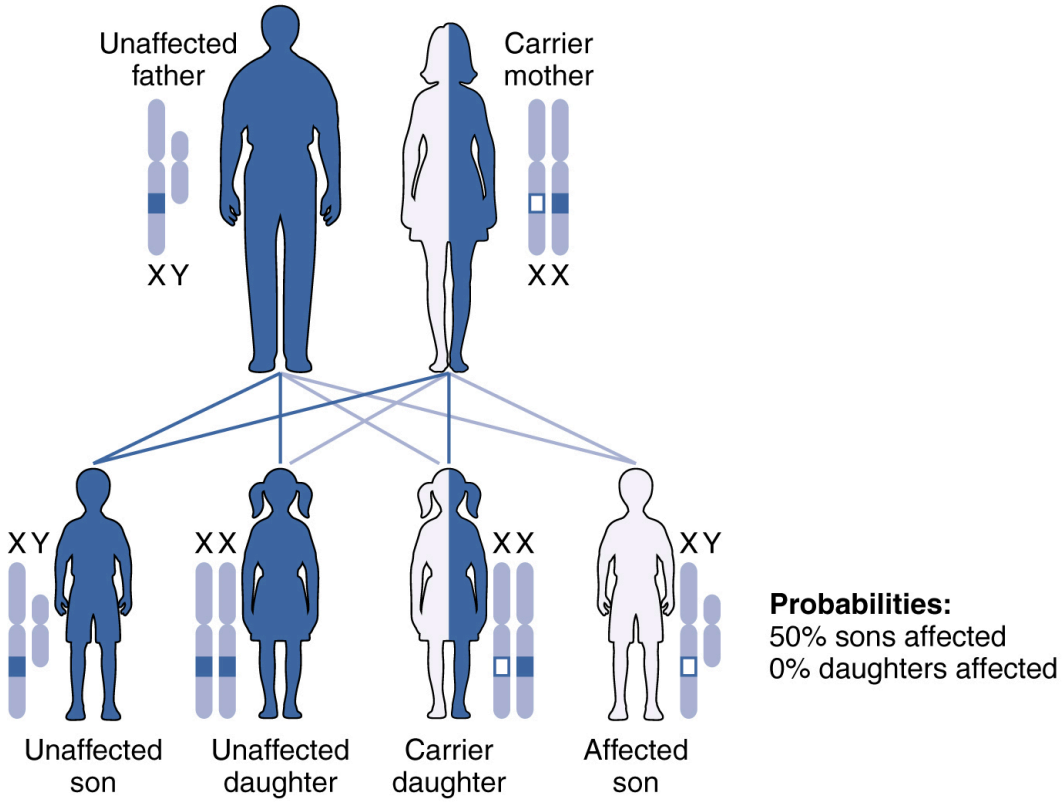


Figure 10. X-Linked Recessive Inheritance. *Given two parents in which the father is normal and the mother is a carrier of an X-linked recessive disorder, a son would have a 50 percent probability of being affected with the disorder, whereas daughters would either be carriers or entirely unaffected. (credit: U.S. National Library of Medicine)*

With X-linked recessive diseases, males either have the disease or are genotypically normal—they cannot be carriers. Females, however, can be genotypically normal, a carrier who is phenotypically normal, or affected with the disease. A daughter can inherit the gene for an X-linked recessive illness when her mother is a carrier or affected, or her father is affected. The daughter will be affected by the disease only if she inherits an X-linked recessive gene from both parents. As you can imagine, X-linked recessive disorders affect many more males than females. For example, color blindness affects at least 1 in 20 males, but only about 1 in 400 females.

Certain combinations of alleles can be lethal, meaning they prevent the individual from developing in utero, or cause a shortened life span. In **recessive lethal** inheritance patterns, a child who is born to two heterozygous (carrier) parents and who inherited the faulty allele from both would not survive. An example of

this is Tay–Sachs, a fatal disorder of the nervous system. In this disorder, parents with one copy of the allele for the disorder are carriers. If they both transmit their abnormal allele, their offspring will develop the disease and will die in childhood, usually before age 5.

Dominant lethal inheritance patterns are much rarer because neither heterozygotes nor homozygotes survive. Of course, dominant lethal alleles that arise naturally through mutation and cause miscarriages or stillbirths are never transmitted to subsequent generations. However, some dominant lethal alleles, such as the allele for Huntington's disease, cause a shortened life span but may not be identified until after the person reaches reproductive age and has children. Huntington's disease causes irreversible nerve cell degeneration and death in 100 percent of affected individuals, but it may not be expressed until the individual reaches middle age. In this way, dominant lethal alleles can be maintained in the human population. Individuals with a family history of Huntington's disease are typically offered genetic counseling, which can help them decide whether or not they wish to be tested for the faulty gene.

Mutations: A **mutation** is a change in the sequence of DNA nucleotides that may or may not affect a person's phenotype. Mutations can arise spontaneously from errors during DNA replication, or they can result from environmental insults such as radiation, certain viruses, or exposure to tobacco smoke or other toxic chemicals. Because genes encode for the assembly of proteins, a mutation in the nucleotide sequence of a gene can change amino acid sequence and, consequently, a protein's structure and function. Spontaneous mutations occurring during meiosis are thought to account for many spontaneous abortions (miscarriages).

Chromosomal Disorders: Sometimes a genetic disease is not caused by a mutation in a gene, but by the presence of an incorrect number of chromosomes. Nondisjunction is the term used in genetics to describe how chromosomes fail to disjoin and move to opposite poles during either Meiosis I or Meiosis II. For example, Down syndrome is caused by having three copies of chromosome 21. This is known as trisomy 21. The most common cause of trisomy 21 is chromosomal nondisjunction during meiosis in the mother. The frequency of nondisjunction events appears to increase with age, so the frequency of bearing a child with Down syndrome increases in women over 36.

Whereas Down syndrome is caused by having three copies of a chromosome, Turner syndrome is caused by having just one copy of the X chromosome. This is known as monosomy. The affected child is considered female. Individuals with Turner syndrome are infertile because their sexual organs do not mature.

Having two copies of the X chromosome and one of the Y is also possible and is known as Klinefelter syndrome. The affected child is genetically male, since the Y chromosome is present, and again is infertile. Individuals are normal intellectually, but the incidence of intellectual disability increases as the number of X chromosomes present increases.

Detecting Genetic Disorders: For many genetic diseases, a DNA test can determine whether a person is a carrier. For instance, carrier status for Fragile X, an X-linked disorder associated with mental retardation, or for cystic fibrosis can be determined with a simple blood draw to obtain DNA for testing. A genetic counselor can educate a couple about the implications of such a test and help them decide whether to undergo testing. For chromosomal disorders, the available testing options include a blood test, amniocentesis (in which amniotic fluid is tested), and chorionic villus sampling (in which tissue from the placenta is tested). Each of these has advantages and drawbacks. A genetic counselor can also help a couple cope with the news that either one or both partners is a carrier of a genetic illness, or that their unborn child has been diagnosed with a chromosomal disorder or other birth defect.